Biomarker Predictors of Clinical Efficacy of the Anti-IgE Biologic, Omalizumab, in Severe Asthma in Adults: Results of the SoMOSA Study

Ratko Djukanović[1], Paul Brinkman[2], Johan Kolmert[3], Cristina Gomez[3] James Schofield[1,4], Joost Brandsma[1,], Andy Shapanis[4], Paul JS Skipp[1,4], Anthony Postle[1], Craig Wheelock[3], Sven-Erik Dahlen[3], Peter J Sterk[2], Thomas Brown[5], David J Jackson[6], Adel Mansur[7], Ian Pavord[8], Mitesh Patel[9], Christopher Brightling[10], Salman Siddiqui[10,23], Peter Bradding[10], Ian Sabroe[11], Dinesh Saralaya[12], Livingstone Chishimba[13], Joanna Porter[14], Douglas Robinson[14], Stephen Fowler[15], Peter H Howarth[1], Louisa Little[16], Thomas Oliver[16], Kayleigh Hill[16], Louise Stanton[16], Alexander Allen[16], Deborah Ellis[16], Gareth Griffiths[16], Tim Harrison[17], Ayobami Akenroye[18,19], Jessica Lasky-Su[19], Liam Heaney[20], Rekha Chaudhuri[21] and Ramesh Kurukulaaratchy[1], on behalf of the SoMOSA study team and the U-BIOPRED study team.

[1] School of Clinical and Experimental Sciences, Faculty of Medicine, University of Southampton and NIHR Southampton Biomedical Research Centre, Southampton, UK.
[2] Dept. Respiratory Medicine, Amsterdam UMC, University of Amsterdam, The Netherlands.
[3] Institute of Environmental Medicine, Karolinska Institutet, and the Department of Respiratory Medicine and Allergy, Karolinska University Hospital, Stockholm, Sweden.
[4] Biological Sciences, University of Southampton, Southampton, UK.
[5] Portsmouth Hospitals University NHS Trust, Queen Alexandra Hospital, Portsmouth, UK.
[6] Guy's Severe Asthma Centre, School of Immunology & Microbial Sciences, King's College London, UK.
[7] University of Birmingham and Heartlands Hospital, University Hospitals Birmingham NHS Foundation Trust, Birmingham, UK.
[8] Oxford Respiratory NIHR BRC, Nuffield Department of Medicine, University of Oxford, Oxford, UK.
[9] Respiratory Medicine and R&D, University Hospitals Plymouth NHS Trust, Plymouth, UK.
[10] Institute for Lung Health and Leicester NIHR Biomedical Research Centre, University of Leicester, Leicester, UK.
[11] Clinical Research Facility, Sheffield Teaching Hospitals NHS Foundation Trust, Sheffield, UK.
[12] Bradford Institute for Health Research and the National Patient Recruitment Centre, Bradford, UK.
[13] Clinical Sciences, Liverpool University Hospitals NHS Foundation Trust, Liverpool, UK.
[14] UCL Respiratory and NIHR University College London Hospitals Biomedical Research Centre, London, UK.
[15] Division of Infection, Immunity and Respiratory Medicine, School of Biological Sciences, The University of Manchester; Manchester Academic Health Science Centre and NIHR Manchester Biomedical Research Centre, Manchester University Hospitals NHS Foundation Trust, Manchester, UK
[16] Southampton Clinical Trials Unit, University of Southampton, and University Hospital Southampton NHS Foundation Trust, Southampton, UK.
[17] Nottingham Respiratory NIHR Biomedical Research Centre, University of Nottingham, Nottingham, UK.
[18] Division of Allergy and Clinical Immunology, Brigham and Women's Hospital, Harvard Medical School, Boston, USA
[19] Channing Division of Network Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, USA
[20] Wellcome-Wolfson Institute for Experimental Medicine, Belfast, Northern Ireland.
[21] Gartnavel General Hospital and School of Infection & Immunity, University of Glasgow, Glasgow, UK.

Corresponding author: Ratko Djukanović, [1]School of Clinical and Experimental Sciences, Faculty of Medicine, University of Southampton and NIHR Southampton Biomedical Research Centre, Southampton, UK.

Word count: 3479

AJRCCM Articles in Press. Published April 18, 2024 as 10.1164/rccm.202310-1730OC

At a glance commentary

Scientific knowledge on the subject: The mechanisms of action of the anti-IgE biologic, omalizumab, in asthma are poorly understood, and commonly measured biomarkers (exhaled nitric oxide, serum IgE, eosinophils) cannot reliably predict the clinical response to treatment. In the age of stratified medicine, the search for reliable ways to predict clinical responses to biologics must be extended to the spectrum of omics biomarkers that have transformed our understanding of the mechanisms of asthma.

What this study adds to the field: This is the first ever study to provide proof-of-concept that omics methods can prospectively identify biomarkers that predict to a high degree whether patients respond to omalizumab, as judged by at least a 50% reduction in acute exacerbations. This study offers a set of volatile organic compounds (VOCs) as the most promising biomarkers for prediction of clinical response and a set of plasma biomarkers for which laboratory methods to measure individual biomarkers would be needed. Prospective studies, comparing clinical responses in patients selected by these biomarkers with those selected according to criteria used in current practice, are needed to validate the candidate biomarkers identified in our study for use in clinical practice.

This article has an online data supplement, which is accessible at the Supplements tab.

AJRCCM Articles in Press. Published April 18, 2024 as 10.1164/rccm.202310-1730OC

Abstract

Background: The anti-IgE monoclonal, omalizumab, is widely used for severe asthma. This study aimed to identify biomarkers that predict clinical improvement during one year of omalizumab treatment.

Methods: 1-year, open-label, **S**tudy **o**f **M**echanisms of action of **O**malizumab in **S**evere **A**sthma (**SoMOSA**) involving 216 severe (GINA step 4/5) uncontrolled atopic asthmatics (≥2 severe exacerbations in previous year) on high-dose inhaled corticosteroids, long-acting β-agonists, ± mOCS. It had two phases: 0-16 weeks, to assess early clinical improvement by Global Evaluation of Therapeutic Effectiveness (GETE), and 16-52 weeks, to assess late responses by ≥50% reduction in exacerbations or dose of maintenance oral corticosteroids (mOCS). All participants provided samples (exhaled breath, blood, sputum, urine) before and after 16 weeks of omalizumab treatment.

Results: 191 patients completed phase 1; 63% had early improvement. Of 173 who completed phase 2, 69% had reduced exacerbations by ≥50%, while 57% (37/65) on mOCS reduced their dose by ≥50%. The primary outcome 2, 3-dinor-11-β-PGF2α, GETE and standard clinical biomarkers (blood and sputum eosinophils, exhaled nitric oxide, serum IgE) did not predict either clinical response. Five volatile organic compounds and 5 plasma lipid biomarkers strongly predicted the ≥50% reduction in exacerbations (receiver operating characteristic area under the curve (AUC): 0.780 and 0.922, respectively) and early responses (AUC:0.835 and 0.949, respectively). In an independent cohort, the GC-MS biomarkers differentiated between severe and mild asthma.

Conclusions: This is the first discovery of omics biomarkers that predict improvement to a biologic for asthma. Their prospective validation and development for clinical use is justified.

AJRCCM Articles in Press. Published April 18, 2024 as 10.1164/rccm.202310-1730OC

Introduction

The anti-IgE monoclonal antibody, omalizumab (Xolair^{R}) is widely used to reduce asthma exacerbations and need for oral corticosteroids (OCS) in severe allergic asthmatics (1-3), but there is no reliable way to predict its benefit. In current practice, patients with at least two severe exacerbations in the past year requiring OCS are given a 16-week therapeutic trial and the response is assessed using the Global Evaluation of Treatment Effectiveness (GETE) (4), a clinical tool based solely on the physician's assessment. GETE-responders are then advised to continue treatment and undergo review after one year of treatment for a reduction in severe acute asthma exacerbations or dose of maintenance OCS (mOCS). While using GETE enriches the responder population (4), a significant proportion of selected patients do not benefit long-term, and there may be GETE non-responders who respond later. Thus, there is an unmet need for predictive biomarkers to optimise the use of omalizumab.

Studies evaluating standard, simple-to-measure clinical biomarkers as predictors of clinical response to omalizumab have had inconsistent results (5); none have assessed biomarker combinations. In order to improve understanding of mechanisms of action of omalizumab and identify predictive biomarkers for clinical practice, we designed a real-world **S**tudy **o**f Mechanisms **o**f action of **O**malizumab in **S**evere **A**sthma (**SoMOSA**). In this article, the focus is on identifying biomarkers that predict which patients improve with treatment. We hypothesised that omics biomarkers (breathomics, proteomics, lipidomics) and urine eicosanoids in readily obtained samples (exhaled breath, blood, sputum, urine) can predict both early responses (using GETE at 16 weeks) and late responses (≥50% reduction in acute exacerbations or mOCS during the first year of treatment), outcomes that are the rationale for prescribing biologics. We measured >1400 omics variables developed by the **U**nbiased **BIO**markers **P**redictive of **RE**spiratory **D**isease outcomes (U-BIOPRED) programme (6, 7), including the prostaglandin D_2 metabolite, 2,3-dinor-11β-PGF2α, and leukotriene E_4 (LTE_4), whose concentrations we have previously found to be lower in severe asthmatics on

AJRCCM Articles in Press. Published April 18, 2024 as 10.1164/rccm.202310-1730OC

omalizumab than in patients receiving standard treatment (8). The predictive value of omics biomarkers was compared with GETE and standard clinical practice biomarkers (FeNO, blood and sputum eosinophils, serum IgE). Evidence for clinical relevance of identified predictive biomarkers was then sought in datasets from two independent cohorts: U-BIOPRED (6, 7) and the Massachusetts General Brigham (MGB) Biobank (9).

Methods

Study design and clinical assessment in the core SoMOSA study

This was an open-label, real-world study; all participants received standard-of-care omalizumab and met current inclusion criteria. After 16 weeks of treatment (study phase 1), patients were evaluated by GETE for early responses. At study end (52 weeks) late responses were defined as ≥50% decrease in asthma exacerbations or dose of mOCS between 16 and 52 weeks of treatment (phase 2). Asthma severity and control were assessed using Asthma Control Questionnaire (ACQ-7), Asthma Control Test (ACT) and Standardised Asthma Quality of Life Questionnaires (AQLQ-S). In contrast to standard practice, patients failing the GETE assessment were also invited to continue treatment in phase 2. The study protocol was approved by the Wales Research Ethics Committee 5, Bangor (15-WA-0302) and patients provided written informed consent.

Two independent cohorts, U-BIOPRED and the MGB Biobank, provided data which served to seek additional clinical value of any identified predictive biomarkers in SoMOSA.

Participants

For the core SoMOSA study, patients from 17 tertiary severe asthma clinics, the inclusion criteria were severe asthma (GINA step 4/5) uncontrolled (ACQ≥1.5, atopic, ≥2 severe asthma exacerbations in past year) despite high-dose ICS and long-acting beta agonists (LABA), ±mOCS, serum total IgE 30-1500 IU/mL, age 18-70 years (see online supplement for complete criteria).

AJRCCM Articles in Press. Published April 18, 2024 as 10.1164/rccm.202310-1730OC

Biomarker datasets from two independent cohorts were identified as suitable for additional analysis of the biomarkers shown in the core SoMOSA cohort as predictive of clinical responses to omalizumab: the U-BIOPRED study (10) and the MGB Biobank ((for details of cohorts and methods, see online supplement).

Standard and omics biomarkers

In the SoMOSA study patients provided exhaled breath, blood, induced sputum and morning urine samples before and after 16 weeks of treatment. Four analytical omics methods, able to quantify large numbers of biomarkers (6, 7), were applied and compared for predictive efficacy with biomarkers often used in clinics (blood and sputum eosinophil counts and FeNO), and with the GETE-based early clinical response tool. Ultra-performance liquid chromatography-tandem mass spectrometry (UPLC-MS/MS) measured urine concentrations of 14 arachidonic acid-derived eicosanoids (11). Exhaled breath was analysed by two methods: 1) Gas Chromatography-Mass Spectrometry (GC-MS) for individual volatile organic compounds (VOCs) and 2) combination of electronic nose (eNose) cross-reactive sensors (12) that produced signatures without VOC identities. Intact lipids in sputum and plasma were measured by ultra-high performance supercritical fluid chromatography–ion mobility-tandem mass spectrometry (UHPSFC-IM-MS/MS) (13). Quantitative data-independent LC/HDMS^{E} was used to measure proteins in sputum and morning urine (7).

The omics methods applied in the U-BIOPRED study were broadly the same as those used in SoMOSA with some technical advances in the latter. Plasma samples from the MGB Biobank underwent global metabolomic profiling (Metabolon, Morrisville, North Carolina, USA) using untargeted liquid chromatography mass spectrometry (LC-MS) platforms which includes amines, amino-acids, and polar and non-polar lipids (14). For more details of U-BIOPRED and MGB Biobank analytical methods, see online supplement.

AJRCCM Articles in Press. Published April 18, 2024 as 10.1164/rccm.202310-1730OC

Power calculation and statistical analysis

The change in urine prostaglandin D_2 metabolite, 2,3-dinor-11β-PGF2α from baseline to 16 weeks post-omalizumab initiation was the selected as the primary outcome and for power calculation using data from a U-BIOPRED study comparing asthmatics on and not on omalizumab (8). Omics biomarkers were prespecified as co-primary outcomes because power calculations are not possible for unbiased omics biomarkers. Assuming 66% of participants would respond (2:1 responder:non-responder ratio), 194 completed participants were required, with sample size adjustment allowed depending on the final responder:non-responder ratio. The same calculation was used to compare exacerbation responders and non-responders. The same participant number was assumed to be required to test the hypothesis that 2,3-dinor-11β-$PGF_{2\alpha}$ in urine would be reduced in participants with ≥50% reduced exacerbations. For more details, see online supplement.

Initial analysis of treatment effects on patient reported outcomes, FEV_1% predicted, FeNO, blood and sputum eosinophils counts applied Analysis of Covariance (ANCOVA) or quantile regression models, depending on the distribution of the data. For the omics analysis, missing values were dealt with as previously described (7, 13), excluding from analysis molecules with detection rates across samples below 40% for proteins and 60% for lipids. Due to differences in methodology between lipidomics and proteomics, missingness was dealt with differently: lipidomics data were imputed using 50% of the lowest limit of detection while for proteomics we used median levels to minimise identification of false positive markers. Data were batch corrected for location, defining GETE and exacerbations as outcomes of interest to preserve variation. Features that detected contaminants due to sample collection and/or processing were removed. Data were then split 50/50 into training and test cohorts; the latter analysed after a final model was produced on the training cohort. Feature selection was performed on the training data. The equal Gini estimator sought to identify the top 5 predictive features for each omics platform data set, which were then used to train the final machine learning prediction model using a random forest algorithm, with 5-fold cross validation repeated 3 times. After

AJRCCM Articles in Press. Published April 18, 2024 as 10.1164/rccm.202310-1730OC

training, the prediction model was tested on the test cohort and results plotted as receiver operating characteristic (ROC) curves. Comparisons of identified predictive biomarkers from the core SoMOSA study were made using the U-BIOPRED and MGB Biobank datasets, using two sample Wilcoxon tests applied to severe and mild to moderate asthmatics in the former and omalizumab responders and non-responders in the latter. Sparse partial least squares discriminant analysis (sPLS-DA) was applied to the U-BIOPRED data set to assess whether those groups of biomarkers identified by random forest analysis to predict clinical responses could differentiate between severe and mild/moderate asthmatics and between patients on and off omalizumab.

RESULTS

Analysis of the SoMOSA study data

Of 811 initially assessed patients, 217 were enrolled; 191 successfully completed phase 1 and 173 completed phase 2, while 43 withdrew (Figure 1, Tables S1, S2). In keeping with the prespecified allowance to adjust the required number of completed patients, recruitment stopped after 191 patients completed phase 1.

Clinical responses

Based on GETE at 16 weeks, 121 of 191 (63%) patients were classified as early responders (Table 1). The majority (n=173, 91%) completed phase 2; of those, 120 (71%) were late responders based on ≥50% acute exacerbation reduction (Table 1) unrelated to age, sex, smoking history or BMI (Table S2). Of 65 patients on mOCS, 37 (57%) reduced the dose by ≥50% without losing asthma control (Table 1). Among early responders not on mOCS, 71.6% also met late-responder criteria; similarly, 70.7% of late-responders not initially on mOCS were also early-responders. Among patients on mOCS at enrolment, 80% of early-responders met the criteria for late-responders, either by acute exacerbations or mOCS use. Taking these two late-response criteria together, 62% of late-responders were also early-responders, while 63% (44 of 70) of early non-responders, who would normally be asked to stop

AJRCCM Articles in Press. Published April 18, 2024 as 10.1164/rccm.202310-1730OC

treatment, were shown in phase 2 to be late-responders, as judged by reduced exacerbations, reduction in mOCS use, or both. Thus, of 36 GETE non-responders not on mOCS, 24 (67%) had a positive response in phase 2. Of the 34 GETE non-responders using mOCS prior to treatment, 20 (59%) had a positive response in phase 2.

The numbers of responders and non-responders, judged by exacerbation reduction (120 and 49, respectively) or GETE (121 and 70, respectively), were deemed sufficient to split the cohort into training and test sets. In contrast, responder by mOCS reduction numbers (37 and 28, respectively) were too small for analysis.

Biomarker measurements

A total of 1408 variables passed quality control. Because individual biomarker molecules can result in multiple mass spectrometry variables that require deconvolution to produce single variables, the 1408 variables were reduced to 14 eicosanoids, 70 breath VOCs, 112 sputum proteins, and 147 urine proteins. A further 158 eNose variables provided signatures without molecular identities. Of the 589 lipid variables in plasma and 305 in sputum, identities were determined only if concentrations were different between responders and non-responders (86 in plasma and 25 in sputum).

Baseline differences in biomarkers between responders and non-responders

Baseline concentrations of 2,3-dinor-11β-$PGF_{2\alpha}$ (primary outcome) did not differentiate early or late-responders and non-responders (Figures 2 and S1). Even though baseline LTE_4 was significantly ($p=0.018$) higher in early responders, LTE_4 and other eicosanoid levels did not differentiate late-responders and non-responders (Figure 2). The same was true for FeNO, blood and sputum eosinophil counts, or IgE (Table 1). In contrast, a total of 368 omics variables were different between responders and non-responders across the four omics platforms (Figure 3): 103, 143 and 122 when comparing

AJRCCM Articles in Press. Published April 18, 2024 as 10.1164/rccm.202310-1730OC

responses by GETE, exacerbation reduction and mOCS use reduction, respectively, 67 being different for more than one outcome.

Prediction of clinical responses to omalizumab by Random Forest analysis

2,3-dinor-11β-$PGF_{2\alpha}$, did not predict either the early GETE-based response (ROC AUC=0.556) or the ≥50% exacerbation reduction during phase 2 (ROC AUC=0.542) (Figure 2), nor did the other urine eicosanoids (data not shown). Similarly, GETE, FeNO, blood or sputum eosinophils, and serum IgE (Figure S2) did not predict exacerbation reductions.

Analysis of all the omics platforms showed that breathomics and plasma lipidomics predicted both early and late responses (Figure 4), while the other omics platforms had weak predictive value (Table S3. One set of 5 exhaled breath VOCs (Benzothiazole, Acetophenone, 2-Pentyl-Furan, Methylene Chloride, 2-Methyl-Butane) predicted early improvement (ROC AUC 0.835); another set of VOCs (2-Ethyl-1-Hexanol, Toluene, 2-Pentene, Nonanal and a VOC of unknown identity, detected as X79.175 by GC-MS) predicted ≥50% exacerbation reduction (ROC AUC 0.780). Two sets of 5 plasma lipids were highly predictive of early and late clinical responses (ROC AUC: 0.949 and 0.922, respectively). The plasma lipids that predicted early responses consisted of four triglycerides (TG(54:6), TG(56:7), TG(55:2) and TG(52:3)) and a currently unidentified lipid. A further set predicted exacerbation reductions: of these, only one could be identified in lipid databases or the wider literature, namely the sphingomyelin peak for SM(d40:2), likely comprising a combination of SM(d18:2/22:0), SM(d16:1/24:1) and SM(d18:1/22:1) molecular species (15). Two further peaks were putatively identified as TG52:3 and ceramide.

Effect of treatment on eicosanoids and standard biomarkers

Urinary 2,3-dinor-11-β-$PGF_{2\alpha}$, decreased significantly (p=0.029) after 16 weeks of treatment, with no difference between responders and non-responders (Figure 2). LTE_4 also reduced (p<0.001) but to a

AJRCCM Articles in Press. Published April 18, 2024 as 10.1164/rccm.202310-1730OC

similar extent in responders and non-responders (Figure 2). The other urine eicosanoids did not change (data not shown).

In the entire cohort, omalizumab reduced blood and sputum eosinophil numbers during phase 1 ($p<0.001$ and 0.023, respectively) and FeNO and blood eosinophils during phase 2 ($p=0.022$, <0.001, respectively), but these changes were not related to treatment responses except for FeNO, which reduced more in early responders ($p=0.014$); however, neither FeNO nor any of the other standard biomarkers discriminated late responders and non-responders by ROC analysis, either in isolation or when combined (Figure S2). We also stratified patients according to FeNO and blood eosinophil count cut-off values used by Hanania and colleagues (2) as biomarker high or low when assessing their clinical response to omalizumab. We found that such stratification did not predict which stratum of patients would respond to omalizumab (Figure S3). Similarly, time to first protocol-defined asthma exacerbation, as demonstrated by Kaplan-Meier curves, was no different (Figure S4) between these strata of patients.

Analysis of the identified predictive biomarkers in the U-BIOPRED and MGB Biobank

A search of the U-BIOPRED data undertaken for matching VOCs and plasma lipids showed that several of the candidate biomarkers that we found predictive of responses to omalizumab were able to differentiate between severe atopic asthmatics and mild/moderate asthmatics (for full details see online supplement) in the U-BIOPRED cohort. In the MGB biobank, the concentrations of plasma sphingomyelin (sphingomyelin (d18:1/22:1, d18:2/22:0, d16:1/24:1) were significantly ($p=0.03$) lower in responders to omalizumab compared to non-responders.

Discussion

To our knowledge, this is the first study to use a multi-omics approach to identify predictive biomarkers for severe asthma, providing proof-of-concept that breathomics and plasma lipidomics

AJRCCM Articles in Press. Published April 18, 2024 as 10.1164/rccm.202310-1730OC

biomarkers can predict who benefits from omalizumab during the first 16 weeks of treatment and with ≥50% reduction in exacerbations during the first year of treatment. In an independent cohort, the biomarkers identified in SoMOSA were shown to differentiate between mild/moderate and severe asthma, including those with more frequent exacerbations who would be candidates for treatment with omalizumab. Development of these biomarkers has significant potential to give patients, their medical teams and payers more certainty of achieving reduced exacerbations with omalizumab, a key objective of asthma treatment.

Consistent with previously reported efficacy, 63% of patients improved within 16 weeks of starting treatment, which suggests that the enrolled cohort is representative of the typical patient considered for omalizumab. In our study, GETE, the clinical tool widely used to assess clinical response to omalizumab, did not predict late improvement (Figure S2); indeed, many patients classified by GETE as non-early responders had a late response (reduced exacerbations or mOCS). Although 2,3-dinor-11B-PGF2a, the co-primary outcome used for power calculation, reduced significantly with treatment, the changes were similar in responders and non-responders and baseline concentrations did not predict either early or late improvement (Figure 2). Similarly, none of the standard biomarkers currently used in asthma management (FeNO, sputum and blood eosinophils and serum IgE) had predictive value (Figure S2).

Breathomics is a growing field in medicine (16). There are several types of eNoses that provide signatures, but not identities of VOCs, and mass spectrometric methods like GC-MS effectively predict clinical and therapeutic outcomes. Whereas the combination of electronic nose (eNose) cross-reactive sensors could not predict clinical improvement, five VOCs (2-ethyl-1-hexanol, toluene, 2-pentene and one unknown VOC), derived by GC-MS, confidently predicted the reductions in exacerbations, while a separate set of five GC-MS-derived VOCs (benzothiazole, acetophenone, 2-pentyl-furan, methylene chloride, and 2-methyl-butane) predicted good early responses. Together, these VOCs differentiated

AJRCCM Articles in Press. Published April 18, 2024 as 10.1164/rccm.202310-1730OC

between mild/moderate asthmatics and all atopic severe asthmatics (ROC AUC 0.931) and between mild/moderate asthma and severe asthmatics prone to exacerbations (≥2 exacerbations per annum), a cut-off for initiating treatment with a biologic. Many of these VOCs have been reported in respiratory studies. Nonanal is associated with neutrophilic asthma and smoking; it has been able to predict exacerbations and discriminate between allergic and non-allergic asthma in children (17-19). Toluene, a common organic solvent, is raised in smokers (20), is related to environmental exposure (21) and has also been associated with asthma (22). We have previously found nonanal within a group of exhaled breath biomarkers in cystic fibrosis patients with sputum positive for *Pseudomonas aeruginosa* (23). The predictive set in our study also included 2-ethyl-1-hexanol for which there is prior evidence of a role in asthma and in lung cancer (reviewed in Sola-Martinet et al. (24)). It is a known indoor pollutant and the main metabolite of di(2-ehylhexyl)phthalate, a solvent and frequent plasticizer of polyvinylchloride (PVC). Concentrations of 2-ethyl-1-hexanol sampled in ambient air are negligible when compared to those in exhaled breath (24), suggesting that, if it is in part inhaled, it is concentrated in the lungs. 2-ethyl-1-hexanol is produced in greater quantities by cancer cells (25). Within the lungs, it acts as an endocrine-disrupting chemical and is associated with oxidative stress and modulation of immune responses (26). The hydrocarbon, 2-pentene, also a solvent and known by-product of thermal cracking of petroleum, is found in ambient air. It is also a volatile derived from lipid-peroxidation, with increased concentrations found by GC-MS in the headspace of bacterial cultures (27). Among the GC-MS variables that predicted early improvement, three have been reported in respiratory conditions: acetophenone in cystic fibrosis patients with *Pseudomonas aeruginosa* (23) and 2-pentylfurane in patients with *Aspergillus fumigatus* (28). Analysis of VOCs in exhaled breath that diagnose ventilator-associated pneumonia has proposed a set of 12 predictive VOCs, among them 2-methyl-butane (29). We could not find any similar reports for benzothiazide.

Lipidomic analysis of plasma also identified two sets of predictive biomarkers. Early improvement was predicted by 4 triglycerides and one unknown lipid species. In comparison to our understanding of the

AJRCCM Articles in Press. Published April 18, 2024 as 10.1164/rccm.202310-1730OC

roles of leukotrienes, knowledge of other lipids in asthma is limited, although obesity is strongly associated with asthma. Serum triglyceride levels are higher in obese people with asthma, even when adjusted for BMI, blood eosinophils, and statin treatment (30). A recent lipidomics study, identifying >1300 plasma lipid species, showed that triglyceride levels, albeit different from the ones in our analysis, differentiated asthma from health and were related to asthma severity (31), with ceramides being related to asthma severity, in keeping with the findings in our study. Ceramide exacerbates inflammation, mucus production and endoplasmic reticulum stress, and increased levels are associated with airway hyperresponsiveness, a key feature of asthma (32). However, these lipids were not good at differentiating between severe and mild/moderate asthma and frequent exacerbators in the U-BIOPRED study; even though concentrations of plasma triglyceride 52:3 and one unidentified lipid were significantly higher in severe atopic asthmatics and in those with ≥2 exacerbations, the ROC AUC indicated weak differentiation (see online supplement). Of note, however, comparison of responders and non-responders to omalizumab (defined by ≥50% reduction in exacerbations) in the MGB cohort showed significantly lower concentrations of sphingomyelin (d18:1/22:1, d18:2/22:0, d16:1/24:1) in responders.

This study has limitations. It could be argued that we should have used a classical randomised controlled trial design, despite ample precedent of similar study design in oncology. Our discussions with the patient advisory group strongly favoured a real-world study design, arguing that a placebo arm would be unethical because it would deny patients a drug known to improve a severe condition, that recruitment into a placebo-controlled trial would be difficult because omalizumab is readily available, and patients expect to be treated. The fact that study recruitment took 26 months and required engagement of 17 severe asthma centres with exclusive rights to prescribe biologics justified this decision. The other limitation of the study is that there were too few patients in whom mOCS treatment was reduced by at least 50%, a measure that is very relevant to patients because of OCS side-effects.

AJRCCM Articles in Press. Published April 18, 2024 as 10.1164/rccm.202310-1730OC

The identified biomarkers should be viewed as candidate biomarkers that require confirmation in a prospective study in which treatment efficacy in patients selected by these biomarkers would be compared with efficacy in patients selected by standard clinical criteria. Further studies are also needed to elucidate how these biomarkers are involved in asthma pathogenesis. Prospective validation of the candidate biomarkers should focus on breathomics, an easy to apply platform, or in combination with plasma lipid measurements. In view of the cost of developing routine analytical methods, the development of single platform assays is likely to be easier, more acceptable to patients, and less expensive. Although lipids had greater predictive power (AUC>0.9) than the VOC biomarkers (AUCs 0.835 for early and 0.780 for late responses), breathomics is, in our view, a superior omics platform because of easier sample collection, more certainty about the VOC identities and, most importantly, easier development of point-of-care instruments for clinical use. Further elucidation of the detected lipids would likely be more complex, costly and with uncertain outcomes, therefore, riskier.

AJRCCM Articles in Press. Published April 18, 2024 as 10.1164/rccm.202310-1730OC

Figure legends

Figure 1. CONSORT diagram

Figure 2. Urine eicosanoids, 2,3-dinor-11-β-$PGF_{2\alpha}$ and LTE_4. (A) Baseline concentrations of 2,3-dinor-11β-PGF2α (primary outcome) and LTE_4 in patients (pg/mL) defined as responders or non-responders based on GETE; (B) Changes in concentrations of 2,3-dinor-11-β-$PGF_{2\alpha}$ and LTE_4 in the entire cohort (responders and non-responders) from baseline to 16 weeks, analysed by Mann-Whitney U test; (C) Receiver Operating Characteristic (ROC) area under the curve (AUC) for 2,3-dinor-11β-PGF2α in respect of prediction of early (GETE-based) and late (acute exacerbation-based) response to omalizumab.

Figure 3. Volcano plots of baseline concentrations of all biomarker variables in responders and non-responders. Responses shown include early response judged by GETE response (3A), late response defined by ≥50% reduction in exacerbations (3B). The red and blue biomarkers (all $p<0.05$) are labelled by numbers (for identities see in Table S4 in the online supplement). Green and red dots represent >1-fold different biomarkers. The data are shown as the means of concentrations in the responders from which the means of the concentrations in the non-responders have been subtracted (i.e. responder minus non-responder). They are shown as log2-transformed data. The p values are obtained by Mann-Whitney U test.

Figure 4. Breath volatile organic compounds (VOC) and plasma lipids that predict early or late clinical responses. The biomarker identities (4A) of the VOCs were derived from the variables detected by gas chromatography-mass spectrometry (GC-MS), while the identities of the plasma lipids were derived from the variables detected by ultra-high performance supercritical fluid chromatography–ion mobility tandem mass spectrometry (UHPSFC-IM-MS/MS). Receiver operating characteristic (ROC) area under the curve (AUC) figures show the prediction by VOCs (4B) and by lipids (4C) of early clinical responses judged by GETE and late responses by reduction in asthma exacerbations. The ROC AUC values for the other omics platforms (sputum lipids, sputum proteins, urine proteins and eicosanoids) are given in table S3.

AJRCCM Articles in Press. Published April 18, 2024 as 10.1164/rccm.202310-1730OC

Tables

	n (%)	ACQ7 (IQR)	ACT (IQR)	AQLQ (IQR)	Ac. Ex. (IQR)	On mOCS n (%)	IgE IU (IQR)	FeNO ppb (IQR)	Blood eos. x10^9/L (IQR)	Sput. eos. % (IQR)
Early responder (16 wk)	121 (63.3)									
Baseline	121	2.9 (2.0-3.6)	12.0 (9.0-17.0)	3.8 (2.9-5.0)	4.0 (3.0-6.0)	41/41 (33.1)	231.0 (114.0-377.0)	33.5 (17.3-59.0)	0.26 (0.11-0.48)	6.0 (0.9 - 23.0)
16 weeks	121	1.6 (1.0-2.1)	18.0 (14.0-21.0)	5.5 (4.5-6.2)	n/a	41/41 (32.2)	n/a	23.5 (12.8-39.0)	0.18 (0.10-0.38)	1.6 (0.5-8.1)
52 weeks	113	1.6 (0.9-2.3)	18.0 (15.0-22.0)	5.7 (4.8-6.3)	1.0 (0.0-2.0)	31/39 (30.6)	n/a	26.0 (16.0-48.5)	0.20 (0.11-0.37)	2.3 (1.0-12.0)
Early non-responder (16 wk)	70 (36.7)									
Baseline	70	2.9 (2.4-3.7)	10.0 (9.0-13.0)	3.7 (2.9-4.7)	4.0 (2.0-6.0)	34/34 (48.6)	157.0 (87.0-302.0)	34.3 (16.0-63.0)	0.23 (0.11-0.41)	2.3 (0.9-13.9)
16 weeks	70	2.6 (2.0-3.3)	13.0 (10.0-16.0)	4.2 (3.4-5.2)	n/a	34/34(47.1)	n/a	29.8 (19.0-52.0)	0.15 (0.08-0.30)	2.3 (0.6-5.6)
52 weeks	60	2.4 (1.1-3.0)	14.5 (10.5-19.0)	4.6 (3.7-5.8)	1.0 (0.0-3.0)	25/29(40.0)	n/a	23.0 (15.0-41.0)	0.17 (0.07-0.38)	2.5 (0.5-10.0)
Ac. ex. Responder (52 wk)	120 (71.0)									
Baseline	120	2.7 (2.0-3.6)	12.0 (9.0-16.0)	3.8 (3.0-5.0)	4.0 (3.0-6.0)	38/38 (31.7)	209.8 (112.5-326.4)	33.0 (16.0-56.0)	0.25 (0.11-0.44)	3.0 (1.0-17.8)
16 weeks	120	1.9 (1.1-2.6)	17.5 (14.0-21.0)	5.2 (4.1-6.1)	n/a	38/38 (30.0)	n/a	24.0 (14.0-37.0)	0.19 (0.10-0.35)	1.8 (0.5-5.8)
52 weeks	120	1.6 (0.9-2.4)	18.5 (13.5-22.0)	5.6 (4.6-6.3)	1.0 (0.0-2.0)	29/38 (30.0)	n/a	22.5 (15.5-42.0)	0.20 (0.10-0.39)	2.3 (0.8-12.3)
Ac. ex. non-responder (52 wk)	49 (29.0)									
Baseline	49	3.0 (2.3-3.7)	10.0 (8.0-13.0)	3.8 (2.9-4.8)	4.0 (2.0-5.0)	26/26 (53.1)	191.7 (98.0-360.0)	31.3 (18.0-53.0)	0.22 (0.11-0.50)	7.8 (1.3-30.5)

AJRCCM Articles in Press. Published April 18, 2024 as 10.1164/rccm.202310-1730OC

16 weeks	49	2.4 (1.6-3.0)	13.0 (11.0-18.0)	4.8 (3.7-5.7)	n/a	26/26 (51.0)	n/a	28.0 (17.5-51.0)	0.12 (0.09-0.30)	2.3 (0.8-8.5)
52 weeks	49	2.4 (1.4-3.0)	14.0 (12.0-18.0)	4.6 (3.9-5.6)	3.0 (2.0-4.0)	24/26 (51.0)	n/a	31.0 (16.5-56.5)	0.18 (0.10-0.38)	2.9 (1.0-6.1)
mOCS responder (52 wk)	37 (56.9)									
Baseline	37	2.4 (2.0-3.6)	13.0 (10.0-19.0)	4.0 (3.2-5.5)	3.0 (2.0-4.0)	37/37 (100)	177.0 (87.0-492.0)	36.0 (19.0-55.5)	0.24 (0.09-0.37)	3.9 (1.5-12.5)
16 weeks	37	2.0 (1.3-2.6)	18.0 (13.0-21.0)	5.1 (4.4-6.1)	n/a	37/37 (100)	n/a	30.0 (17.5-44.0)	0.16 (0.07-0.30)	3.5 (0.5-13.5)
52 weeks	37	2.0 (0.9-2.9)	17.0 (13.0-19.0)	5.2 (4.6-6.1)	1.0 (0.0-2.0)	23/37 (62.2%)	n/a	31.0 (19.5-63.0)	0.25 (0.16-0.42)	7.5 (1.0-21.0)
mOCS non-responder (52 wk)	28 (43.1)									
Baseline	28	3.0 (2.4-3.9)	11.0 (9.0-17.0)	4.0 (3.2-4.6)	4.0 (2.0-6.0)	28/28 (100)	152.0 (101.0-254.0)	30.0 (15.8-72.0)	0.13 (0.06-0.26)	11.0 (0.8-13.0)
16 weeks	28	2.4 (1.1-3.5)	15.5 (10.5-21.5)	4.9 (3.5-6.0)	n/a	28/28 (100)	n/a	27.5 (20.8-61.8)	0.10 (0.04-0.26)	2.6 (1.5-8.5)
52 weeks	28	2.3 (1.4-3.9)	16.0 (10.5-22.0)	4.6 (3.5-6.1)	2.0 (1.0-3.5)	28/28 (100)	n/a	27.0 (15.5-63.5)	0.12 (0.06-0.32)	4.9 (1.3-24.8)
P values (ANCOVA/ quantile regression										
16 wk		<0.001	<0.001	<0.001	n/a	n/a	n/a	0.014	0.180	0.227
52 wk Ac. ex.		0.005	0.002	0.001	n/a	n/a	n/a	0.077	0.099	0.344
52 wk mOCS		0.479	0.853	0.442	n/a	n/a	n/a	0.504	0.166	0.251

Table 1. Demographic and main clinical outcomes

Ac. Ex.: acute exacerbations; mOCS: maintenance oral corticosteroids; IU: international units; ACQ: asthma control questionnaire; AQLQ-S: asthma quality of life questionnaire; ACT: asthma control test; mOCS: maintenance oral corticosteroids; IQR: interquartile range; FeNO: fractional exhaled nitric oxide; ppb: parts per billion; FEV_1: forced expiratory volume in one second; Blood eos: blood

AJRCCM Articles in Press. Published April 18, 2024 as 10.1164/rccm.202310-1730OC

eosinophil counts; Sput. eos.: sputum eosinophil counts. P-values obtained from an Analysis of Covariance (ANCOVA) model or quantile regression model, depending on the distribution of the data. Model: comparison of changes in variable values from baseline to 16 or 52 weeks, with variable after 16/52 weeks of treatment = intercept + response group + variable at baseline. The numbers of participants who enrolled and remained in the study at 16 and 52 weeks are shown in the second column. In the seventh column, the mOCS use data are shown as the numbers of participants on mOCS at the time of assessment as a proportion of the total numbers of patients assessed (i.e., still in the study) at that time-point.

AJRCCM Articles in Press. Published April 18, 2024 as 10.1164/rccm.202310-1730OC

Contributors:

Study design: Ratko Djukanović, Anthony Postle, Craig Wheelock, Peter J Sterk, Paul Skipp, Gareth Griffiths, Louisa Little, Borislav Dimitrov (deceased), Liam Heaney, Rekha Chaudhuri and Peter Bradding.

Study management group: Ratko Djukanovic, Liam Heaney, Rekha Chaudhuri, Tim Harrison, Robert Niven, Tom Stokes, patient member of the SMG, Thomas Oliver.

Principal investigators at study sites: Rekha Chaudhuri, Robin Gore, Liam Heaney, Adel Mansur, Dinesh Saralaya, Ian Pavord, Mitesh Patel (initially Matthew Masoli), Christopher Brightling, David Jackson, Tim Harrison, Thomas Brown, Ian Sabroe, Livingstone Chishimba, Ramesh Kurukulaaratchy (initially Peter H Howarth), Joanna Porter (initially Douglas Robinson), Stephen Fowler (initially Robert Niven).

Study site clinical teams: Vanessa Brown, Alison Clinton, Christine McAnally (Belfast), Rahul Shrimanker, Gareth Hynes, Mary Green (Oxford), Salman Siddiqui, Peter Bradding, Christobelle White, Richard Russell, Michelle Bourne, Beverley Hargadon, Vijay Mistry, Tracey Thornton, Sarah Diver (Leicester), Joanne Kavanagh, Diana Roque (GKT, London), Ben Durham, Linda Kay, Matthew Austin, Ian Smith, Mercy Korley, Sam Anderson, Liam Haslam (Sheffield), Katherine Smith, Helen Smith (Nottingham), Christopher Brereton, Kamran Tariq, Laura Presland, Clair Barber, Jon Ward (Southampton), Jagdeep Sahota, Joel Solis, Akif Khawaja (UCL, London), Valarie Magaya, Sarah Davies, Mary Bellamy, Lesley Horton, Tamika Thompson (Birmingham), Jacqui Galloway, Tom Dymond (Addenbrookes, Cambridge), Matthew Masoli, Julie Anderton, Natasha Wilmhurst, Jennifer Kingdon (Plymouth), Thomas Jones, Jonathon Winter, Scott Elliott (Portsmouth), Daniela Ferreira, Jesús Reiné Esther German, John Blakey, Elena Mitsi, Carla Solorzano, Elissavet Nikolaou, Catherine Lowe (Liverpool), Freda Yang, Steven Smith, Diane Murray, Tracyanne Grandison, Marc Jones (Glasgow).

Trial management team: Gareth Griffiths, Thomas Oliver, Louisa Little, Jessica Rajaram, Nicholas Das, Zina Eminton, Jacqui Nuttall, Susi Renz, Claire Forbes.

AJRCCM Articles in Press. Published April 18, 2024 as 10.1164/rccm.202310-1730OC

Data managements: Alexander Allen, Deborah Ellis, Emma Tilt, Emma Wrixon, Nicola Scott.

Statistical analysis: Kayleigh Hill, Kerensa Thorne, Angeliki Galanopolous, Louise Stanton, Paul Brinkman, James Schofield, Johan Kolmert, Joost Brandsma, Gareth Griffith, Ratko Djukanović and Borislav Dimitrov (deceased).

Massachusetts General Brigham Biobank study: Ayobami Akenroye, Jessica Lasky-Su

Quality control: Kathryn Clark, Keira Fines, Benjamin Sale.

Conflict of Interest statements from all authors:

Ratko Djukanovic discloses shares and consultancy for Synairgen plc, which he has co-founded, speaker honoraria from GlaxoSmithKline, and consultancy for Sanofi and KyMab.

Paul Brinkman discloses to have received financial support from Amsterdam UMC, Stichting Astma Bestrijding (SAB), Vertex Pharmaceuticals, Boehringer Ingelheim, Eurostars and the Horizon Europe Framework Programme.

Johan Kolmert discloses consultancy fees from Gesynta Pharma AB and Lipum AB.

Cristina Gomez has nothing to disclose.

James Schofield's salary and scientific research were funded by Novartis during the period of the study, and he is now founder, director and shareholder in TopMD Precision Medicine Ltd.

Joost Brandsma has nothing to disclose.

Andy Shapanis has nothing to disclose.

Paul JS Skipp discloses shares and consultancy fees in TopMD Precision Medicine Ltd.

Anthony Postle has nothing to disclose.

Craig Wheelock has nothing to disclose.

Sven-Erik Dahlén discloses research grants, consulting fees or lecture honoraria from ASTRAZENECA, Cayman Chemicals, GlaxoSmithKline, Regeneron, Sanofi, and Teva.

Peter J Sterk is scientific advisor for, and has an officially non-substantial interest in, the SME Breathomix that produces eNoses and the accompanying software.

Thomas Brown discloses speaker fees from Astra Zeneca, Glaxo Smith Kline, Sanofi, Teva, Novartis and Chiesi; honoraria for advisory board attendance from Astra Zeneca, Sanofi and Teva; sponsorship to attend international scientific meetings from Sanofi, GLAXOSMITHKLINE, Teva, Chiesi and Napp Pharmaceuticals.

David J Jackson discloses speaker fees and consultancy fees for AstraZeneca, Sanofi and GlaxoSmithKline.

Adel Mansur discloses speaker and advisory board fees and conferences sponsorship from AstraZeneca, GlaxoSmithKline, Sanofi, Boehringer Ingelheim, and Novartis.

AJRCCM Articles in Press. Published April 18, 2024 as 10.1164/rccm.202310-1730OC

Ian Pavord discloses speaker honoraria for speaking at sponsored meetings from AstraZeneca, Boehringer Ingelheim, Aerocrine, Almirall, Novartis, Teva, Chiesi, Sanofi/Regeneron, Menarini and GlaxoSmithKline and payments for organising educational events from AstraZeneca, GlaxoSmithKline, Sanofi/Regeneron and Teva. He has received honoraria for attending advisory panels with Genentech, Sanofi/Regeneron, AstraZeneca, Boehringer Ingelheim, GlaxoSmithKline, Novartis, Teva, Merck, Circassia, Chiesi and Knopp and payments to support FDA approval meetings from GlaxoSmithKline. He has received sponsorship to attend international scientific meetings from Boehringer Ingelheim, GlaxoSmithKline, AstraZeneca, Teva and Chiesi.

Mitesh Patel has nothing to disclose.

Christopher Brightling has received grants and consultancy paid to his Institution from AstraZeneca, GlaxoSmithKline, Regeneron, Sanofi, Genentech, Roche, Chiesi, Areteia, Mologic, 4DPharma.

Salman Siddiqui discloses consultancy fees from GlaxoSmithKline, Astra Zeneca, Chiesi, Areteia therapeutics, CSL Behring, and speaker's honorarium from AstraZeneca, GlaxoSmithKline & Chiesi.

Peter Bradding discloses support to attend scientific meetings from Sanofi and Chiesi, and an honorarium from AstraZeneca.

Ian Sabroe discloses an educational meeting grant from GlaxoSmithKline and a research grant and advisory board fees (paid to his institution) from AstraZeneca.

Dinesh Saralaya has nothing to disclose.

Livingstone Chishimba discloses speaker honorarium for GlaxoSmithKline and support from Chiesi-attending the ERS congress.

Joanna Porter discloses consultancy fees for Istesso and Tacit Fusion and speaker honorarium for The Limbic.

Douglas Robinson has nothing to disclose.

Stephen Fowler has nothing to disclose.

Peter Howarth discloses employment by GlaxoSmithKline.

Louisa Little has nothing to disclose.

Thomas Oliver has nothing to disclose.

Kayleigh Hill has nothing to disclose.

Louise Stanton has nothing to disclose.

Alexander Allen has nothing to disclose.

Deborah Ellis has nothing to disclose.

Gareth Griffiths discloses speaker honorarium at CPD Workshops by AstraZeneca and AbbVie.

Tim Harrison is an employee of AstraZeneca and holds share options in AstraZeneca.

Ayobami Akenroye has nothing to disclose.

Jessica Lasky-Su – disclosure awaited

AJRCCM Articles in Press. Published April 18, 2024 as 10.1164/rccm.202310-1730OC

Liam Heaney has received grant funding, participated in advisory boards and given lectures at meetings supported by Amgen, AstraZeneca, Boehringer Ingelheim, Chiesi, Circassia, Hoffmann la Roche, GlaxoSmithKline, Novartis, Theravance, Evelo Biosciences, Sanofi, and Teva; he has received grants from MedImmune, Novartis UK, Roche/ Genentech Inc, and GlaxoSmithKline, Amgen, Genentech/Hoffman la Roche, AstraZeneca, MedImmune, GlaxoSmithKline, Aerocrine and Vitalograph; he has received sponsorship for attending international scientific meetings from AstraZeneca, Boehringer Ingelheim, Chiesi, GlaxoSmithKline and Napp Pharmaceuticals; he has also taken part in asthma clinical trials sponsored by AstraZeneca, Boehringer Ingelheim, Hoffmann la Roche, and GlaxoSmithKline for which his institution received remuneration; he is the Academic Lead for the Medical Research Council Stratified Medicine UK Consortium in Severe Asthma which involves industrial partnerships with a number of pharmaceutical companies including Amgen, AstraZeneca, Boehringer Ingelheim, GlaxoSmithKline, Hoffmann la Roche, and Janssen.

Rekha Chaudhuri has received lecture fees from GlaxoSmithKline, AstraZeneca, Teva, Chiesi, Sanofi and Novartis; honoraria for Advisory Board Meetings from GlaxoSmithKline, AstraZeneca and Celltrion; sponsorship to attend international scientific meetings from Chiesi, Sanofi and GlaxoSmithKline and a research grant to her Institute from AstraZeneca for a UK multi-centre study.

Ramesh Kurukulaaratchy has nothing to disclose.

Acknowledgments: Funding for this study was from Novartis. RASP-UK (funded by the MRC, provided additional funding for biobanking. We acknowledge the following for their excellent help to coordinate and deliver the study: University Hospital Southampton NHS Foundation Trust R&D and all R&D departments from individual study sites, NIHR Respiratory Translational Research Collaboration for facilitating contracts and for general support of the study's research objectives. We thank the Medical Research Council and the RASP-UK Trial Steering Committee for oversight of the programme during annual reviews. Special thanks to the patient advisory group for their decisive input into the study design.

We wish to pay special tribute to Dr Borislav Dimitrov, a statistician at the University of Southampton, who undertook the power calculations and helped to plan the statistical analysis plan but, sadly, passed away before the end of the study.

REFERENCES

1. Hanania NA, Alpan O, Hamilos DL, Condemi JJ, Reyes-Rivera I, Zhu J, Rosen KE, Eisner MD, Wong DA, Busse W. Omalizumab in severe allergic asthma inadequately controlled with standard therapy: a randomized trial. *Ann Intern Med* 2011; 154: 573-582.
2. Hanania NA, Wenzel S, Rosen K, Hsieh HJ, Mosesova S, Choy DF, Lal P, Arron JR, Harris JM, Busse W. Exploring the effects of omalizumab in allergic asthma: an analysis of biomarkers in the EXTRA study. *Am J Respir Crit Care Med* 2013; 187: 804-811.
3. Bousquet J, Cabrera P, Berkman N, Buhl R, Holgate S, Wenzel S, Fox H, Hedgecock S, Blogg M, Cioppa GD. The effect of treatment with omalizumab, an anti-IgE antibody, on asthma exacerbations and emergency medical visits in patients with severe persistent asthma. *Allergy* 2005; 60: 302-308.
4. Bousquet J, Rabe K, Humbert M, Chung KF, Berger W, Fox H, Ayre G, Chen H, Thomas K, Blogg M, Holgate S. Predicting and evaluating response to omalizumab in patients with severe allergic asthma. *Respir Med* 2007; 101: 1483-1492.
5. Kroes JA, Zielhuis SW, van Roon EN, Ten Brinke A. Prediction of response to biological treatment with monoclonal antibodies in severe asthma. *Biochem Pharmacol* 2020; 179: 113978.
6. Wheelock CE, Goss VM, Balgoma D, Nicholas B, Brandsma J, Skipp PJ, Snowden S, Burg D, D'Amico A, Horvath I, Chaiboonchoe A, Ahmed H, Ballereau S, Rossios C, Chung KF, Montuschi P, Fowler SJ, Adcock IM, Postle AD, Dahlen SE, Rowe A, Sterk PJ, Auffray C, Djukanovic R, Group UBS. Application of 'omics technologies to biomarker discovery in inflammatory lung diseases. *Eur Respir J* 2013; 42: 802-825.
7. Schofield JPR, Burg D, Nicholas B, Strazzeri F, Brandsma J, Staykova D, Folisi C, Bansal AT, Xian Y, Guo Y, Rowe A, Corfield J, Wilson S, Ward J, Lutter R, Shaw DE, Bakke PS, Caruso M, Dahlen SE, Fowler SJ, Horváth I, Howarth P, Krug N, Montuschi P, Sanak M, Sandström T, Sun K, Pandis I, Riley J, Auffray C, De Meulder B, Lefaudeux D, Sousa AR, Adcock IM, Chung KF, Sterk PJ, Skipp PJ, Djukanović R. Stratification of asthma phenotypes by airway proteomic signatures. *J Allergy Clin Immunol* 2019; 144: 70-82.
8. Kolmert J, Gómez C, Balgoma D, Sjödin M, Bood J, Konradsen JR, Ericsson M, Thörngren JO, James A, Mikus M, Sousa AR, Riley JH, Bates S, Bakke PS, Pandis I, Caruso M, Chanez P, Fowler SJ, Geiser T, Howarth P, Horváth I, Krug N, Montuschi P, Sanak M, Behndig A, Shaw DE, Knowles RG, Holweg CTJ, Wheelock Å M, Dahlén B, Nordlund B, Alving K, Hedlin G, Chung KF, Adcock IM, Sterk PJ, Djukanovic R, Dahlén SE, Wheelock CE. Urinary Leukotriene E4 and Prostaglandin D2 Metabolites Increase in Adult and Childhood Severe Asthma Characterized by Type-2 Inflammation. *Am J Respir Crit Care Med* 2020.
9. Boutin NT, Schecter SB, Perez EF, Tchamitchian NS, Cerretani XR, Gainer VS, Lebo MS, Mahanta LM, Karlson EW, Smoller JW. The Evolution of a Large Biobank at Mass General Brigham. *J Pers Med* 2022; 12.
10. Shaw DE, Sousa AR, Fowler SJ, Fleming LJ, Roberts G, Corfield J, Pandis I, Bansal AT, Bel EH, Auffray C, Compton CH, Bisgaard H, Bucchioni E, Caruso M, Chanez P, Dahlen B, Dahlen SE, Dyson K, Frey U, Geiser T, Gerhardsson de Verdier M, Gibeon D, Guo YK, Hashimoto S, Hedlin G, Jeyasingham E, Hekking PP, Higenbottam T, Horvath I, Knox AJ, Krug N, Erpenbeck VJ, Larsson LX, Lazarinis N, Matthews JG, Middelveld R, Montuschi P, Musial J, Myles D, Pahus L, Sandstrom T, Seibold W, Singer F, Strandberg K, Vestbo J, Vissing N, von Garnier C, Adcock IM, Wagers S, Rowe A, Howarth P, Wagener AH, Djukanovic R, Sterk PJ, Chung KF, Group UBS. Clinical and inflammatory characteristics of the European U-BIOPRED adult severe asthma cohort. *Eur Respir J* 2015; 46: 1308-1321.
11. Gómez C, Gonzalez-Riano C, Barbas C, Kolmert J, Hyung Ryu M, Carlsten C, Dahlén SE, Wheelock CE. Quantitative metabolic profiling of urinary eicosanoids for clinical phenotyping. *J Lipid Res* 2019; 60: 1164-1173.

AJRCCM Articles in Press. Published April 18, 2024 as 10.1164/rccm.202310-1730OC

12. Brinkman P, Wagener AH, Hekking PP, Bansal AT, Maitland-van der Zee AH, Wang Y, Weda H, Knobel HH, Vink TJ, Rattray NJ, D'Amico A, Pennazza G, Santonico M, Lefaudeux D, De Meulder B, Auffray C, Bakke PS, Caruso M, Chanez P, Chung KF, Corfield J, Dahlén SE, Djukanovic R, Geiser T, Horvath I, Krug N, Musial J, Sun K, Riley JH, Shaw DE, Sandström T, Sousa AR, Montuschi P, Fowler SJ, Sterk PJ. Identification and prospective stability of electronic nose (eNose)-derived inflammatory phenotypes in patients with severe asthma. *J Allergy Clin Immunol* 2019; 143: 1811-1820.e1817.
13. Brandsma J, Goss VM, Yang X, Bakke PS, Caruso M, Chanez P, Dahlén SE, Fowler SJ, Horvath I, Krug N, Montuschi P, Sanak M, Sandström T, Shaw DE, Chung KF, Singer F, Fleming LJ, Sousa AR, Pandis I, Bansal AT, Sterk PJ, Djukanović R, Postle AD. Lipid phenotyping of lung epithelial lining fluid in healthy human volunteers. *Metabolomics* 2018; 14: 123.
14. Ryals J, Lawton K, Stevens D, Milburn M. Metabolon, Inc. *Pharmacogenomics* 2007; 8: 863-866.
15. Quehenberger O, Armando AM, Brown AH, Milne SB, Myers DS, Merrill AH, Bandyopadhyay S, Jones KN, Kelly S, Shaner RL, Sullards CM, Wang E, Murphy RC, Barkley RM, Leiker TJ, Raetz CR, Guan Z, Laird GM, Six DA, Russell DW, McDonald JG, Subramaniam S, Fahy E, Dennis EA. Lipidomics reveals a remarkable diversity of lipids in human plasma. *J Lipid Res* 2010; 51: 3299-3305.
16. Ibrahim W, Natarajan S, Wilde M, Cordell R, Monks PS, Greening N, Brightling CE, Evans R, Siddiqui S. 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* 2021; 7: 00030-02021.
17. Schleich FN, Zanella D, Stefanuto PH, Bessonov K, Smolinska A, Dallinga JW, Henket M, Paulus V, Guissard F, Graff S, Moermans C, Wouters EFM, Van Steen K, van Schooten FJ, Focant JF, Louis R. Exhaled Volatile Organic Compounds Are Able to Discriminate between Neutrophilic and Eosinophilic Asthma. *Am J Respir Crit Care Med* 2019; 200: 444-453.
18. Caldeira M, Perestrelo R, Barros AS, Bilelo MJ, Morête A, Câmara JS, Rocha SM. Allergic asthma exhaled breath metabolome: a challenge for comprehensive two-dimensional gas chromatography. *J Chromatogr A* 2012; 1254: 87-97.
19. Jareño-Esteban JJ, Muñoz-Lucas M, Carrillo-Aranda B, Maldonado-Sanz J, de Granda-Orive I, Aguilar-Ros A, Civera-Tejuca C, Gutiérrez-Ortega C, Callol-Sánchez LM. Volatile organic compounds in exhaled breath in a healthy population: effect of tobacco smoking. *Arch Bronconeumol* 2013; 49: 457-461.
20. Capone S, Tufariello M, Forleo A, Longo V, Giampetruzzi L, Radogna AV, Casino F, Siciliano P. Chromatographic analysis of VOC patterns in exhaled breath from smokers and nonsmokers. *Biomed Chromatogr* 2018; 32.
21. Maniscalco M, De Laurentiis G, Pentella C, Mormile M, Sanduzzi A, Carratù P, Sofia M. Exhaled breath condensate as matrix for toluene detection: a preliminary study. *Biomarkers* 2006; 11: 233-240.
22. Dragonieri S, Schot R, Mertens BJ, Le Cessie S, Gauw SA, Spanevello A, Resta O, Willard NP, Vink TJ, Rabe KF, Bel EH, Sterk PJ. An electronic nose in the discrimination of patients with asthma and controls. *J Allergy Clin Immunol* 2007; 120: 856-862.
23. Kos R, Brinkman P, Neerincx AH, Paff T, Gerritsen MG, Lammers A, Kraneveld AD, Heijerman HGM, Janssens HM, Davies JC, Majoor CJ, Weersink EJ, Sterk PJ, Haarman EG, Bos LD, Maitland-van der Zee AH. Targeted exhaled breath analysis for detection of Pseudomonas aeruginosa in cystic fibrosis patients. *J Cyst Fibros* 2022; 21: e28-e34.
24. Sola-Martínez RA, Lozano-Terol G, Gallego-Jara J, Morales E, Cantero-Cano E, Sanchez-Solis M, García-Marcos L, Jiménez-Guerrero P, Noguera-Velasco JA, Cánovas Díaz M, de Diego Puente T, Candel-Torralba ME, Garcia-Marcos L, Gimenez-Banon MJ, Martinez-Torres A, Morales E, Perez-Fernandez V, Sanchez-Solis M, Nieto A, Prieto-Sanchez MT, Sanchez-Ferrer M, Fernandez-Palacios L, Gomez-Gomez VP, Martinez-Gracia C, Peso-Echarri P, Ros-Berruezo G, Santaella-Pascual M, Gazquez A, Larque E, Pastor-Fajardo MT, Sanchez-Campillo M, Serrano-

AJRCCM Articles in Press. Published April 18, 2024 as 10.1164/rccm.202310-1730OC

Munuera A, Zornoza-Moreno M, Jimenez-Guerrero P, Adomnei E, Arense-Gonzalo JJ, Mendiola J, Navarro-Lafuente F, Torres-Cantero AM, Salvador-Garcia C, Segovia-Hernández M, Yagüe-Guirao G, Valero-Guillén PL, Aviles-Plaza FV, Cabezas-Herrera J, Martinez-Lopez A, Martinez-Villanueva M, Noguera-Velasco JA, Cantero-Cano E, Franco-Garcia A, Garcia-Serna AM, Hernandez-Caselles T, Martin-Orozco E, Norte-Muñoz M, Cánovas Díaz M, de Diego Puente T, Pastor JM, Sola-Martínez RA, Esteban-Gil A, Fernández-Breis JT, Alcántara MV, Hernández S, López-Soler C, the Nsg. Exhaled volatilome analysis as a useful tool to discriminate asthma with other coexisting atopic diseases in women of childbearing age. *Scientific Reports* 2021; 11: 13823.

25. SPONRING A, FILIPIAK W, MIKOVINY T, AGER C, SCHUBERT J, MIEKISCH W, AMANN A, TROPPMAIR J. Release of Volatile Organic Compounds from the Lung Cancer Cell Line NCI-H2087 In Vitro. *Anticancer Research* 2009; 29: 419-426.
26. Wakayama T, Ito Y, Sakai K, Miyake M, Shibata E, Ohno H, Kamijima M. Comprehensive review of 2-ethyl-1-hexanol as an indoor air pollutant. *J Occup Health* 2019; 61: 19-35.
27. Boots AW, Smolinska A, van Berkel JJBN, Fijten RRR, Stobberingh EE, Boumans MLL, Moonen EJ, Wouters EFM, Dallinga JW, Van Schooten FJ. Identification of microorganisms based on headspace analysis of volatile organic compounds by gas chromatography–mass spectrometry. *Journal of Breath Research* 2014; 8: 027106.
28. Chambers ST, Syhre M, Murdoch DR, McCartin F, Epton MJ. Detection of 2-Pentylfuran in the breath of patients with Aspergillus fumigatus. *Medical Mycology* 2009; 47: 468-476.
29. Schnabel R, Fijten R, Smolinska A, Dallinga J, Boumans M-L, Stobberingh E, Boots A, Roekaerts P, Bergmans D, van Schooten FJ. Analysis of volatile organic compounds in exhaled breath to diagnose ventilator-associated pneumonia. *Scientific Reports* 2015; 5: 17179.
30. van Zelst CM, de Boer GM, Türk Y, van Huisstede A, In't Veen J, Birnie E, Boxma-de Klerk BM, Tramper-Stranders GA, Braunstahl GJ. Association between elevated serum triglycerides and asthma in patients with obesity: An explorative study. *Allergy Asthma Proc* 2021; 42: e71-e76.
31. Jiang T, Dai L, Li P, Zhao J, Wang X, An L, Liu M, Wu S, Wang Y, Peng Y, Sun D, Zheng C, Wang T, Wen X, Cheng Z. Lipid metabolism and identification of biomarkers in asthma by lipidomic analysis. *Biochim Biophys Acta Mol Cell Biol Lipids* 2021; 1866: 158853.
32. Siddique MM, Li Y, Chaurasia B, Kaddai VA, Summers SA. Dihydroceramides: From Bit Players to Lead Actors. *J Biol Chem* 2015; 290: 15371-15379.

AJRCCM Articles in Press. Published April 18, 2024 as 10.1164/rccm.202310-1730OC

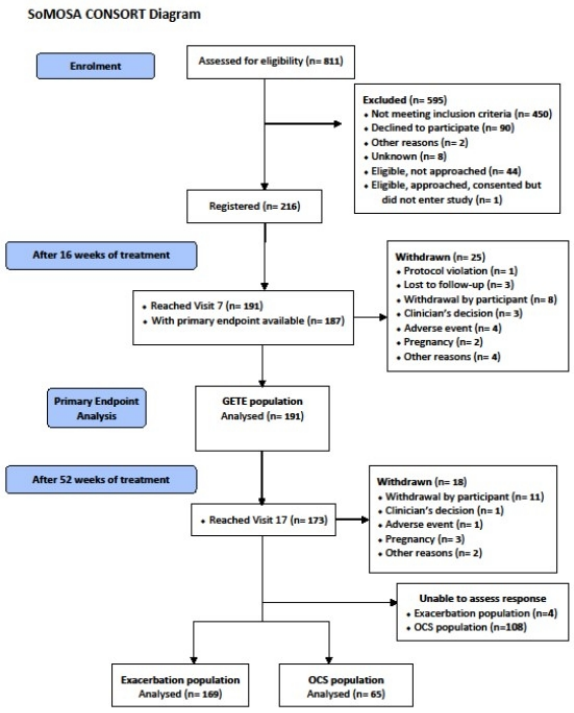


Fig 1

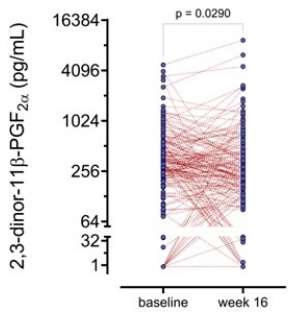
Figure 1. CONSORT diagram

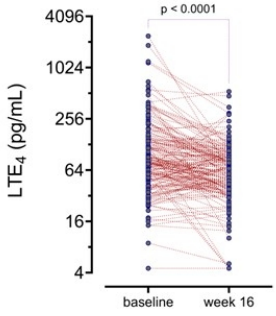
254x190mm (96 x 96 DPI)

AJRCCM Articles in Press. Published April 18, 2024 as 10.1164/rccm.202310-1730OC

A

Clinical outcome	Urine Eicosanoid Metabolite	Responder Median (IQR)	Non-Responder Median (IQR)	*p*-value
GETE	2,3-dinor-11β-$PGF_{2α}$	325 (206-554)	323 (177-531)	0.827
GETE	LTE_4	88 (57-159)	71 (40-112)	0.018
Exacerbation	2,3-dinor-11β-$PGF_{2α}$	321 (199-553)	296 (185-557)	0.929
Exacerbation	LTE_4	81 (51-137)	79 (47-138)	0.640
OCS	2,3-dinor-11β-$PGF_{2α}$	342 (210-644)	242 (105-555)	0.350
OCS	LTE_4	75 (52-186)	68 (40-101)	0.156





C

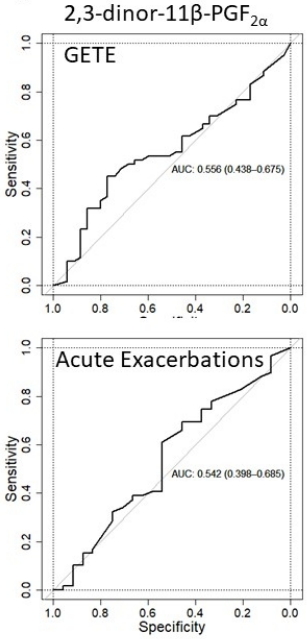


Fig 2

Figure 2. Urine eicosanoids, 2,3-dinor-11-β-PGF2α and LTE4. (A) Baseline concentrations of 2,3-dinor-11β-PGF2α (primary outcome) and LTE4 in patients (pg/mL) defined as responders or non-responders based on GETE; (B) Changes in concentrations of 2,3-dinor-11-β-PGF2α and LTE4 in the entire cohort (responders and non-responders) from baseline to 16 weeks, analysed by Mann-Whitney U test; (C) Receiver Operating Characteristic (ROC) area under the curve (AUC) for 2,3-dinor-11β-PGF2α in respect of prediction of early (GETE-based) and late (acute exacerbation-based) response to omalizumab.

254x190mm (96 x 96 DPI)

AJRCCM Articles in Press. Published April 18, 2024 as 10.1164/rccm.202310-1730OC

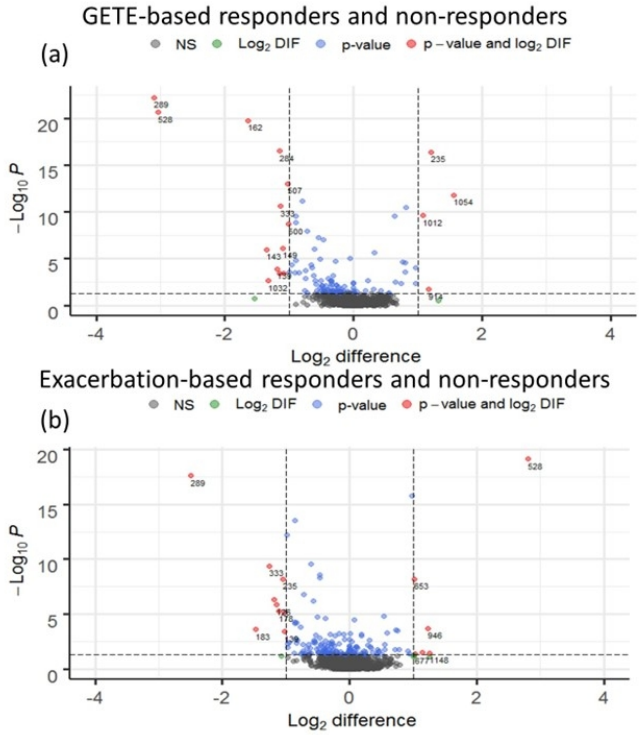


Fig 3

Figure 3. Volcano plots of baseline concentrations of all biomarker variables in responders and non-responders. Responses shown include early response judged by GETE response (3A), late response defined by ≥50% reduction in exacerbations (3B). The red and blue biomarkers (all $p<0.05$) are labelled by numbers (for identities see in Table S4 in the online supplement). Green and red dots represent >1-fold different biomarkers. The data are shown as the means of concentrations in the responders from which the means of the concentrations in the non-responders have been subtracted (i.e. responder minus non-responder). They are shown as log2-transformed data. The p values are obtained by Mann-Whitney U test.

254x190mm (96 x 96 DPI)

AJRCCM Articles in Press. Published April 18, 2024 as 10.1164/rccm.202310-1730OC

	GETE		Acute exacerbations	
Biomarker	Variable	Identity	Variable	Identity
VOCs	X103.1157	Benzothiazole	X55.1.902	2-Ethyl-1-Hexanol
	X50.1.977	Acetophenone	X63.539	Toluene
	X109.7.861	2-Pentyl-Furan	X79.175	unknown
	X35.157	Methylene Chloride	X49.328	2-Pentene
	X39.2.265	2-Methyl-Butane	X55.1.999	Nonanal
Plasma lipids	PlP_1.20_878.8511_n	TG(54:6)	PlP_1.34_467.4848_mz	unknown
	PlP_1.22_881.9598_n	Unknown	PlP_2.96_879.9375_mz	TG 52:3?
	PlP_2.54_927.8584_mz	TG(56:7)	PlP_3.25_548.6240_mz	Ceramide
	PlP_2.62_918.8734_mz	TG(55:2)	PlP_3.47_368.4908_mz	unknown
	PlP_2.96_879.9375_mz	TG 52:3?	PlP_3.76_784.7540_n	SM(d40:2)

Fig 4A

Figure 4. Breath volatile organic compounds (VOC) and plasma lipids that predict early or late clinical responses. The biomarker identities (4A) of the VOCs were derived from the variables detected by gas chromatography-mass spectrometry (GC-MS), while the identities of the plasma lipids were derived from the variables detected by ultra-high performance supercritical fluid chromatography–ion mobility tandem mass spectrometry (UHPSFC-IM-MS/MS). Receiver operating characteristic (ROC) area under the curve (AUC) figures show the prediction by VOCs (4B) and by lipids (4C) of early clinical responses judged by GETE and late responses by reduction in asthma exacerbations. The ROC AUC values for the other omics platforms (sputum lipids, sputum proteins, urine proteins and eicosanoids) are given in table S3.

254x190mm (96 x 96 DPI)

AJRCCM Articles in Press. Published April 18, 2024 as 10.1164/rccm.202310-1730OC

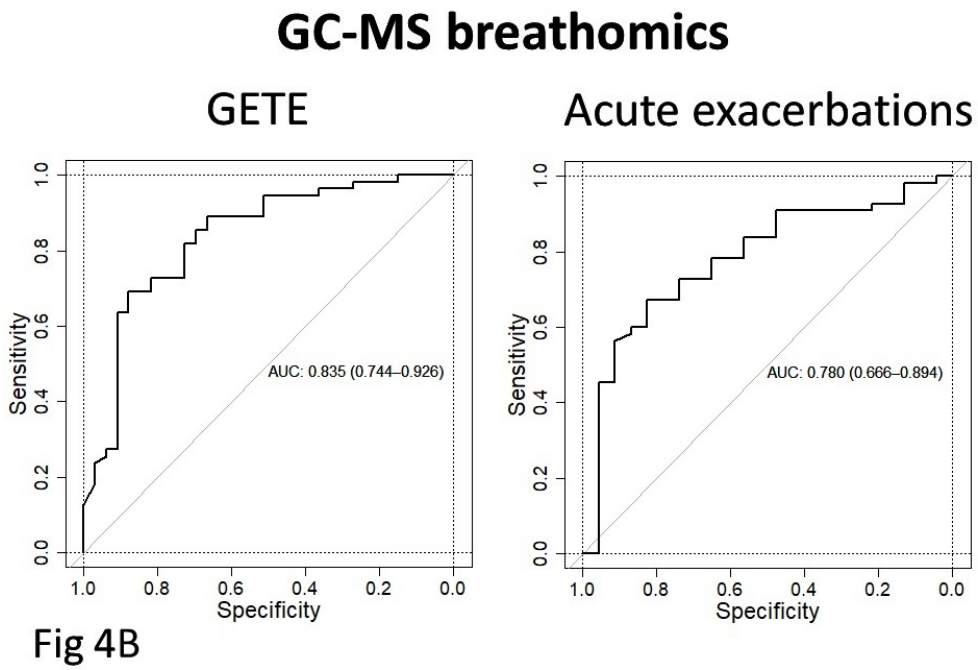


Figure 4. Breath volatile organic compounds (VOC) and plasma lipids that predict early or late clinical responses. The biomarker identities (4A) of the VOCs were derived from the variables detected by gas chromatography-mass spectrometry (GC-MS), while the identities of the plasma lipids were derived from the variables detected by ultra-high performance supercritical fluid chromatography–ion mobility tandem mass spectrometry (UHPSFC-IM-MS/MS). Receiver operating characteristic (ROC) area under the curve (AUC) figures show the prediction by VOCs (4B) and by lipids (4C) of early clinical responses judged by GETE and late responses by reduction in asthma exacerbations. The ROC AUC values for the other omics platforms (sputum lipids, sputum proteins, urine proteins and eicosanoids) are given in table S3.

254x190mm (96 x 96 DPI)

AJRCCM Articles in Press. Published April 18, 2024 as 10.1164/rccm.202310-1730OC

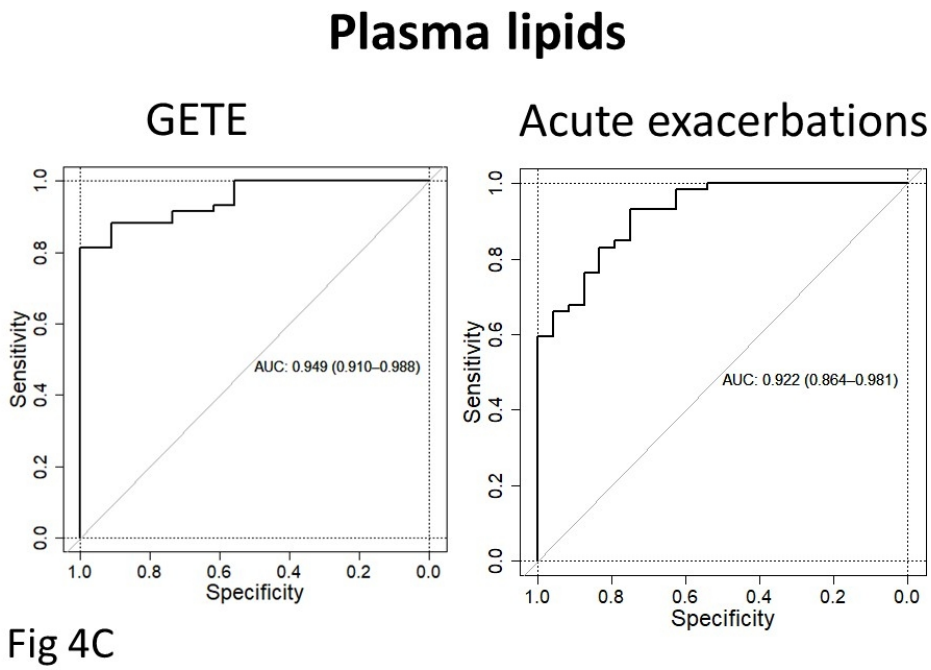


Figure 4. Breath volatile organic compounds (VOC) and plasma lipids that predict early or late clinical responses. The biomarker identities (4A) of the VOCs were derived from the variables detected by gas chromatography-mass spectrometry (GC-MS), while the identities of the plasma lipids were derived from the variables detected by ultra-high performance supercritical fluid chromatography–ion mobility tandem mass spectrometry (UHPSFC-IM-MS/MS). Receiver operating characteristic (ROC) area under the curve (AUC) figures show the prediction by VOCs (4B) and by lipids (4C) of early clinical responses judged by GETE and late responses by reduction in asthma exacerbations. The ROC AUC values for the other omics platforms (sputum lipids, sputum proteins, urine proteins and eicosanoids) are given in table S3.

254x190mm (96 x 96 DPI)

AJRCCM Articles in Press. Published April 18, 2024 as 10.1164/rccm.202310-1730OC

Biomarker Predictors of Clinical Efficacy of the Anti-IgE Biologic, Omalizumab, in Severe Asthma in Adults: Results of the SoMOSA Study

Ratko Djukanović, Paul Brinkman, Johan Kolmert, Cristina Gomez, James Schofield, Joost Brandsma, Andy Shapanis, Paul JS Skipp, Anthony Postle, Craig Wheelock, Sven-Erik Dahlen, Peter J Sterk, Thomas Brown, David J Jackson, Adel Mansur, Ian Pavord, Mitesh Patel, Christopher Brightling, Salman Siddiqui, Peter Bradding, Ian Sabroe, Dinesh Saralaya, Livingstone Chishimba, Joanna Porter, Douglas Robinson, Stephen Fowler, Peter H Howarth, Louisa Little, Thomas Oliver, Kayleigh Hill, Louise Stanton, Alexander Allen, Deborah Ellis, Gareth Griffiths, Tim Harrison, Ayobami Akenroye, Jessica Lasky-Su, Liam Heaney, Rekha Chaudhuri and Ramesh Kurukulaaratchy, on behalf of the SoMOSA study team and the U-BIOPRED study team

ONLINE DATA SUPPLEMENT

AJRCCM Articles in Press. Published April 18, 2024 as 10.1164/rccm.202310-1730OC

Online supplement

Patient involvement in the study

Patients with severe asthma engaged in the study as an advisory group from the beginning of the programme of the Refractory asthma Stratification programme (RASP-UK) within which the SoMOSA study was one of the work-packages. Their discussions were crucial in respect of the inclusion criteria and the open-label, real-life design.

Tom Stokes, the patient member of the Study Management Group that led and coordinated the study delivery submits the following statement:

I would like to confirm that this trial was conducted at all times with the patient welfare foremost in the thoughts and actions of the trial. I was consulted at all times as to what I thought of the procedures, and if any obstacle was reached, I was privy as to how they were dealt with and asked if I would be happy as a Patient. I had complete confidence that the Patient welfare was always of paramount importance.

Tom Stokes

Inclusion Criteria:

1. Severe uncontrolled asthma (GINA step 4 and 5) despite daily treatment with high-dose inhaled corticosteroids (ICS) and long-acting beta agonists (LABA). (High-dose ICS will be a minimum twice daily dose of 800 mcg of beclomethasone dipropionate equivalent inhaler for at least 8 weeks before screening). Participants will need to fulfil the criteria for uncontrolled asthma as judged by their Asthma Control Questionnaire (ACQ) score ≥1.5 during the screening period.
2. Participants on maintenance treatment with oral corticosteroids will also be included and will also have to meet the same ACQ inclusion criterion (ACQ≥1.5).
3. Atopic, as identified by positive skin prick test or in vitro reactivity to a perennial aeroallergen.
4. Two or more documented severe asthma exacerbations within the previous 12 months that require courses of prednisolone; defined as increased asthma symptoms requiring treatment in the community or in hospital with systemic corticosteroid rescue therapy or an increase in daily oral corticosteroids for patients already on maintenance oral corticosteroids for >2 months.
5. Frequent daytime symptoms or night-time awakenings.
6. IgE level of 30 to 1500 IU/mL
7. Body weight less than 150 kg
8. Age 18-70
9. Able to give written informed consent prior to participation in the study, which includes ability to comply with the requirements and restrictions listed in the consent form.
10. Able to read, comprehend, and write at a sufficient level to complete study related materials.

Exclusion Criteria:

1. An exacerbation requiring treatment with systemic corticosteroids (or an increase in the baseline dose of OCS) within the 30 days before screening.
2. Active lung disease other than asthma.
3. Treatment with Xolair or another biologic in the 12 months before screening.
4. Elevated serum IgE levels for reasons other than allergy (for example, parasite infections, the hyperimmunoglobulin E syndrome, the Wiskott–Aldrich syndrome, or bronchopulmonary aspergillosis).
5. The following medication is not allowed during the run-in and treatment period and should not have been taken for at least 3 months prior to screening: methotrexate, cyclosporine, intravenous immunoglobulin or immunosuppressants.
6. Current smoker. Smoking in the past year is also an exclusion. Ex-smokers will have to be confirmed by a negative cotinine test. If there is a history of smoking for >10 pack years, then asthma diagnosis should have been made before the age of 40 and objective evidence of reversibility of FEV1>12% and 200ml should be available [either previously recorded or done as part of screening for this study]. Patients where an asthma/COPD overlap is suspected should not be included.
7. The participant has a history of current recreational drug use or other allergy, which, in the opinion of the responsible physician, contra-indicates their participation.
8. Female patient who is pregnant or lactating or up to 6 weeks post- partum or 6 weeks cessation of breast feeding.

AJRCCM Articles in Press. Published April 18, 2024 as 10.1164/rccm.202310-1730OC

9. Those participants who, in the opinion of the investigator, have a risk of non-compliance with study procedures.
10. The subject has a recent history of incapacitating psychiatric disorders.
11. History or current evidence of an upper or lower respiratory infection or symptoms (including common cold) within 4 weeks of baseline assessments (in such participant assessments should be deferred until after 4 weeks have lapsed from the infection).

Data collection

Data collection, storage and quality control were conducted following the standard operating procedures of the Southampton Clinical Trials Unit, using the Medidata RAVE electronic data capture platform.

Power calculation and statistical analysis

The change from baseline to 16 weeks post-omalizumab initiation in urine concentrations of the prostaglandin D2 metabolite, 2,3-dinor-11β-PGF2α, was selected for power calculation/sample size estimation using data from a previous U-BIOPRED study comparing severe asthmatics on and those not on omalizumab (8). For details of power calculation see online supplement. Omics biomarker levels were treated as co-primary outcomes because standard power calculation was not possible at this stage of biomarker development. Power calculation was performed in nQuery v7.0, assuming a mean (SD) concentration of 70.3 (40.9) ng/mmol creatinine at baseline and 49 (23) ng/mmol creatinine in GETE responders, with non-responder levels reducing only by 5% (to 66.79 ng/mmol creatinine). Assuming a pooled SD of 39.2 ng/mmol creatinine, at 80% power and two-sided $p<0.05$, and assuming 66% of participants would respond (responder to non-responder ratio of 2:1), a total of 194 completed participants were required. Allowance was made to adjust the sample size depending on the final responder:non-responder ratio. The same calculation was used to compare exacerbation responders and non-responders, assuming 66% of participants would be classed as exacerbation responders. The same number of participants was assumed to be required to test the hypothesis that 2,3-dinor-11β-PGF2α in urine was reduced in participants who subsequently have reduced exacerbations.

Dealing with missing values

For this study, we opted to use untargeted mass spectrometry (MS) in order to screen for a wide range of proteins and lipids. Whilst this approach enables broad coverage of both analyte types, it has the disadvantages of being less sensitive and more prone to picking up "noise". As a result of these limitations, and the high dimensionality of the information collected, u MS data usually include large amounts of "missing data" and require additional processing before further statistical and data analysis. It is often difficult to determine whether any null measurements for a given analyte were caused by it being below the limit-of-detection (MNAR: missing not-at-random), because of suboptimal sample processing or instrument setup (MAR: missing at-random), or because of some unknown factor out of the analyst's control (MCAR: missing completely at-random) (1). Proteomics and lipidomics have different ways of processing the spectral data, which leads to different ways of dealing with missing data; hence, these are described separately.

We used methods that are well established within the field of mass spectrometry. Thus, for proteins we undertook data filtration and normalisation as follows and as reported in detail the Journal of Proteome Research (2). Protein identifications collated from the ion accounting files were further quality filtered by allowing only identifications with the following criteria: identification in at least two separate samples (not including replicate injections), a process that required at least three high quality unmodified peptides using the Top 3 method, and 2 peptides with at least 4 fragment ions for each protein. High abundance proteins are generally more frequently identified, but many high abundance proteins are also identified at lower frequency. These abundant proteins are expected to be observed across multiple time points in the same patient, although there may be a proportion that are not replicated because of biological variation or where they are near the limits of detection. To illustrate this approach, please also see the figure in our previous publication, Burg et al. (2).

More than 2000 proteins can be identified in the sputum in ≥1 individual(s): 2,354 proteins in ≥2 individual(s), 284 proteins in ≥40% and 73 in ≥90% of individuals). High abundance proteins are generally more frequently identified, but many high abundance proteins are also identified at lower frequency. These abundant proteins are expected to be observed across multiple time points in the same patient, although there may be a proportion that

AJRCCM Articles in Press. Published April 18, 2024 as 10.1164/rccm.202310-1730OC

are not replicated because of biological variation or where they are near the limits of detection. With consideration for these effects, we have defined the sputum proteome in two ways, the 'core' and 'extended' sputum proteome. The 284 proteins identified in ≥40% of participants were defined as the 'core sputum proteome' and were used in the statistical analysis. The 'core' proteome represents the most commonly detected proteins within the sputum samples. The cut-off for the current study was defined at ≥40%, since at this frequency of identification, the frequency vs. protein rank curve was close to the point of inflection, where even a slight increase in the frequency of identification 'cut off', significantly increased the sparsity of the dataset and, hence, the total number of missing values. For more information, please see our paper by Burg et al. (2).

In untargeted MS-based lipidomics analyses, more than 1000 individual ions can be detected in any given sample. However, the vast majority of these ions are present in only one or a few of the samples, but absent or below the limit of detection in the rest. For example, we previously published an untargeted MS analysis of the induced sputum lipidome, in which only 32 ions were consistently detected in all samples, 141 ions in 90% or more of the samples, 214 ions in 80% or more, 291 ions in 60% or more, etcetera, in a sigmoidal fashion (3). Investigations of missingness in metabolomics studies have shown that it is primarily of the MNAR and MAR types, and not MCAR (1, 4). In other words, these null measurements are mostly caused by a combination of low abundance and low prevalence analytes, with only limited amounts of measurements that were missed completely at random by the MS instrument. This observation guides the strategy for dealing with missing values in the lipidomics data, which is commonly to select a certain cut-off (in our case, 40% missing) and remove all analytes with higher levels of missingness. The remaining missing values are then assumed to be of the MNAR type (e.g., primarily caused by low concentration analytes) and imputed with a small value, typically half the lower limit of detection. This left-censoring approach is the most widely used in high-dimensional lipidomics and metabolomics studies using untargeted MS approaches. Nevertheless, we acknowledge its downsides in potentially skewing parts of the data and introducing a bias for analytes prone to MAR-type missingness (e.g., lipids for which the analytical pipeline was less than optimal). This reinforces the need for validation of any lipid biomarkers in follow-on studies using targeted MS methods.

AJRCCM Articles in Press. Published April 18, 2024 as 10.1164/rccm.202310-1730OC

Table S1. Whole cohort demographics

n	191
Age (yr) median (IQR)	47 (35-55)
Sex (female/male) n (%)	116 (61) / 75 (39)
BMI (kg/m2) median (IQR)	29.1 (25.8-33.5)
ACQ7 median (IQR)	2.9 (2.1-3.6)
ACT median (IQR)	11.0 (9.0-16.0)
AQLQ median (IQR)	3.8 (2.9-5.0)
Acute exacerbation/12 months median (IQR)	4.0 (3.0-6.0)
mOCS: n (%)	74 (38.7)
Serum IgE (IU): median (IQR)	200.7 (105.0-345.9)
FeNO (ppb): median (IQR)	34.0 (16.7-60.0)
Blood eosinophils (x109/L): median (IQR)	0.25 (0.11-0.48)
Sputum eosinophils (%) median (IQR)	3.53 (0.88-19.1)
FEV_1 (% of predicted) median (IQR)	72.8 (56.7-85.1)

AJRCCM Articles in Press. Published April 18, 2024 as 10.1164/rccm.202310-1730OC

Table S2. Demographics of exacerbation responders and non-responders

Characteristic	**Exacerbation Responders (n=120)**	**Exacerbation Non-responders (n=49)**	**P- Value**
Age (yrs) at baseline[1]			
Median	47.5	45.0	0.4324
IQR	37.0 to 56.0	35.0 to 55.0	
Range	19.0 to 74.0	19.0 to 79.0	
Gender – n (%)[2]			
Male	44 (36.7%)	21 (42.9%)	0.4529
Female	76 (63.3%)	28 (57.1%)	
History of smoking – n (%)[2]			
Yes	29 (24.2%)	14 (28.6%)	0.5508
No	91 (75.8%)	35 (71.4%)	
BMI (kg/m2)[1]			
Median	29.4	28.7	
IQR	26.1 to 34.4	25.8 to 32.8	0.5227
Range	19.5 to 45.6	20.2 to 46.6	

AJRCCM Articles in Press. Published April 18, 2024 as 10.1164/rccm.202310-1730OC

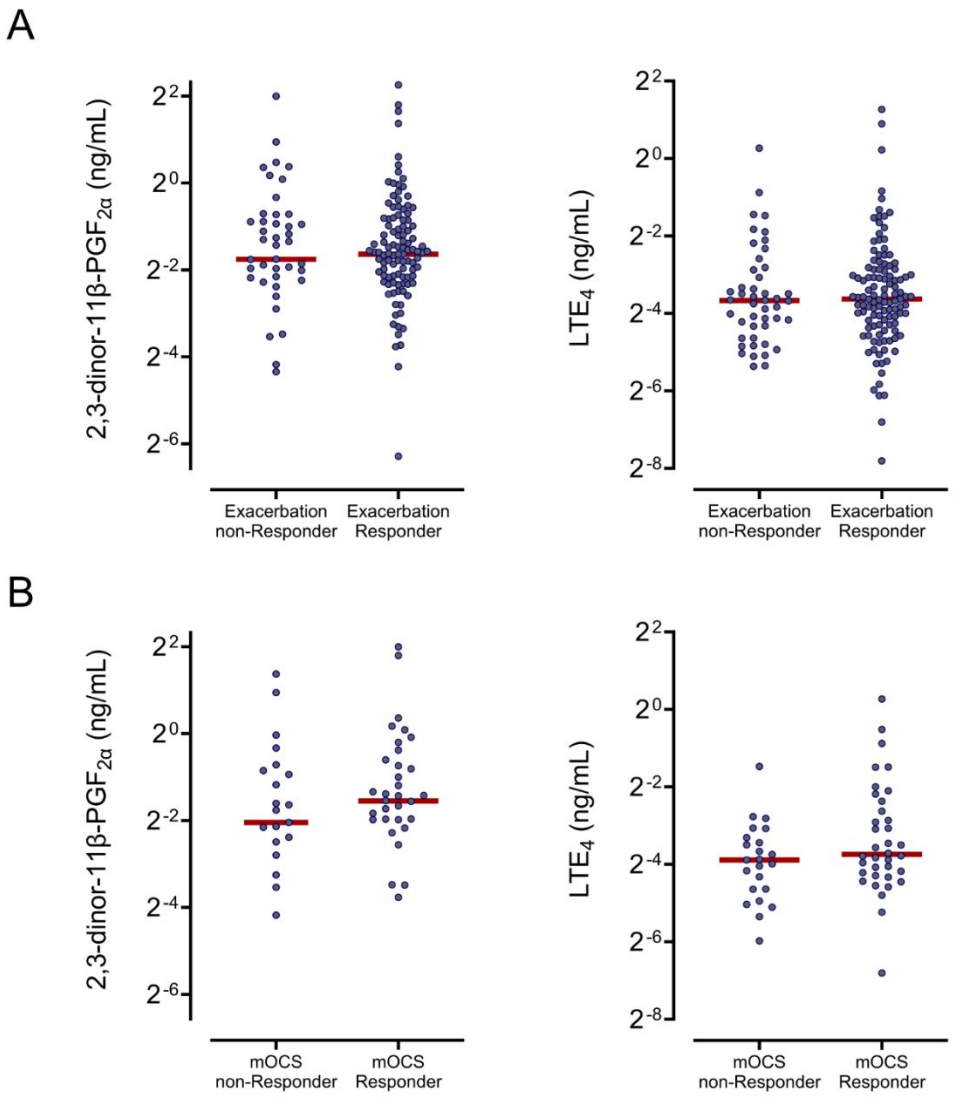


Figure S1. Urinary concentrations (ng/mL normalized to specific gravity) in patients stratified by non-responder/responder according to: (A) a reduction in exacerbations, (B) a reduction in mOCS use. Group median value shown in red. Concentrations of 2,3-dinor-11β-PGF2α were below LOD in 13 patients (panel A) and 7 patients (panel B) and were, therefore, excluded from the graphs.

AJRCCM Articles in Press. Published April 18, 2024 as 10.1164/rccm.202310-1730OC

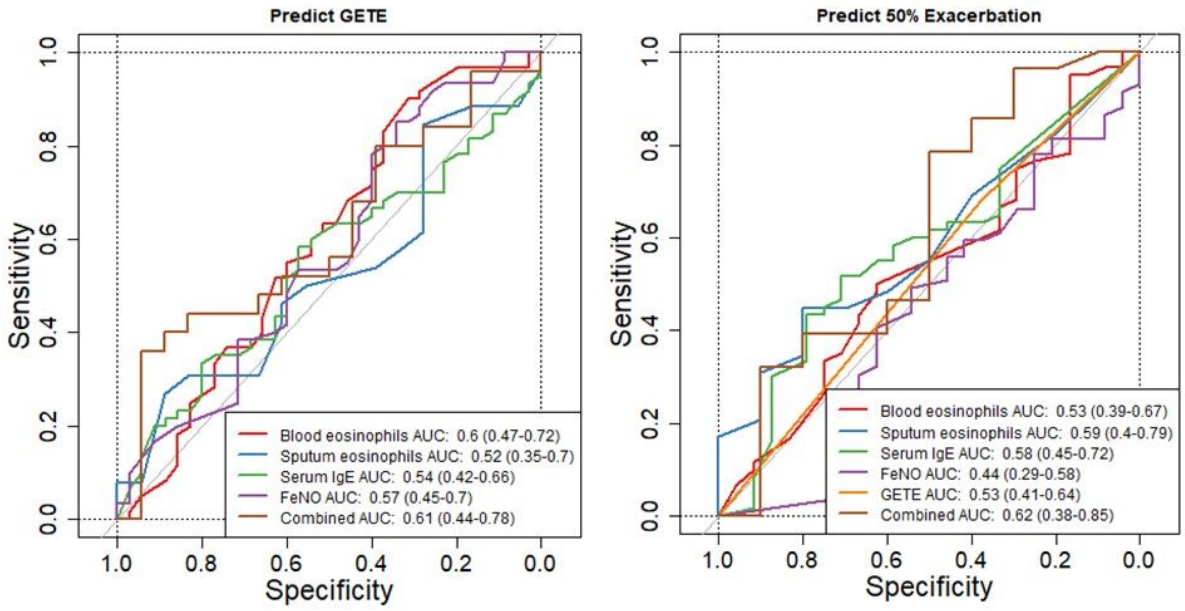


Fig S2

Figure S2. Prediction of early and late clinical improvement during treatment with omalizumab. Random Forest analysis was applied to training and test sets and results are shown as receiver operating characteristic (ROC) area under the curve (AUC). Early responses were assessed at 16 weeks using the Global Evaluation of Treatment Effectiveness (GETE) method and late responses were judged by ≥50% reduction in acute exacerbation rates. Standard clinical biomarkers assessed for predictive power included blood and sputum eosinophil counts, total serum IgE concentrations, and fractional exhaled nitric oxide (FeNO). Additionally, GETE itself was tested for the ability to predict acute exacerbation rate reduction. The predictive value of the combination of all biomarkers was also tested (combined AUC).

AJRCCM Articles in Press. Published April 18, 2024 as 10.1164/rccm.202310-1730OC

Stratification of the SoMOSA cohort using T2-high biomarkers, FeNO and blood eosinophil counts

A previous study (EXTRA) by Hanania et al, using T2-high biomarkers, FeNO, blood eosinophils and periostin, showed significant differences in exacerbation frequency following 48 weeks of omalizumab treatment between patients stratified (by all three biomarkers) as either T2-high or T2-low at study onset. In that study, stratification as either T2-high or T2-low was based on the following biomarker cut-off values: median blood eosinophil count (260 cells/ul), median FeNO in the EXTRA study (19.5 ppb), the FeNO cutoff recommended by the American Thoracic Society (24 ppb), as well as periostin levels of 50 ng/ml) from a previous trial of Lebrikizumab. Of note, all three biomarkers were predictive individually of better response with omalizumab, as evidenced both by comparing total rates within the strata and by assessing the time (expressed in months) to exacerbation. In our study, we stratified the SoMOSA cohort using three of the criteria used by Hanania et al., namely the EXTRA study's median blood eosinophil count (260 cells/ul) and median FeNO (19.5 ppb), and we also applied the FeNO cutoff recommended by the American Thoracic Society (24 ppb), as well as the periostin cutoff of 50 ng/ml from the trial of Lebrikizumab, but we could not stratify periostin levels as this biomarker was not measured in our study. As shown below, this analysis did not confirm the value of either FeNO or blood eosinophils, evident both by comparing total exacerbation rates of the two strata (Fig. S3) and by assessing the time (expressed in months) to exacerbation (Fig S4).

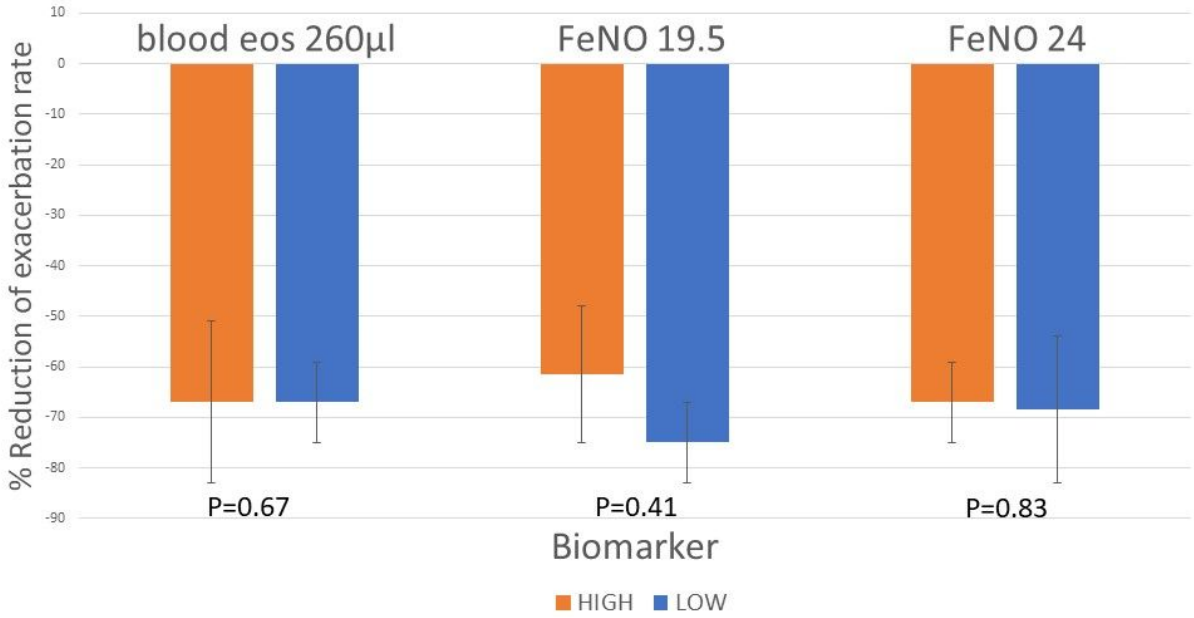
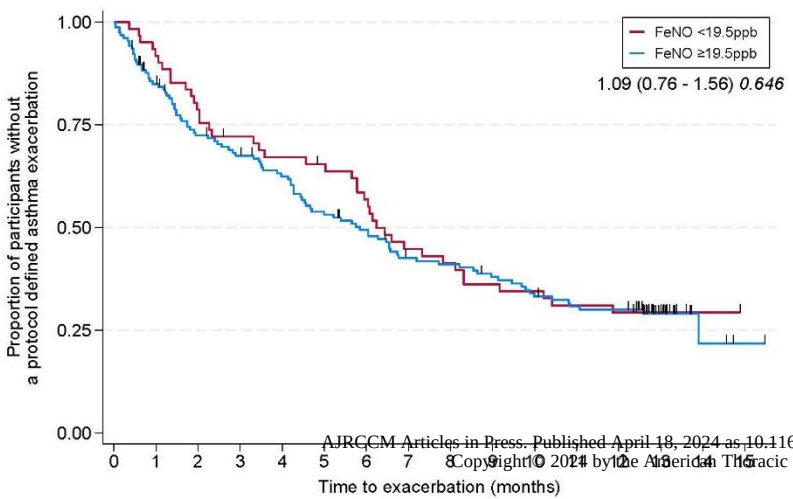


Fig. S3

Figure S3. Median percent reduction (95% Confidence interval) in protocol-defined asthma exacerbation rates in the low- and high-biomarker subgroups defined by the following cut-offs: blood eosinophils 260 ul/ml, FeNO 19.5 ppb (median values in the study by Hanania et al. Am J Respir Crit Care Med 2013;187:804-811) and FeNO 24 ppb (cut-off based on the ATS guidelines (Dweik et al. Am J Respir Crit Care Med 2011;184:602-615) . Comparison between biomarker high and biomarker low by Mann Whitney U test.



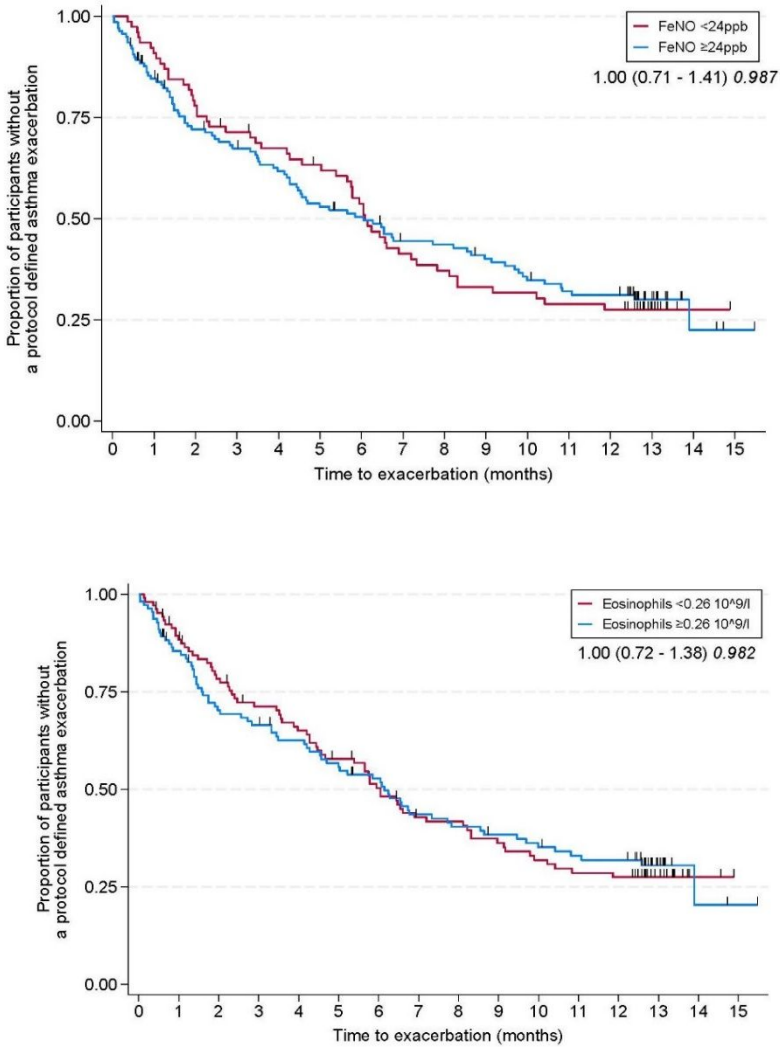


Figure S4. Time to first protocol-defined asthma exacerbation in baseline fractional exhaled nitric oxide (FeNO) low (<19.5 ppb) and high (≥19.5 ppb) subgroups (top), fractional exhaled nitric oxide (FeNO) low (<24 ppb) and high (≥24 ppb) subgroups (middle), and blood eosinophil low (<0.26 10^9/l) and high (≥0.26 10^9/l) subgroups (bottom). Hazard ratio with 95% confidence interval and p-value is presented within each Kaplan-Meier curve. For patients who did not experience an exacerbation, a hash mark is shown on the graph at the corresponding timepoint of censoring.

AJRCCM Articles in Press. Published April 18, 2024 as 10.1164/rccm.202310-1730OC

Omics platform	GETE ROC AUC	Acute exacerbations ROC AUC
Breathomics (GC-MS)	**0.835**	**0.780**
eNose	0.537	0.592
Plasma lipids	**0.949**	**0.922**
Sputum lipids	0.537	0.654
Sputum proteins	0.615	0.531
Urine proteins	0.479	0.587
Eicosanoids	0.533	0.576

Table S3

Table S3. Analysis of the ability of all omics platforms used in study to predict an early (GETE-based) or late (exacerbation-based) improvement by omalizumab treatment. The breathomics and lipidomics results are highlighted as these two platforms provided two highly predictive biomarker sets.

AJRCCM Articles in Press. Published April 18, 2024 as 10.1164/rccm.202310-1730OC

Omics platforms used in SoMOSA, U-BIOPRED and the MGB Biobank Cohort

Biomarker datasets from two independent cohorts were identified as suitable for additional analysis of the biomarkers shown in the core SoMOSA cohort as predictive of clinical responses to omalizumab: the U-BIOPRED study (311 non-smoking severe asthmatics and 88 mild/moderate nonsmoking asthmatics, none on omalizumab or another biologic) and the MGB Biobank (53 adult asthmatics given omalizumab: 22 responders and 31 non-responders to omalizumab treatment, with response defined as ≥50% reduction in exacerbations).

The analytical platforms in the independent cohorts differed more or less from those applied in SoMOSA.

UBIOPRED breath samples were analysed at the Philips Research laboratory (Eindhoven, The Netherlands) using GC-TOF-MS [Time-of-Flight; ref: https://pubmed.ncbi.nlm.nih.gov/31515400/] and SoMOSA was analysed at the Amsterdam UMC, location AMC, Netherlands, using GC-MS [Quadrupole]. The breath sampling technique applied was identical for both projects. The applied strategy to match exhaled VOC data from both studies was as following: 1) NIST library profiles of the ten SoMOSA predictor VOCs were selected, 2) UBIOPRED samples were analysed on the presence of these VOCs in an automated manner using the AMDIS software package, 3) VOCs (fragments; table S3-second column) were extracted from the UBIOPRED dataset, 4) univariate (wilcoxon signed rank test) and multivariate (sPLSDA; sparse Partial Least Squares Discriminant Analysis) analysis were conducted.

The lipidomics platform in U-BIOPRED used direct-infusion shotgun mass spectrometry on a quadrupole time-of-flight [Q-ToF] instrument for semi-quantitative measurement of individual plasma samples. This was followed by reversed-phase liquid chromatography separation [UPLC] and multistage mass spectrometry [MS/MS] analysis of a smaller number of pooled plasma samples for compound identification. In contrast, the lipidomics platform in SoMOSA utilized supercritical fluid chromatography [SFC] separation of lipid classes, followed by ion-mobility tandem mass spectrometry [IM-MS/MS] for identification, fragmentation and further cross-sectional separation of individual lipids. Quantification within each of the major classes was achieved using a dedicated set of stable isotope-labelled internal standards. All samples were aligned using dedicated vendor software (Progenesis QI, Waters Inc, Milford, MS, USA) and identities were obtained from external lipidomic databases queried by the same software.

For the MGB Biobank cohort, metabolomics profiling was conducted by Metabolon (North Carolina, USA). EDTA plasma samples from 53 patients with moderate to severe persistent allergic asthma on omalizumab, which had been stored at -80^{o}C, underwent liquid chromatography mass spectrometry using Metabolon's untargeted platform. This platform measures: amines, amino acids, polar lipids, non-polar lipids, fatty acids, and bile acids. Metabolites are identified by cross-referencing to a reference database. We followed a standard quality control pipeline for metabolomics. First, we evaluated missingness in all measured metabolites and excluded those with over 75% missingness. Thereafter, we imputed missing values for each metabolite with half the minimum value for that metabolite. Then, we evaluated the distribution of the metabolites, log-10 transformed, and conducted pareto scaling.

Analysis of the identified predictive biomarkers in the U-BIOPRED and MBM Biobank

A search of the U-BIOPRED data was undertaken for matching VOCs and plasma lipids, and comparisons were made between severe atopic asthmatics and mild/moderate asthmatics, with additional sub-stratification by exacerbation rates.

Nine of the ten matching VOCs were identified in the U-BIOPRED dataset (Table S4). Analysis by sPLSDA/ROC showed the following ROC AUCs: 0.771 [0.589-0.953] for comparisons of atopic severe vs. mild/moderate asthma (p=0.025) (Figure S5), and 0.931 [0.814-1.000] (p=0.0016) when severe asthmatics with ≥2 exacerbations per annum were selected for comparison (Figure S6).

A search of the U-BIOPRED data set found only 5 of the 9 lipids shown in the SoMOSA cohort to be predictive of early and/or late clinical responses (Table S4). Comparison of severe atopic and mild-moderate asthmatics from the U-BIOPRED data showed the concentrations of the following plasma lipids to be significantly higher in the severe patients: triglyceride (52:3) (PlP_2.96_879.9375_mz) (p<0.004) and an unknown lipid (PIP_1.20_881.9598_n) (p=0.045). Analysis of the atopic severe asthma dataset stratified into two categories (<2

AJRCCM Articles in Press. Published April 18, 2024 as 10.1164/rccm.202310-1730OC

or ≥2 exacerbations) showed similarly significant differences between more and less frequent exacerbators. However, sPLSDA showed weakly prognostic values, with ROC AUCs not much higher than 0.6 (see table S4).

In the MGB biobank, the concentrations of plasma sphingomyelin (sphingomyelin (d18:1/22:1, d18:2/22:0, d16:1/24:1) were significantly (p=0.03) lower in responders compared to non-responders to omalizumab.

Biomarker identities			**Comparison between ALL UBIOPRED non-smoking ATOPIC severe asthmatics and mild-moderate asthmatics: excluding patients on omalizumab**		**Comparison between ALL UBIOPRED non-smoking severe atopic asthmatics with ≥2 EXACERBATIONS and mild-moderate asthmatics: excluding patients on omalizumab**	
GETE Responder predictor VOCs SoMOSA IDs	**GETE Responder predictor VOCs U-BIOPRED IDs**	**Variable**	**p value**	**ROC AUC**	**p value**	**ROC AUC**
X103.1157	X135.1298.11	Benzothiazole	0.75		0.96	
X50.1.977	X120.1127.88	Acetophenone	0.71		0.37	
X109.7.861	X138.1041.33	2-Pentyl-Furan	0.63		0.48	
X35.157	X86.504.60	Methylene Chloride	0.045		0.074	
X39.2.265	X72.483.85	2-Methyl-Butane	0.26		0.093	
ROC AUC GETE predictors			0.025	0.771 (0.580-0.962)	0.027	0.819 (0.614-1.0)
Acute Exacerbations Responder predictor VOCs						
X55.1.902	X112.1081.03	2-Ethyl-1-Hexanol	0.13		0.17	
X63.539	X92.798.78	Toluene	0.29		0.2	
X79.175	n.a.	unknown				
X49.328	X70.490.26	2-Pentene	0.22		0.14	
X55.1.999	X124.1155.91	Nonanal	0.79		0.89	
ROC AUC Exacerb. Predictors			0.15	0.680 (0.466-0.893)	0.011	0.861 (0.663-1.0)
ROC AUC Combined GETE and Exacerb. Predictors			0.025	0.771 (0.589-0.953)	0.0016	0.931 (0.814-1.0)
GETE Responder Predictor plasma lipids						
PIP_1.20_878.8511_n		TG(54:6)	0.47		0.6	
PIP_1.22_881.9598_n		Unknown	0.045		0.061	
PIP_2.54_927.8584_mz		TG(56:7)	0.096		0.2	
PIP_2.62_918.8734_mz		TG(55:2)				
PIP_2.96_879.9375_mz		TG 52:3?	0.0039		0.028	
ROC AUC GETE predictors			0.015	0.592 (0.522-0.662)	0.038	0.585 (0.507-0.662)
Acute Exacerbations Responder predictor plasma lipids						
PIP_1.34_467.4848_mz		unknown				
PIP_2.96_879.9375_mz		TG 52:3?				
PIP_3.25_548.6240_mz		Ceramide				
PIP_3.47_368.4908_mz		unknown				
PIP_3.76_784.7540_n		SM(d40:2)	0.19		0.88	
ROC AUC Exacerb. Predictors			0.005	0.606 (0.537-0.675)	0.0055	0.614 (0.537-0.690)
ROC AUC Combined GETE and Exacerb. Predictors			0.013	593 (0.523-0.0663)	0.019	0.596 (0.519-0.674)

Table S4. Analysis of the SoMOSA predictive VOCs (top half) and plasma lipids (bottom half) from the U-BIOPRED database. ROC AUCs are shown for sets of biomarkers predictive of either the GETE-based early improvement or the late response judged by ≥50% reduced exacerbation rates. Shown in larger font are the ROC AUCs for combinations of GETE- and Exacerbation-based clinical improvements.

AJRCCM Articles in Press. Published April 18, 2024 as 10.1164/rccm.202310-1730OC

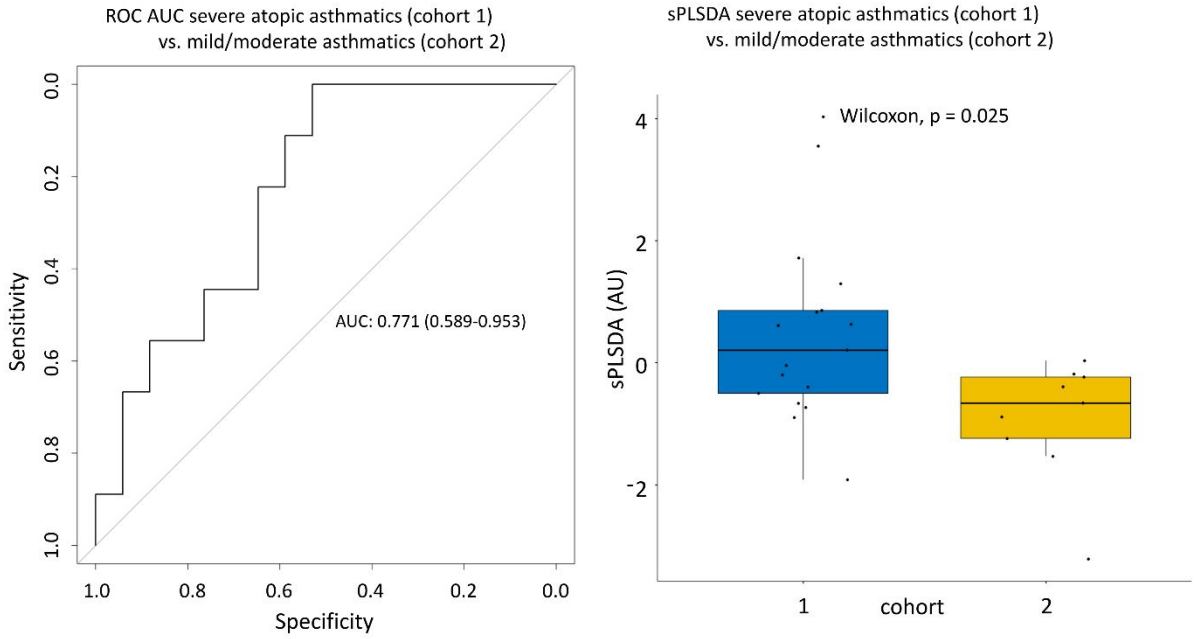


Figure S5. Receiver operating characteristic (ROC) area under the curve (AUC) comparing severe atopic asthmatics and mild/moderate asthmatics in the U-BIOPRED study.

AJRCCM Articles in Press. Published April 18, 2024 as 10.1164/rccm.202310-1730OC

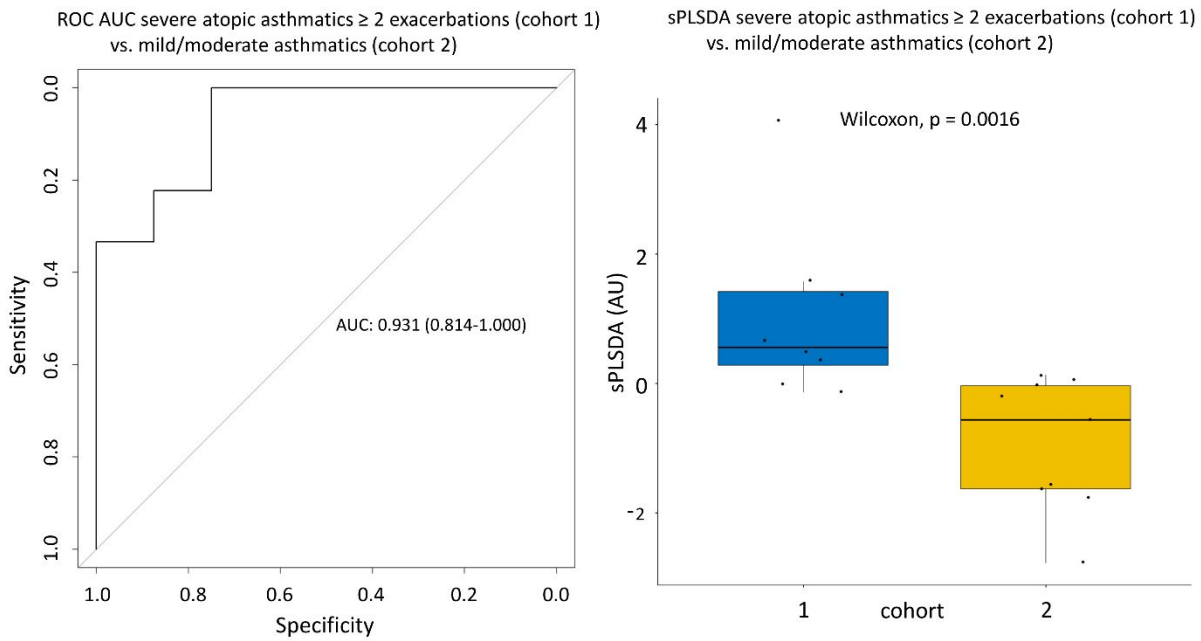


Figure S6. Receiver operating characteristic (ROC) area under the curve (AUC) comparing severe atopic asthmatics with at least 2 exacerbations in the past year and mild/moderate asthmatics in the U-BIOPRED study.

AJRCCM Articles in Press. Published April 18, 2024 as 10.1164/rccm.202310-1730OC

References

1. Davis TJ, Firzli TR, Higgins Keppler EA, Richardson M, Bean HD. Addressing Missing Data in GC × GC Metabolomics: Identifying Missingness Type and Evaluating the Impact of Imputation Methods on Experimental Replication. *Anal Chem* 2022; 94: 10912-10920.
2. Burg D, Schofield JPR, Brandsma J, Staykova D, Folisi C, Bansal A, Nicholas B, Xian Y, Rowe A, Corfield J, Wilson S, Ward J, Lutter R, Fleming L, Shaw DE, Bakke PS, Caruso M, Dahlen SE, Fowler SJ, Hashimoto S, Horváth I, Howarth P, Krug N, Montuschi P, Sanak M, Sandström T, Singer F, Sun K, Pandis I, Auffray C, Sousa AR, Adcock IM, Chung KF, Sterk PJ, Djukanović R, Skipp PJ, The UBSG. Large-Scale Label-Free Quantitative Mapping of the Sputum Proteome. *J Proteome Res* 2018; 17: 2072-2091.
3. Brandsma J, Goss VM, Yang X, Bakke PS, Caruso M, Chanez P, Dahlén SE, Fowler SJ, Horvath I, Krug N, Montuschi P, Sanak M, Sandström T, Shaw DE, Chung KF, Singer F, Fleming LJ, Sousa AR, Pandis I, Bansal AT, Sterk PJ, Djukanović R, Postle AD. Lipid phenotyping of lung epithelial lining fluid in healthy human volunteers. *Metabolomics* 2018; 14: 123.
4. Do KT, Wahl S, Raffler J, Molnos S, Laimighofer M, Adamski J, Suhre K, Strauch K, Peters A, Gieger C, Langenberg C, Stewart ID, Theis FJ, Grallert H, Kastenmüller G, Krumsiek J. Characterization of missing values in untargeted MS-based metabolomics data and evaluation of missing data handling strategies. *Metabolomics* 2018; 14: 128.

AJRCCM Articles in Press. Published April 18, 2024 as 10.1164/rccm.202310-1730OC