

# BMJ Open Identification of responders to inhaled corticosteroids in a chronic obstructive pulmonary disease population using cluster analysis

David R Hinds,<sup>1</sup> Rachael L DiSantostefano,<sup>1</sup> Hoa V Le,<sup>1,2</sup> Steven Pascoe<sup>3</sup>

**To cite:** Hinds DR, DiSantostefano RL, Le HV, *et al*. Identification of responders to inhaled corticosteroids in a chronic obstructive pulmonary disease population using cluster analysis. *BMJ Open* 2016;**6**:e010099. doi:10.1136/bmjopen-2015-010099

► Prepublication history and additional material is available. To view please visit the journal (<http://dx.doi.org/10.1136/bmjopen-2015-010099>).

Received 24 September 2015  
Revised 29 April 2016  
Accepted 3 May 2016



CrossMark

<sup>1</sup>Department of Worldwide Epidemiology, GSK, Research Triangle Park, North Carolina, USA

<sup>2</sup>PAREXEL International, Durham, North Carolina, USA

<sup>3</sup>Respiratory Medicines Development, GSK, King of Prussia, Pennsylvania, USA

## Correspondence to

David R Hinds;  
David.R.Hinds@gsk.com

## ABSTRACT

**Objectives:** To identify clusters of patients who may benefit from treatment with an inhaled corticosteroid (ICS)/long-acting  $\beta_2$  agonist (LABA) versus LABA alone, in terms of exacerbation reduction, and to validate previously identified clusters of patients with chronic obstructive pulmonary disease (COPD) (based on diuretic use and reversibility).

**Design:** Post hoc supervised cluster analysis using a modified recursive partitioning algorithm of two 1-year randomised, controlled trials of fluticasone furoate (FF)/vilanterol (VI) versus VI alone, with the primary end points of the annual rate of moderate-to-severe exacerbations.

**Setting:** Global.

**Participants:** 3255 patients with COPD (intent-to-treat populations) with a history of exacerbations in the past year.

**Interventions:** FF/VI 50/25  $\mu\text{g}$ , 100/25  $\mu\text{g}$  or 200/25  $\mu\text{g}$ , or VI 25  $\mu\text{g}$ ; all one time per day.

**Outcome measures:** Mean annual COPD exacerbation rate to identify clusters of patients who benefit from adding an ICS (FF) to VI bronchodilator therapy.

**Results:** Three clusters were identified, including two groups that benefit from FF/VI versus VI: patients with blood eosinophils  $>2.4\%$  (RR=0.68, 95% CI 0.58 to 0.79), or blood eosinophils  $\leq 2.4\%$  and smoking history  $\leq 46$  pack-years, experienced a reduced rate of exacerbations with FF/VI versus VI (RR=0.78, 95% CI 0.63 to 0.96), whereas those with blood eosinophils  $\leq 2.4\%$  and smoking history  $>46$  pack-years were identified as non-responders (RR=1.22, 95% CI 0.94 to 1.58). Clusters of patients previously identified in the fluticasone propionate/salmeterol (SAL) versus SAL trials of similar design were not validated; all clusters of patients tended to benefit from FF/VI versus VI alone irrespective of diuretic use and reversibility.

**Conclusions:** In patients with COPD with a history of exacerbations, those with greater blood eosinophils or a lower smoking history may benefit more from ICS/LABA versus LABA alone as measured by a reduced rate of exacerbations. In terms of eosinophils, this finding is consistent with findings from other studies; however, the validity of the 2.4% cut-off and the impact of smoking history require further investigation.

## Strengths and limitations of this study

- A key strength of this study was the availability of two large, global, randomised study populations, which allowed patients with chronic obstructive pulmonary disease to be well characterised such that differential responders to inhaled corticosteroids (ICS) could be identified and validated against nearly identical controlled trials.
- This analysis was among the first to identify that raised blood eosinophil levels may be predictive of treatment response to ICS, which may inform disease management and prescribing in clinical practice.
- A limitation of this analysis is that it is hypothesis-generating and failed to validate previous clusters identified using similar methodology. New findings regarding eosinophil levels require further external validation because this factor was not measured in the prior studies with an identical design.

**Trial registration numbers:** NCT01009463; NCT01017952; Post-results.

## INTRODUCTION

Chronic obstructive pulmonary disease (COPD) is an increasing burden on health-care systems globally.<sup>1</sup> Along with the persistent and progressive airflow limitation that characterises COPD,<sup>2</sup> exacerbations of COPD significantly impair patients' lives and further increase the burden on healthcare services.<sup>2 3</sup> Moderate-to-severe exacerbations are significant events associated with poor patient prognosis, including worse quality of life, more rapid disease progression and, for severe events, increased mortality.<sup>2 3</sup> Addition of an inhaled corticosteroid (ICS) to bronchodilator maintenance therapy has been shown to decrease exacerbations among patients

with COPD<sup>4–10</sup> and, as a result, the current Global Initiative for Chronic Lung Disease (GOLD) guidelines recommend a fixed-dose combination therapy of ICS/long-acting  $\beta_2$  agonist (LABA) for patients at risk of frequent exacerbations (ie, two or more per year).<sup>2</sup>

The identification of patients who may respond more favourably to one treatment option versus another is an important consideration for patients and healthcare providers, allowing for optimal patient management and evaluation of risks versus benefits when allocating limited healthcare resources.<sup>5 6 11</sup> Cluster analysis is one method of identifying groups of patients who are more likely to benefit from one treatment versus another based on clinical characteristics, and this analysis can be particularly useful in complex, heterogeneous diseases such as COPD.<sup>12</sup>

Cluster analysis has previously been used to identify patients who had greater exacerbation reduction with an LABA (salmeterol (SAL)) in combination with an ICS, fluticasone propionate (SAL/fluticasone combination, SFC) compared with SAL alone.<sup>13</sup> Clusters of patients receiving diuretics and those not receiving diuretics but with a baseline bronchodilator reversibility of  $\geq 12\%$  were found to have significantly greater reductions in exacerbations when treated with SFC versus SAL, based on clinical trial data comparing treatment with SFC versus SAL. No significant differences were found between the treatments in a third cluster of patients not receiving diuretics and with baseline bronchodilator reversibility of  $< 12\%$ .

The present analysis evaluated data from two recent clinical trials comparing the treatment with an ICS/LABA versus an LABA alone<sup>6</sup> and aimed to identify clusters of patients with COPD who may benefit from the addition of an ICS to an LABA in terms of exacerbation reduction, as well as to validate the clusters identified from the previous analysis of SFC versus SAL.<sup>13</sup>

## METHODS

### Clinical study design and patients

The full methodology for the two clinical trials included in this analysis (NCT01009463 and NCT01017952) has been previously reported.<sup>6</sup> Briefly, these were randomised, double-blind, parallel group, 52-week, multicentre studies. The primary end point of both trials was the annual rate of moderate-to-severe exacerbations. Eligible patients were aged  $\geq 40$  years and had a history of COPD, a smoking history of  $\geq 10$  pack-years and a forced expiratory volume in 1 s (FEV<sub>1</sub>) of  $\leq 70\%$  predicted and an FEV<sub>1</sub>/forced vital capacity (FVC) ratio of  $\leq 0.70$  after bronchodilator use; they also had a history of at least one documented moderate or severe disease exacerbation in the year before screening. Patients with a prior history of asthma were eligible if they had a current diagnosis of COPD, but those with a current diagnosis of asthma were excluded. For 4 weeks prior to randomisation, patients received an open-label combination of

fluticasone propionate (250  $\mu\text{g}$ ) and SAL (50  $\mu\text{g}$ ) two times per day to establish adherence to treatment and a stable baseline. In both studies, patients were randomised to receive one time per day fluticasone furoate (FF)/vilanterol (VI) 50/25, 100/25 or 200/25  $\mu\text{g}$ , or one time per day VI 25  $\mu\text{g}$  monotherapy using the ELLIPTA<sup>1</sup> dry powder inhaler.

These clinical trials had identical study designs to the two trials of SFC versus SAL that were used in the cluster analysis by DiSantostefano *et al.*,<sup>13</sup> with the exception of study medication, and presented the unique opportunity to validate the previously identified clusters of patients who benefited from the addition of an ICS in COPD. The previously identified cluster modelling was therefore applied to the FF/VI versus VI data in an attempt to validate the findings.

### Cluster analysis methodology

This exploratory retrospective analysis was performed using a modified recursive partitioning algorithm to carry out supervised cluster analysis that identified new groups of patients who responded better to an ICS/LABA combination compared with an LABA alone. The data-driven algorithm maximised differential responses to treatment between clusters in terms of exacerbation reduction (FF/VI vs VI), based on clinical characteristics from the intent-to-treat populations. For these analyses, all strengths of FF/VI were collapsed and considered as one treatment group.

### Variables

The baseline characteristics considered to identify clusters of patients are summarised in [table 1](#). In the case of missing values for continuous variables, the median value for the full data set was used as an imputed value. For categorical variables, the value of the most frequent category was used as an imputed value. Missing values were generally infrequent and only imputed for variables in the clustering model. The standard rpart clustering methodology addresses missing values by including only observations (patients) with non-missing values. Considering the size of this data set, missing values for variables in the clustering model were imputed to optimise the data and to be consistent with previously conducted cluster analyses.<sup>13</sup> To avoid confounding by similar variables (eg, membership of a specific age group such as 65–75 years vs age as a continuous variable), each variable was assessed for co-linearity with every other variable using Pearson's correlation coefficients. Where two similar variables exhibited a Pearson correlation coefficient value of  $\geq 0.7$ , only one variable was retained. Baseline information was summarised for each cluster. As cluster analysis requires a sufficiently large overall sample size ( $> \sim 500$  patients) to

<sup>1</sup>ELLIPTA is a trademark of the GSK group of companies.

**Table 1** Variables considered for the initial cluster analysis

	Variables
Demography and medical history	Age (years) Sex Race Ethnicity Body mass index (kg/m <sup>2</sup> ) Pneumococcal vaccination status (baseline) Influenza vaccination status (baseline)
COPD and exacerbation history	Duration of COPD, bronchitis/emphysema, moderate and severe exacerbation variables at baseline, smoking status/history Outcome data: mean annual moderate/severe exacerbation data to reproduce primary end point modelling
Concomitant medication assessment	Medications at run-in/prior medications (including ATC codes and flags to identify prior period if needed)
Outcome/efficacy assessments	Annual rate of moderate/severe COPD exacerbations as defined in the two clinical trials. Moderate: worsening symptoms of COPD that required treatment with oral corticosteroids and/or antibiotics; severe: worsening symptoms of COPD that required treatment with in-patient hospitalisation
Screening and baseline spirometry	Pre-bronchodilator FEV <sub>1</sub> (L) Post-bronchodilator FEV <sub>1</sub> (L) Per cent predicted post-bronchodilator FEV <sub>1</sub> (%) Post-bronchodilator FEV <sub>1</sub> /FVC (%) Per cent reversibility FEV <sub>1</sub> (%) FEV <sub>1</sub> reversibility (mL) Reversibility at screening Baseline FEV <sub>1</sub> (L)
Physical examination	Screening/baseline medical conditions page: (body system level yes/no) with exception of cardiac disorders where each condition is listed explicitly from screening/baseline medical conditions page: <ul style="list-style-type: none"> <li>▶ Cardiac disorders: congestive heart failure, coronary disease, myocardial infarction and arrhythmia</li> <li>▶ Eye disorders</li> <li>▶ Metabolism and nutrition disorders</li> <li>▶ Vascular disorders</li> <li>▶ Infections and infestations: pneumonia</li> <li>▶ Endocrine disorders</li> </ul>
Laboratory assessments/ECG	Baseline lymphocytes, WBC, eosinophils, neutrophils, BUN/urea, RBC, haemoglobin, ECG (p-wave dispersion)

ATC, anatomical therapeutic chemical; BUN, blood urea nitrogen; COPD, chronic obstructive pulmonary disease; FEV<sub>1</sub>, forced expiratory volume in 1 s; FVC, forced vital capacity; RBC, red blood cells; WBC, white blood cell counts.

identify maximum differences in response in a robust manner, the algorithms were set to allow clusters of no <100 patients.

### Modelling

For this analysis, a data-driven, modified, recursive partitioning technique<sup>14</sup> was employed in a similar manner to that used previously for a similar study.<sup>13</sup> Computations were performed using the rpart package (T Therneau, B Atkinson, B Ripley. RPART: Recursive Partitioning. R package version 4.0-1. 2012. <http://cran.r-project.org/web/packages/rpart/index.html> (accessed 6 Jan 2014)) for the statistical software R (R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing, 2016. <http://www.R-project.org> (accessed

26 Apr 2016)). In brief, the frequency of each variable was examined for sparse values prior to inclusion into the tree with the minimal subgroup set at 100 patients. The best split of the tree was determined by maximising the treatment differences between clusters, and cluster membership was assigned to patients based on the selected tree. Splits in the clustering tree (including the value on which the tree splits for continuous variables) are determined by the clustering algorithm. The primary model for the FF/VI versus VI studies was used in maximising treatment differences between clusters. The primary model from the trials was a negative binomial model of the number of moderate and severe exacerbations adjusted for smoking status at screening, geographical region, FEV<sub>1</sub>% predicted at randomisation and Study 1 versus Study 2 indicator variable, with log

time on treatment per patient as an offset. Imbalances in baseline characteristics by treatment within each cluster were also evaluated and added to the final model ( $p < 0.05$ ).

Rate ratios (RRs) and 95% CIs were used to estimate the differences in annual mean exacerbation rates for each cluster. Internal validation was performed using a split sample, so that a random sample of 50% of the patients was selected to create the tree and the remaining half was used for the computation of RRs and CIs to test statistical significance. Baseline variables within each cluster are presented as proportions for categorical variables and medians (interquartile range) for continuous variables. Differences between variables in clusters were assessed using  $\chi^2$  for categorical variables and the Wilcoxon rank-sum test for continuous variables.

### Validation

To validate the clusters from the previous analysis of SFC versus SAL,<sup>13</sup> the same negative binomial model and final clusters from the recursive partitioning tree from that analysis were evaluated using the FF/VI versus VI data. The SFC versus SAL model was adjusted for baseline FEV<sub>1</sub>% predicted, reversibility status (yes/no for  $\geq 12\%$  and  $\geq 200$  mL improvement) and investigator region as a random effect, with log time on treatment per patient as an offset. The prior clusters were considered validated if the resulting RRs were in the same direction, and the  $p$  value for the overall treatment by cluster interaction was  $< 0.05$  utilising the FF/VI versus VI data. The clusters were also considered conceptually validated if the new clusters identified from the FF/VI versus VI data were defined by similar characteristics as the SFC versus SAL clusters (reversibility and/or diuretic use). For the new clusters identified in the FF/VI versus VI analysis, external validation using the SFC versus SAL data was considered, using the same RR and  $p$  value criteria described above.

## RESULTS

A total of 3255 patients were included from the intent-to-treat populations of the FF/VI clinical trials, with ~800 patients included in each of the four treatment arms. The median age of patients was 63.6 years, and most were men and of Caucasian ethnicity. Further baseline demographics have been previously presented.<sup>6</sup>

### Identification of clusters

Three clusters of patients were identified from the FF/VI versus VI data. The clusters were defined by blood eosinophil levels (as a percentage of white cell counts) and smoking history (in pack-years). Using the primary trial model, cluster 1 included patients with blood eosinophils  $> 2.4\%$  (RR=0.66, 95% CI 0.57 to 0.78), cluster 2 included patients with blood eosinophils  $\leq 2.4\%$  and a smoking history of  $\leq 46$  pack-years (RR=0.76, 95% CI 0.62 to 0.94) and cluster 3 included patients with blood

eosinophils  $\leq 2.4\%$  and a smoking history of  $> 46$  pack-years (RR=1.23, 95% CI 0.95 to 1.60). Patients in clusters 1 and 2 exhibited statistically significant reductions in mean annual exacerbation rate, whereas no reduction was observed for those in cluster 3.

Within clusters, the number of hospitalisations due to exacerbations, number of steroid/antibiotic-treated exacerbations, COPD type (bronchitis and emphysema), ethnicity and reversibility subgroup at baseline were imbalanced ( $p < 0.05$ ) by treatment (see online supplementary table S1). These variables were therefore added to the model of moderate-to-severe exacerbations for the final model used in this analysis. After controlling for explanatory variables from the original clinical trial model and imbalances within clusters, treatment with FF/VI reduced the rate of annual moderate and severe exacerbations by 22% (RR=0.78, 95% CI 0.70 to 0.88) compared with VI alone (figure 1). Statistically significant reductions in mean annual exacerbations were observed in cluster 1, patients with blood eosinophils  $> 2.4\%$  (RR=0.68, 95% CI 0.58 to 0.79), and cluster 2, patients with blood eosinophils  $\leq 2.4\%$  and smoking history of  $\leq 46$  pack-years (RR=0.78, 95% CI 0.63 to 0.96), no reduction was observed in cluster 3, patients with blood eosinophils  $\leq 2.4\%$  and smoking history of  $> 46$  pack-years (RR=1.22, 95% CI 0.94 to 1.58). The percentage of eosinophils was the primary determinant of exacerbation reduction in the clustering tree.

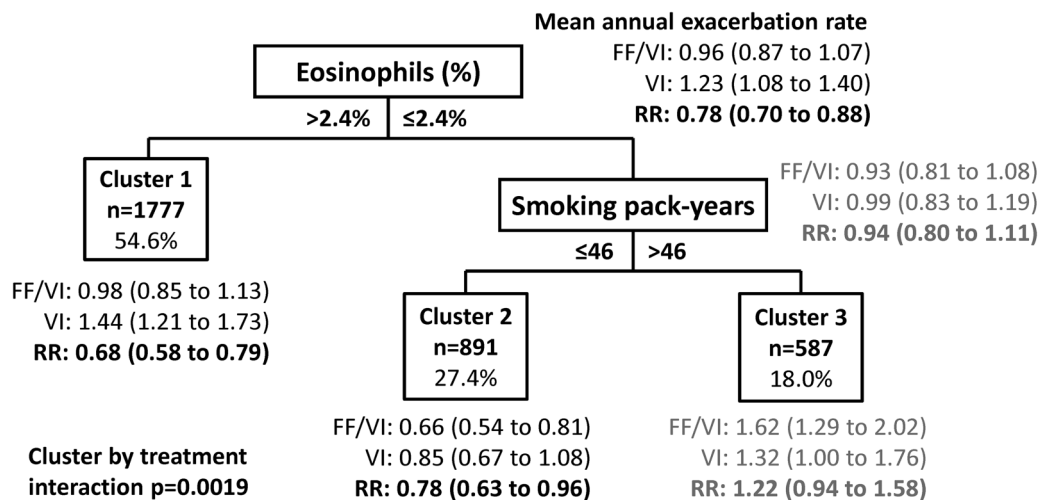
The baseline patient characteristics for the three clusters, including specific COPD characteristics (as determined by each patient's clinician) and laboratory values, are presented in table 2. Cluster 2 tended to differ from clusters 1 and 3 and had more women, fewer comorbidities and less medication use, less emphysematic COPD and less reversibility. Clusters 1 and 3 were generally similar, although cluster 1 had more bronchitic COPD relative to emphysematic, whereas cluster 3 had similar frequencies of the two COPD types.

### Sensitivity analyses

Sensitivity analyses were conducted to understand better the primary split in the clustering tree of percentage eosinophils. The relationship between blood eosinophil level and treatment effect was generally linear (figure 2); therefore, the split in clusters at a blood eosinophil cut-off of 2.4% was driven by the cluster analysis methodology. In creating the clusters, the algorithm split the clusters near the median value of eosinophils (2.6%).

Sensitivity analyses were also conducted to examine the effect of imputing missing values for the 77 patients with missing values for blood eosinophils. The missing values for these patients were imputed with the median eosinophil value for the data set (ie, they were assigned to cluster 1, which included patients with eosinophils  $> 2.4\%$ ). In sensitivity analyses where these patients were removed from the analysis and forced to the other side of the eosinophil split, the results did not appreciably





**Figure 1** Clusters maximising treatment differences in the FF/VI versus VI population (independent clustering solution, fully adjusted model). The model adjusted for investigator region, FEV<sub>1</sub>% predicted at randomisation, number of hospitalised exacerbations, number of steroid/antibiotic-treated exacerbations, COPD type—bronchitis, COPD type—emphysema, study ID, smoking status, ethnicity and reversibility subgroup. Numbers in parentheses are 95% CIs. COPD, chronic obstructive pulmonary disease; FEV<sub>1</sub>: forced expiratory volume in 1 s; FF, fluticasone furoate; RR, rate ratio; VI, vilanterol.

change. The patients were further classified into clusters 2 or 3 depending on their smoking history.

### Validation

The clusters of patients identified from the FF/VI versus VI trials could not be validated with the SFC versus SAL trial data that were used in the previous cluster analysis by DiSantostefano *et al.*<sup>13</sup> Eosinophils as a percentage of white cell counts were not captured in the SFC versus SAL trials, so the clustering model could not be used with these data.

Additionally, the FF/VI trial data did not validate the clusters identified in the prior analysis of SFC versus SAL<sup>13</sup> (figure 3). Within each of the clusters defined by the SFC versus SAL analysis, patients receiving FF/VI had lower exacerbation rates compared with those who received VI alone. These findings are consistent with the previous analysis, except for the cluster of patients who were not receiving diuretics and were not reversible (<12%) and who were previously identified as a group of non-responders.<sup>13</sup>

Despite the identical methods, the SFC versus SAL and FF/VI versus VI trial populations differed. This may have been related to the SFC versus SAL trials being conducted in the USA and Canada, whereas the FF/VI versus VI studies included a global population. Patients in the FF/VI versus VI trials appeared to have less severe disease than those in the SFC versus SAL trials with less severe airflow limitation (FEV<sub>1</sub>% predicted), fewer exacerbations during baseline and less medication use for comorbid disease. A sensitivity analysis restricting the FF/VI data to North American data only was conducted. This analysis did not validate the previously identified clusters, but it did attenuate many of the imbalances between the patient populations.

### DISCUSSION

This analysis was performed to identify clusters of patients with COPD who might benefit from an ICS added to bronchodilator therapy, as measured by a reduced rate of moderate and/or severe exacerbations, and to validate the clusters previously identified by DiSantostefano *et al.*<sup>13</sup> Three clusters were identified from the FF/VI versus VI data. The blood eosinophil level was the primary driver of treatment response, with a greater treatment effect observed in patients with blood eosinophil levels >2.4% when treated with FF/VI versus VI alone. The treatment effect was generally linear, with the mean exacerbation rate decreasing as eosinophil levels increased. Patients with blood eosinophils ≤2.4% were further differentiated by smoking history. Patients with a shorter smoking history (≤46 pack-years) responded well to the combined treatment, but those with longer smoking history did not respond.

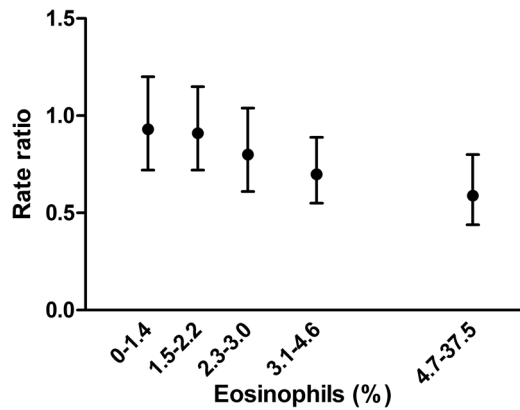
This analysis was unable to validate the previously identified clusters from the analysis of SFC versus SAL, where patients using diuretics and those who were reversible (>12%) but not using diuretics showed a significant reduction in exacerbation when treated with SFC versus SAL. Patients who were not using diuretics and who were not reversible did not show a treatment benefit of SFC versus SAL, but they did show a treatment benefit with FF/VI versus VI. The inability of this analysis to validate the clusters previously identified in the analysis of SFC versus SAL<sup>13</sup> is potentially due to differences between the patient populations. Besides different treatments across studies, one of the biggest differences was that the SFC versus SAL studies were primarily conducted in the USA and Canada, whereas the FF/VI versus VI studies included a global patient population. Other differences that may explain the inconsistent

**Table 2** Patient characteristics of the three clusters identified in the final model

Patient characteristics	Cluster 1 eosinophils >2.4% (N=1777)	Cluster 2 Eosinophils ≤2.4%, pack-years ≤46 (N=891)	Cluster 3 eosinophils ≤2.4%, pack-years >46 (N=587)	p Value
Age, years	65 (58–70)	62 (55–69)	65 (60–71)	<0.0001
Female	693 (39.0)	480 (53.9)	212 (36.1)	<0.0001
Hispanic or Latino	298 (16.8)	192 (21.5)	94 (16.0)	0.0041
Current smoker	714 (40.2)	443 (49.7)	282 (48.0)	<0.0001
Number of pack-years	40 (28–58)	30 (20–39)	61 (51–80)	<0.0001
History of				
Cardiovascular disease	1108 (62.4)	519 (58.2)	382 (65.1)	0.0219
Cardiac disorders	306 (17.2)	113 (12.7)	107 (18.2)	0.0036
Coronary artery disease	186 (10.5)	65 (7.3)	73 (12.4)	0.0031
Myocardial infarction	104 (5.9)	26 (2.9)	35 (6.0)	0.0027
Eye disorders	225 (12.7)	87 (9.8)	95 (16.2)	0.0012
Metabolism disorders	729 (41.0)	321 (36.0)	244 (41.6)	0.0277
Medication use				
Antithrombotic agents	486 (27.3)	178 (20.0)	167 (28.4)	<0.0001
Antihypertensives	87 (4.9)	27 (3.0)	35 (6.0)	0.0195
Lipid-modifying agents	517 (29.1)	187 (21.0)	177 (30.2)	<0.0001
Anti-inflammatory and antirheumatic products	643 (36.2)	242 (27.2)	209 (35.6)	<0.0001
Antihistamines for systemic use	216 (12.2)	73 (8.2)	54 (9.2)	0.0036
Exacerbations treated with steroid/antibiotics in the 12 months before screening visit				
0	129 (7.3)	67 (7.5)	54 (9.2)	0.47
1	1104 (62.4)	547 (61.4)	341 (58.1)	.
2	363 (20.4)	194 (21.8)	136 (23.2)	.
≥3	181 (10.2)	83 (9.3)	56 (9.5)	.
Number of hospitalised exacerbations in the 12 months before screening visit				
0	1428 (80.4)	696 (78.1)	469 (79.9)	0.1044
1	291 (16.4)	152 (17.1)	104 (17.7)	.
≥2	58 (3.3)	43 (4.8)	14 (2.4)	.
COPD type				
Bronchitis	1156 (65.4)	634 (71.2)	366 (62.8)	0.0012
Emphysema	1021 (57.8)	426 (47.9)	362 (62.1)	<0.0001
Baseline WBC, GI/L	6.90 (5.80–8.10)	7.20 (6.10–8.60)	7.40 (6.20–8.90)	<0.0001
Baseline lymphocytes, %	27.30 (22.30–33.00)	25.80 (21.10–31.30)	25.30 (19.70–30.10)	<0.0001
Baseline eosinophils, GI/L	0.28 (0.21–0.39)	0.11 (0.07–0.14)	0.11 (0.07–0.15)	<0.0001
Baseline eosinophils, %	4.00 (3.10–5.60)	1.50 (1.00–2.00)	1.50 (1.00–2.00)	<0.0001
Baseline eosinophils ≥2%	1700 (100)	233 (26.2)	158 (26.9)	<0.0001
Baseline neutrophils, %	61.10 (55.10–66.70)	66.40 (60.60–71.50)	67.10 (62.10–72.70)	<0.0001
Baseline neutrophils, GI/L	4.19 (3.37–5.11)	4.81 (3.81–5.98)	4.99 (3.85–6.27)	<0.0001
Baseline haemoglobin, G/L	145.0 (136.0–153.0)	145.0 (136.0–154.0)	147.0 (138.0–156.0)	0.0025
Baseline urea/BUN, mmol/L	5.50 (4.50–7.00)	5.20 (4.20–6.50)	5.45 (4.50–6.50)	0.0024
Lung function at screening				
Pre-bronchodilator FEV <sub>1</sub> , L	1.10 (0.81–1.41)	1.12 (0.81–1.42)	1.02 (0.75–1.40)	0.0457
Pre-bronchodilator % predicted FEV <sub>1</sub>	39.90 (30.00–50.00)	41.70 (31.10–51.60)	38.00 (28.20–50.30)	0.0015
Post-bronchodilator FEV <sub>1</sub> , L	1.24 (0.96–1.57)	1.22 (0.94–1.55)	1.17 (0.89–1.57)	0.0603
Post-bronchodilator % predicted FEV <sub>1</sub>	45.70 (35.45–56.00)	46.50 (36.50–56.70)	44.00 (32.60–54.95)	0.0160
Post-bronchodilator FEV <sub>1</sub> /FVC	45.10 (36.80–53.90)	46.80 (37.80–55.40)	43.20 (35.35–52.50)	0.0002
Reversibility at screening				
Reversibility, %	13.00 (5.10–23.20)	10.80 (3.20–19.40)	12.70 (5.20–22.70)	<0.0001
Reversibility, mL	140.00 (50.00–230.0)	120.00 (30.00–200.0)	130.00 (50.00–220.0)	<0.0001
Reversibility subgroup (≥12% and ≥200 mL increase)	578 (33.0)	224 (25.5)	170 (29.2)	0.0004
Lung function at baseline				
Post-bronchodilator FEV <sub>1</sub> , L	1.17 (0.86–1.53)	1.17 (0.88–1.49)	1.10 (0.82–1.48)	0.0222
Post-bronchodilator % predicted FEV <sub>1</sub>	42.70 (32.00–53.95)	44.10 (33.50–53.80)	41.20 (29.85–52.65)	0.0075

Data are n (%) or median (IQR) unless otherwise stated. Data presented include characteristics statistically significantly different ( $p < 0.05$ ) between clusters or imbalanced by treatment ( $p < 0.05$ ) within clusters (see online supplementary table S1).

BUN, blood urea nitrogen; COPD, chronic obstructive pulmonary disease; FEV<sub>1</sub>, forced expiratory volume in 1 s; FVC, forced vital capacity; WBC, white cell count.



**Figure 2** Treatment effect by blood eosinophil percentage. Each value represents the average treatment effect for each 20% of individuals ordered by eosinophil level, against the median of eosinophil within the group. The model adjusted for investigator region, study ID, forced expiratory volume in 1 s % predicted at randomisation and smoking status (model from the source clinical trials). p Value for linear trend of treatment effect:  $p=0.0081$ .

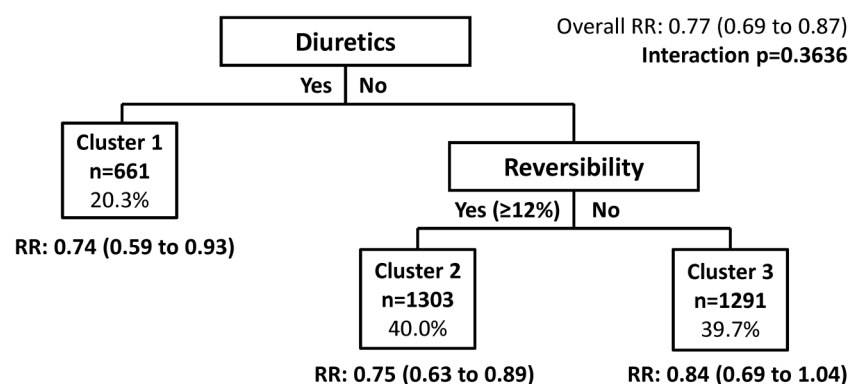
results were that patients in the SFC versus SAL studies tended to have greater airflow limitation, fewer exacerbations at baseline and less medication use.<sup>6, 13</sup> Limiting the FF/VI analysis to North America still did not validate the prior findings, suggesting that other factors contribute to differences between the trial populations. Most notably, guidelines for COPD management and diagnosis of COPD evolved during the time between the studies and may have resulted in differences in the populations recruited into the two sets of clinical trials.

A key strength of this analysis is that it was conducted using data from two large, global clinical trials,<sup>6</sup> and the results have therefore been generated from a robust data set of controlled trials with thorough data collection. On the other hand, the limitations of these clinical trials are also limitations of our analysis and, in particular, the generalisability of our findings is limited to patients with a history of exacerbations, as per the study inclusion criteria. Although cluster analysis is a useful tool to provide important information about differential treatment response among some clusters of patients who

may not respond in a manner similar to the total clinical trial population, the standard limitations of any cluster analysis apply.<sup>12</sup> These analyses are hypothesis-generating using a data-driven algorithm and depend on the baseline patient data collected, and the results may or may not be clinically meaningful. For example, history of asthma was not collected in the trial and, therefore, could not be described by a cluster. Eosinophils are established as a measure of inflammation in asthma phenotypes;<sup>15</sup> therefore, similar to the previous cluster analysis,<sup>13</sup> patients with asthma-COPD overlap could be a potential explanation for some of the observed treatment effect. Although history of asthma was not collected, the clustering algorithm did not split on reversibility variables and similar percentages of each cluster exhibited  $\geq 12\%$  reversibility and  $\geq 200$  mL improvement in FEV<sub>1</sub>. In addition, the discrimination of asthma from COPD in clinical practice for some patients is difficult, particularly for patients presenting with smoking-related COPD. The analysis also relied on the baseline data collected regarding the variables in the clustering model; 77 patients had missing eosinophil per cent data, no patients had missing data on smoking pack-years and small numbers of patients (0–44) had missing data for one of the variables included in the model. Owing to the size of the data set, these missing values were imputed, but the results did not appreciably change in sensitivity analyses where values were not imputed. Other limitations to cluster analysis include that the algorithm cannot adjust for multiple comparisons, which could yield spurious results. We were unable to externally validate the clusters defined by eosinophils and smoking history using the SFC versus SAL data because baseline eosinophil levels were not collected in the SFC versus SAL studies. This is an important limitation in light of the fact that the findings from the previous cluster analysis of SFC versus SAL were not able to be externally validated within the present analysis.

Eosinophil levels are becoming established as a clinically relevant biomarker of oral corticosteroid and ICS response in COPD<sup>16–23</sup> and also in asthma.<sup>24–26</sup> Earlier work focused on sputum eosinophils, but blood eosinophils are now being accepted as an alternative

**Figure 3** Validation test of clusters maximising treatment differences based on the SAL/fluticasone combination cluster analysis model (total fluticasone furoate/vilanterol trial population). The model adjusted for baseline forced expiratory volume in 1 s % predicted, reversibility status (yes/no for  $\geq 12\%$  improvement and  $\geq 200$  mL) and investigator region as a random effect. Numbers in parentheses are 95% CIs. RR, rate ratio.



measurement that is more practical to use in a clinical setting.<sup>27 28</sup> The clinical relevance of the 2.4% cut-off in blood eosinophil levels requires additional exploration; however, it lies between the 3% value commonly used to define sputum eosinophilia<sup>8 22 29</sup> and the 2% value that was identified as a relevant cut-off point in patients with COPD based on an analysis of the Evaluation of COPD Longitudinally to Identify Predictive Surrogate End points (ECLIPSE) cohort.<sup>30</sup> No minimal clinically important difference has yet been defined for eosinophilia in COPD,<sup>29</sup> but clinically meaningful differences have been found in other studies and analyses based on either a 2% or 2.5% cut-off.<sup>3 16 20 23</sup> In particular, this work complements the post hoc subgroup analysis of the same clinical trials of FF/VI versus VI that was conducted by Pascoe *et al*.<sup>23</sup> That analysis found that patients with COPD with blood eosinophils  $\geq 2\%$  at baseline experienced a reduction in exacerbation rates of 29% ( $p < 0.001$ ) when treated with FF/VI (all strengths) versus VI alone; the corresponding reduction in patients with eosinophils  $< 2\%$  was 10% ( $p = 0.283$ ). This analysis additionally found progressively greater exacerbation reduction when eosinophils were stratified into  $< 2\%$ , 2% to  $< 4\%$ , 4% to  $< 6\%$  and  $\geq 6\%$ , which is consistent with the generally linear trend that we have observed. The results from the present analysis confirm these findings while using a different and data-driven methodology, reinforcing eosinophil levels as an important biomarker of ICS treatment response. Owing to the generally linear relationship between eosinophils and treatment, the recursive partitioning algorithm split for the data around the median, which may not be a true inflection point in the data; although the split chosen in the data is consistent with the existing literature on eosinophils. Also of note, due to the cluster analysis methodology and the correlation between eosinophil per cent and eosinophil count in the data set, eosinophil per cent was kept as a variable on which the algorithm could split the tree rather than count. Eosinophil per cent was kept based on fewer missing values relative to count<sup>23</sup> and use in previous analyses.<sup>3 16 21 23</sup> As a sensitivity analysis, eosinophils were added as an explanatory variable in the base modelling and the tree was not allowed to split on them. The resulting tree was based on smoking history, with the split at 46 pack-years (data not shown). The group of non-responders from this analysis (ie, patients with lower eosinophil levels and a smoking history of  $> 46$  pack-years) has been previously identified.<sup>30</sup> Further investigation of patients with lower eosinophils and variable smoking history is needed.

If these findings regarding blood eosinophils being predictive of ICS treatment response can be validated externally, they may provide important information on disease phenotypes that benefit most from the addition of ICS to LABA therapy and inform on the benefit–risk profile of ICS-containing regimens. Sputum eosinophilia has been identified in 20–40% of patients with COPD,<sup>8</sup> and a simple blood test may allow for

straightforward identification of ICS responders in clinical practice.

To be useful in clinical practice, external validation of these results, potentially through clinical trials, is required along with further investigations to identify the relevant cut-off points for blood eosinophil levels. In addition, the relevance of smoking history also warrants further investigation. Within asthma, the identification of blood eosinophils has provided a step forward for the identification of patients who benefit from the use of anti-IL-5 monoclonal antibody treatments, potentially opening the door to a new treatment option within asthma.<sup>8</sup> Such advances within COPD, particularly with regard to minimising exacerbations, may also contribute to more optimised disease management.

**Acknowledgements** The authors thank Hao Li, a former employee of GSK, for his contribution to the statistical analyses. Editorial support in the form of developing the first draft of this manuscript from a detailed outline and study report, assembling tables and figures, collating author comments and editorial comments to drafts, copyediting, fact checking, referencing and graphic services was provided by Rhiannon Owen, BSc (Hons) at Gardiner-Caldwell Communications (Macclesfield, UK); this support was funded by GSK.

**Contributors** All authors contributed substantially to the development of the manuscript and approved the final version. DRH, RLD, HVL and SP have made contributions to conception and design as well as data interpretation. DRH, RLD and HVL involved in data analysis.

**Funding** This work was supported by GSK, study number WEUSKOP6069. It describes a cluster analysis using results from two clinical trials that were sponsored by GSK, study numbers HZC102871 (clinicaltrials.gov registration: NCT01009463) and HZC102970 (NCT01017952).

**Competing interests** DRH, RLD, HVL and SP were employees of, and held stock in, GSK, the study sponsor, at the time these analyses were conducted.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data sharing statement** No additional data are available.

**Open Access** This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>

## REFERENCES

1. De Soyza A, Calverley PM. Large trials, new knowledge: the changing face of COPD management. *Eur Respir J* 2015;45:1692–703.
2. Global Initiative for Chronic Obstructive Lung Disease (GOLD). Global strategy for the diagnosis, management and prevention of COPD: 2015 update to guidelines. <http://www.goldcopd.org/> (accessed 23 Mar 2015).
3. Aaron SD. Management and prevention of exacerbations of COPD. *BMJ* 2014;349:g5237.
4. Agarwal R, Aggarwal AN, Gupta D, *et al*. Inhaled corticosteroids vs placebo for preventing COPD exacerbations: a systematic review and meta-regression of randomized controlled trials. *Chest* 2010;137:318–25.
5. Crim C, Calverley PM, Anderson JA, *et al*. Pneumonia risk in COPD patients receiving inhaled corticosteroids alone or in combination: TORCH study results. *Eur Respir J* 2009;34:641–7.
6. Dransfield MT, Bourbeau J, Jones PW, *et al*. Once-daily inhaled fluticasone furoate and vilanterol versus vilanterol only for prevention of exacerbations of COPD: two replicate double-blind, parallel-group, randomised controlled trials. *Lancet Respir Med* 2013;1:210–23.
7. Ferguson GT, Anzueto A, Fei R, *et al*. Effect of fluticasone propionate/salmeterol (250/50 microg) or salmeterol (50 microg) on COPD exacerbations. *Respir Med* 2008;102:1099–108.



8. Saha S, Brightling CE. Eosinophilic airway inflammation in COPD. *Int J Chron Obstruct Pulmon Dis* 2006;1:39–47.
9. Sharafkhaneh A, Southard JG, Goldman M, *et al*. Effect of budesonide/formoterol pMDI on COPD exacerbations: a double-blind, randomized study. *Respir Med* 2012;106:257–68.
10. Szafranski WW, Cukier A, Ramirez A, *et al*. Efficacy and safety of budesonide/formoterol in the management of chronic obstructive pulmonary disease. *Eur Respir J* 2003;21:74–81.
11. DeCoster DA, Jones M. Tailoring of corticosteroids in COPD management. *Curr Respir Care Rep* 2014;3:121–32.
12. Weatherall M, Shirtcliffe P, Travers J, *et al*. Use of cluster analysis to define COPD phenotypes. *Eur Respir J* 2010;36:472–4.
13. DiSantostefano RL, Li H, Rubin DB, *et al*. Which patients with chronic obstructive pulmonary disease benefit from the addition of an inhaled corticosteroid to their bronchodilator? A cluster analysis. *BMJ Open* 2013;3:e001838.
14. Breiman L. *Classification and regression trees*. New York, NY, USA: Chapman and Hall, 1983.
15. Green RH, Pavord I. Stability of inflammatory phenotypes in asthma. *Thorax* 2012; 67:665–7.
16. Bafadhel M, McKenna S, Terry S, *et al*. Blood eosinophils to direct corticosteroid treatment of exacerbations of chronic obstructive pulmonary disease: a randomized placebo-controlled trial. *Am J Respir Crit Care Med* 2012;186:48–55.
17. Bathoorn E, Liesker JJ, Postma DS, *et al*. Anti-inflammatory effects of combined budesonide/formoterol in COPD exacerbations. *COPD* 2008;5:282–90.
18. Brightling CE, Monteiro W, Ward R, *et al*. Sputum eosinophilia and short-term response to prednisolone in chronic obstructive pulmonary disease: a randomised controlled trial. *Lancet* 2000;356:1480–5.
19. Brightling CE, McKenna S, Hargadon B, *et al*. Sputum eosinophilia and the short term response to inhaled mometasone in chronic obstructive pulmonary disease. *Thorax* 2005;60:193–8.
20. Eltboli O, Brightling CE. Eosinophils as diagnostic tools in chronic lung disease. *Expert Rev Respir Med* 2013;7:33–42.
21. Kitaguchi Y, Komatsu Y, Fujimoto K, *et al*. Sputum eosinophilia can predict responsiveness to inhaled corticosteroid treatment in patients with overlap syndrome of COPD and asthma. *Int J Chron Obstruct Pulmon Dis* 2012;7:283–9.
22. Leigh R, Pizzichini MM, Morris MM, *et al*. Stable COPD: predicting benefit from high-dose inhaled corticosteroid treatment. *Eur Respir J* 2006;27:964–71.
23. Pascoe S, Locantore N, Dransfield MT, *et al*. Blood eosinophil counts, exacerbations and response to the addition of inhaled fluticasone furoate to vilanterol in patients with chronic obstructive pulmonary disease: a secondary analysis of data from two parallel randomised controlled trials. *Lancet Respir Med* 2015;3:435–42.
24. Cowan DC, Taylor DR, Peterson LE, *et al*. Biomarker-based asthma phenotypes of corticosteroid response. *J Allergy Clin Immunol* 2015;135:877–83.e1.
25. Fleming L, Wilson N, Regamey N, *et al*. Use of sputum eosinophil counts to guide management in children with severe asthma. *Thorax* 2012;67:193–8.
26. Green RH, Brightling CE, McKenna S, *et al*. Asthma exacerbations and sputum eosinophil counts: a randomised controlled trial. *Lancet* 2002;360:1715–21.
27. Eltboli O, Mistry V, Barker B, *et al*. Relationship between blood and bronchial submucosal eosinophilia and reticular basement membrane thickening in chronic obstructive pulmonary disease. *Respirology* 2015;20:667–70.
28. Hillas G, Loukides S, Kostikas K, *et al*. Biomarkers obtained by non-invasive methods in patients with COPD: where do we stand, what do we expect? *Curr Med Chem* 2009;16:2824–38.
29. Gibson PG. Why inflammatory phenotyping is necessary for successful drug evaluation in asthma and COPD. *Eur Respir J* 2013;42:891–2.
30. Singh D, Kolsum U, Brightling CE, *et al*. Eosinophilic inflammation in COPD: prevalence and clinical characteristics. *Eur Respir J* 2014;44:1697–700.