The severity of airways obstruction as a determinant of treatment response in COPD

Peter MA Calverley[1]
Romain A Pauwels[2,*]
Paul W Jones[3]
Julie A Anderson[4]
Jørgen Vestbo[5]

[1]Clinical Research Centre, University Hospital Aintree, Liverpool, UK; [2]Department of Pulmonary Disease, University Hospital, Ghent, Belgium; [3]Department of Physiological Medicine, St George's Hospital, London, UK; [4]Biomedical Data Sciences, GlaxoSmithKline Research and Development, Greenford, UK; [5]North West Lung Centre, Wythenshawe Hospital, Manchester, UK

*Sadly, Professor Pauwels died during the preparation of this manuscript.

Abstract: Guidelines recommend that patients with COPD are stratified arbitrarily by baseline severity (FEV_1) to decide when to initiate combination treatment with a long-acting β_2-agonist and an inhaled corticosteroid. Assessment of baseline FEV_1 as a continuous variable may provide a more reliable prediction of treatment effects. Patients from a 1-year, parallel-group, randomized controlled trial comparing 50 µg salmeterol (Sal), 500 µg fluticasone propionate (FP), the combination (Sal/FP) and placebo, (bid), were categorized post hoc into FEV_1 <50% and FEV_1 ≥50% predicted subgroups (n=949/513 respectively). Treatment effects on clinical outcomes – lung function, exacerbations, health status, diary card symptoms, and adverse events – were investigated. Treatment responses based on a pre-specified analysis explored treatment differences by severity as a continuous variable. Lung function improved with active treatment irrespective of FEV_1; Sal/FP had greatest effect. This improvement appeared additive in milder disease; synergistic in severe disease. Active therapy significantly reduced exacerbation rate in patients with FEV_1 <50% predicted, not in milder disease. Health status and breathlessness improved with Sal/FP irrespective of baseline FEV_1; adverse events were similar across subgroups. The spirometric response to Sal/FP varied with baseline FEV_1, and clinical benefits were not restricted to patients with severe disease. These data have implications for COPD management decisions, suggesting that arbitrary stratifications of baseline severity are not necessarily indicative of treatment efficacy and that the benefits of assessing baseline severity as a continuous variable should be assessed in future trials.

Keywords: chronic obstructive pulmonary disease, FEV_1, inhaled corticosteroid, long-acting β_2-agonist, subgroups

Introduction

Patients with COPD are characterized by a reduced FEV_1 and a tendency to experience symptomatic exacerbations and health status impairment (Calverley and Walker 2003). Not all of these problems are present to the same degree at all stages of the illness, with exacerbations that require treatment occurring more frequently as lung function deteriorates. Although there is a statistically significant relationship between health status and the degree of FEV_1 impairment, individual confidence intervals for any particular percentage-predicted FEV_1 are wide.

There is now clear evidence that currently prescribed inhaled drugs, whether long-acting β_2-agonists (LABAs) or inhaled corticosteroids (ICS), alone or in combination, have beneficial effects in stable COPD (Mahler et al 2002; Calverley et al 2003a, 2003b; Szafranski et al 2003). While guidelines recommend that combined use of these drugs be reserved for advanced patients (BTS 1997; American Thoracic Society/European Respiratory Society Task Force 2004; GOLD 2005), it is not clear whether such treatment is equally effective at all stages of disease severity. Furthermore, whether individual treatment outcomes, such as lung function, exacerbation rate, and health status, differ as FEV_1 worsens is also uncertain.

Correspondence: Professor PMA Calverley
Clinical Science Centre, University Hospital Aintree, Longmoor Lane, Liverpool L9 7AL, UK
Tel + 44 151 529 5886
Fax + 44 151 529 5888
Email pmacal@liverpool.ac.uk

The most frequently used method of classifying severity is to apply threshold values of FEV_1. Several expert groups recommend the separation of disease based on a predicted FEV_1 threshold level of 50% (BTS 1997; American Thoracic Society/European Respiratory Society Task Force 2004; National Collaborating Centre for Chronic Conditions 2004; GOLD 2005). Such classifications are somewhat arbitrary and may not accurately represent clinical response at all points on the continuum of disease severity. Perhaps a less judgmental approach would be to treat FEV_1 as a continuous variable and relate this to the subsequent treatment response. Given the different drugs used in COPD therapy and the several different outcomes to be examined, a large number of patients are required to address these issues.

The Trial of Inhaled Steroids and Long-acting β_2-Agonists (TRISTAN) was a one-year, double-blind, placebo-controlled study of patients with stable COPD (Calverley et al 2003b). Patients with a range of FEV_1 severities (25%–70% predicted) were randomized to receive either salmeterol (Sal) or fluticasone propionate (FP) alone or in a fixed-dose combination or an identical placebo, all twice daily. The results showed that Sal/FP significantly improved pretreatment FEV_1, health status, and daily symptoms after 12 months compared with placebo or monotherapy alone. The effect of Sal/FP in reducing the rate of exacerbations compared with placebo was greater in patients with FEV_1 <50% predicted than in those with FEV_1 ≥50%. However, the relationship between baseline severity and outcomes in TRISTAN was not fully analyzed by the time of the publication. In addition, data on the trend of effect along a continuous FEV_1 variable was outside the scope of the original publication, data that would be a useful addition to the clinical evidence base.

In this new, exploratory analysis, we have used data from TRISTAN to test the hypothesis that the severity of airflow obstruction, as reflected by the pretreatment FEV_1, is an important determinant of the subsequent change in the specified outcomes of treatment. To do this, post hoc analyses of the TRISTAN population were conducted to provide new data, whereby patients were categorized according to the arbitrary FEV_1 threshold, and the potential effect of treatment on a variety of clinical outcomes assessed. In addition, treatment responses were evaluated by a pre-specified analysis, in which FEV_1 was assessed as a continuous variable. It was anticipated that these approaches would clarify the validity of FEV_1 in defining treatment response in COPD.

Methods

Full details of the study methodology, patient selection, and outcomes for the intention-to-treat (ITT) analysis have been presented previously (Calverley et al 2003b).

Subjects

Briefly, we recruited 1465 outpatients with COPD aged 40–79 years who were current or ex-smokers with at least a 10-pack-year history, had an initial FEV_1 between 25% and 70% predicted, and who showed limited bronchodilator reversibility. All patients had a history of daily cough with sputum, reported previous exacerbations that required treatment, and fulfilled the diagnostic criteria for COPD both clinically and spirometrically, as defined elsewhere (BTS 1997; GOLD 2005).

We obtained approval from local ethics committees at each participating site, and all patients provided written informed consent.

Experimental design

Patients entered a 2-week run-in period during which any patient using ICS or LABAs had these medicines discontinued. The use of inhaled salbutamol as a relief medication, and regular treatment with anticholinergics, mucolytics, and/or theophylline was permitted throughout the study. Patients clinically stable at the end of the run-in were randomized to one of the following treatments inhaled from a dry powder Diskus device (GlaxoSmithKline Inc., Research Triangle Park, NC, USA), twice daily for the subsequent 52 weeks: a combination of 50 μg Sal and 500 μg FP, 50 μg Sal or 500 μg FP alone, or placebo.

Assessments

Patient evaluations were conducted at weeks 0, 2, 4, 8, 16, 24, 32, 40, and 52. At each visit, spirometry was recorded before use of salbutamol or the morning study medication. The number of COPD exacerbations – defined as episodes of symptomatic worsening that had required medical treatment with antibiotics and/or oral corticosteroids or an emergency hospital visit – that occurred since the previous visit was also noted.

Health status was assessed by the St George's Respiratory Questionnaire (SGRQ) (Jones et al 1992). Additionally, a daily diary record was kept to record the number of times relief medication was used each day and the number of times the patient woke from sleep. Symptoms were scored as: breathlessness, 0 (none) to 4 (breathless at

rest); cough, 0 (none) to 3 (severe); and sputum production, 0 (no sputum) to 4 (dark yellow–green). Adverse events were noted at each clinic visit by recording spontaneously reported complaints from patients, by asking about potentially treatment-related problems and by specifically examining the throat for candidiasis and the forearm for spontaneous bruising.

Statistical analysis

Data are expressed as means and 95% confidence intervals unless otherwise stated.

The analysis population and methods were defined and written in the trial Data Analysis Plan prior to the unblinding of the treatments for post-hoc analysis. Covariates used for analysis, where applicable, were age, sex, country, smoking status, baseline % predicted FEV_1, and baseline value of the response. The study was not powered to detect interactions, so in order to check for significant interactions, it was more conservative to test at a lower significance level. Therefore, the interactions between treatment and baseline were tested for statistical significance at the 0.10 level in the ITT population. This testing was performed when analyzing the primary efficacy variable, clinic FEV_1, prior to use of salbutamol, and the secondary efficacy variable of the number of moderate and/or severe COPD exacerbations. For all pair-wise comparisons, the null hypothesis was that of no treatment effect.

Pre-bronchodilator FEV_1 was analyzed using repeated measures analysis (Brown and Prescot 1999). Time was included as a categorical variable and an unstructured variance–covariance matrix was fitted with SAS PROC MIXED software version 6.12. These methods were also used to analyze SGRQ.

The number of exacerbations was analyzed by a maximum likelihood Poisson regression with the amount of time a patient had received treatment as an offset variable. The time to first exacerbation and time to withdrawal were analyzed using Cox's proportional hazards model (Cox 1972).

For the use of rescue medication, the median data for weeks 1–52 were analyzed using the van Elteren extension to the Wilcoxon rank sum test (Lehmann 1975; van Elteren 1960) and stratified by smoking status, and the confidence limits were calculated with the Hodges–Lehmann method (Hollander and Wolfe 1973).

Analyses by baseline severity were conducted for the following parameters: pre-bronchodilator FEV_1, exacerbations, health status (using the SGRQ), and diary card symptoms.

Two approaches were used to assess the influence of FEV_1 on treatment response. In the first method, which was pre-specified but not fully analyzed at the time of publication, a continuous variable for baseline % predicted pre-bronchodilator FEV_1 was included in parametric statistical analyses (eg, FEV_1 or exacerbations). This type of analysis was pre-specified in the study data analysis plan prior to unblinding for the major efficacy variables as a check that the treatment effects held true for a wide range of severities. The interaction between baseline severity and treatment was investigated using a model including a baseline severity-by-treatment interaction term, and predicted values were obtained from the model baseline % predicted FEV_1 values of 33%, 44%, and 55% (these were the quartile values which had been pre-specified in the data analysis plan and were essentially arbitrary). For measures requiring non-parametric analysis, such as diary card symptoms, this approach was inappropriate.

In the second method, undertaken post hoc, the subjects were categorized into two groups: those with pre-bronchodilator FEV_1 <50% predicted at baseline and those with 50% or more. Where any parametric statistical analyses were performed (eg, SGRQ [Jones et al 1992]), this factor of severity was included in the model, along with an interaction term for treatment-by-severity. This allowed for the possibility that the two severity groups had different treatment effects. The least squares means and treatment differences were produced from this model containing the interaction term. Where non-parametric analyses were performed, the subgroups were analyzed separately.

Results

Demographics and baseline characteristics

The demographic and baseline characteristics according to baseline severity are given in Table 1. As anticipated, patients in the FEV_1 <50% subgroup had worse health status, were more likely to use inhaled medication before randomization, and were less likely to be current smokers than those in the FEV_1 ≥50% subgroup. However, the number of pack-years smoked and the degree of bronchodilator reversibility were comparable between the two populations.

Following randomization to treatment, more patients in the FEV_1 <50% subgroup withdrew than in the FEV_1 ≥50%

Table 1 Demographic and baseline characteristics by baseline FEV_1 severity

	Total population (n=1465)	FEV_1 <50% predicted (n=949)	FEV_1 ≥50% predicted (n=513)
Male	1060 (72%)	714 (75%)	343 (67%)
Age	63 (9)	64 (8)	62 (9)
ICS at screen	630 (51%)	525 (55%)	226 (44%)
LABAs at screen	506 (41%)	413 (44%)	178 (35%)
Current smokers	746 (51%)	456 (48%)	288 (56%)
Pack-years	43 (22)	43 (22)	41 (22)
SGRQ total score	48.2 (16.3)	50.7 (15.8)	43.5 (16.3)
Pre-bronchodilator FEV_1 (L)	1.27 (0.48)	1.03 (0.31)	1.71 (0.40)
Pre-bronchodilator FVC (L)	2.47 (0.79)	2.23 (0.71)	2.90 (0.76)
% predicted FEV_1	44.6 (13.9)	36.1 (8.1)	60.3 (6.9)
% reversibility in predicted FEV_1	3.8 (4.4)	4.0 (4.4)	3.5 (4.3)

NOTES: Data are number (%) or mean (SD) unless otherwise indicated. Three patients have missing baseline FEV_1, so are not included in either subgroup.
Abbreviations: FVC, forced vital capacity; ICS, inhaled corticosteroids; LABAs, long-acting β_2-agonists; SGRQ, St George's Respiratory Questionnaire.

subgroup (Figure 1). Moreover, fewer patients receiving combination treatment withdrew than those on placebo or individual components alone, irrespective of baseline FEV_1 severity (Figure 1). The main reason for withdrawals across both subgroups and all treatment arms was the presence of adverse events, particularly exacerbations of COPD.

Lung function

The ICS and LABA combination produced the greatest improvements in lung function compared with placebo and individual components alone in both severity subgroups (all p<0.001; Table 2). This was particularly apparent in the FEV_1 <50% subgroup, where the treatment effect was twice the sum of individual components (Table 2). Treatment responses following monotherapy were less consistent. While Sal alone led to significant improvements (p≤0.02) in both severity populations, FP was only significantly effective (p<0.001) in patients with less severe disease.

Overall, no statistically significant interaction (p=0.102) between baseline FEV_1 and treatment response was observed, indicating that treatment effects between subgroups were comparable.

Similar effects were observed when FEV_1 was analyzed as a continuous variable (Table 2). Patients who received combination therapy showed larger improvements in lung function compared with those receiving placebo and individual components alone, irrespective of their baseline FEV_1. This response was most evident at the 25% percentile, in patients with particularly low lung function (FEV_1 33% predicted), where the treatment effect following Sal/FP combination therapy was twice the sum of individual components (Table 2). In contrast, FP monotherapy produced progressively smaller changes in FEV_1 as baseline FEV_1 declined. Nonetheless, no significant treatment interactions by severity were noted across the whole data set (p=0.147).

Exacerbations

In the post hoc subgroup analyses a higher proportion of patients experienced an exacerbation (60%) in the severe population (FEV_1 <50% predicted) as compared with those in the FEV_1 ≥50% predicted subgroup (44%).

All active treatments reduced the number of exacerbations relative to placebo, with proportionately

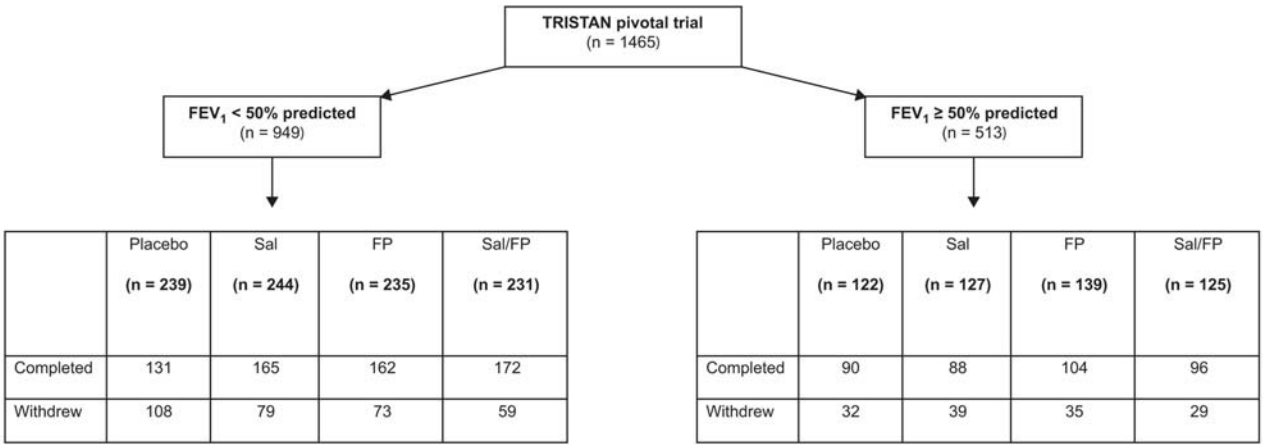

Figure 1 Study withdrawals by baseline severity and therapy group.
Abbreviations: FP 500, fluticasone propionate 500 μg; Sal 50, salmeterol 50 μg; Sal/FP 50/500, salmeterol/fluticasone propionate 50/500 μg.

Table 2 Effect of 52 weeks' treatment on pre-bronchodilator FEV_1 according to baseline severity and therapy group

Analysis	Population (n)	Placebo (353)	Sal/FP (345)	Sal (361)	FP (371)
Categorical	*FEV_1 <50% predicted*				
	Patients (n)	234	226	238	233
	Mean baseline FEV_1(L)	1.02	1.03	1.03	1.03
	Treatment difference[a] vs placebo (95% CI)	-	110 (75, 145)	42 (7, 76)	18 (–16, 53)
	Treatment difference vs Sal/FP (95% CI)	-	-	69 (34, 103)	92 (57, 126)
	p-value vs placebo	-	<0.001	0.017	0.296
	p-value vs Sal/FP	-	-	<0.001	<0.001
	FEV_1 ≥50% predicted				
	Patients (n)	119	119	123	138
	Mean baseline FEV_1(L)	1.73	1.84	1.65	1.65
	Treatment difference vs placebo (95% CI)	-	176 (128, 225)	96 (48, 143)	79 (33, 126)
	Treatment difference vs Sal/FP (95% CI)	-	-	81 (33, 129)	97 (51, 144)
	p-value vs placebo	-	<0.001	<0.001	<0.001
	p-value vs Sal/FP	-	-	<0.001	<0.001
Continuous	*At 25th percentile (33% predicted)*				
	Treatment difference vs placebo (95% CI)	-	113 (77, 149)	38 (2, 74)	18 (–18, 54)
	Treatment difference vs Sal/FP (95% CI)	-	-	75 (39, 110)	94 (58, 130)
	p-value vs placebo	-	<0.001	0.037	0.316
	p-value vs Sal/FP	-	-	<0.001	<0.001
	At median (44% predicted)				
	Treatment difference vs placebo (95% CI)	-	132 (104, 160)	61 (33, 89)	40 (13, 67)
	Treatment difference vs Sal/FP (95% CI)	-	-	71 (43, 99)	92 (64, 119)
	p-value vs placebo	-	<0.001	<0.001	0.004
	p-value vs Sal/FP	-	-	<0.001	<0.001
	At 75th percentile (55% predicted)				
	Treatment difference vs placebo (95% CI)	-	151 (116, 186)	84 (49, 119)	62 (27, 96)
	Treatment difference vs Sal/FP (95% CI)	-	-	67 (32, 102)	89 (55, 124)
	p-value vs placebo	-	<0.001	<0.001	<0.001
	p-value vs Sal/FP	-	-	<0.001	<0.001

[a]Treatment difference=mL.

Abbreviations: FP, fluticasone propionate; Sal, salmeterol; Sal/FP, salmeterol and fluticasone propionate combination.

greater effects in those episodes requiring treatment with oral corticosteroids. With all treatments, reductions in exacerbation rates compared with placebo were statistically significant in subjects with more severe COPD (FEV_1 <50% predicted), but not significant in those with FEV_1 ≥50% predicted (Table 3). Exacerbation rates were consistently lower with combination therapy than with Sal or FP monotherapy in all subgroups, although these effects did not reach statistical significance (Table 3). Moreover, no significant treatment interactions by severity group for all exacerbations (p=0.410) or exacerbations requiring oral corticosteroids (p=0.675) were noted.

Again, broadly similar results were seen when the estimated annual rates of exacerbation were analyzed according to baseline severity as a continuous variable (Figure 2). As baseline FEV_1 declined, more exacerbations occurred, although there was no significant treatment interaction by severity (p=0.139). All treatments significantly reduced the rate of exacerbations at the arbitrary FEV_1 44% and 33% predicted quartiles, but only Sal monotherapy produced a significant reduction within the less severe 55% predicted quartile. With regard to comparisons between active treatments, the reduction in exacerbation rate seen with combination treatment was significantly greater than that seen with Sal alone at the FEV_1 33% predicted quartile (p=0.043). However, no significant differences were seen between combination treatment and FP alone.

There was a delay in the time to first exacerbation with the active treatments compared with placebo in the more severe subgroup, this delay being particularly apparent in patients receiving combination treatment (median 109 days vs 47 days; p=0.003). The time to first exacerbation was slightly longer in the less severe population with all active

treatments compared with placebo, but this was not statistically significant (Figure 3).

Health status and symptoms

The largest improvement in health status compared with placebo, indicated by a reduction in SGRQ score, occurred in patients receiving combination therapy. There was no relationship between treatment effect on health status and baseline spirometry (p=0.598). The treatment effect compared with placebo reached statistical significance only in patients with baseline FEV_1 <50% who were treated with combination therapy (–2.3 units, 95% CI –3.7 to –0.8, p=0.002). In the less severe subgroup, there was a similar improvement in health status with combination therapy which, although not statistically significant, approached significance (–1.9 units, 95% CI –3.9 to 0.1, p=0.058) (Table 4).

There were no consistent effects of baseline FEV_1 severity on diary card symptom scores. Diary card data in both severity subgroups showed that patients receiving combination treatment were significantly less likely to use relief medication than those who received placebo (p≤0.001), Sal alone (p<0.03), or FP alone (p≤0.004). Combination treatment was also more effective than placebo (p<0.004) in improving breathlessness in both subgroups, night-time awakenings in the more severe subgroup (p=0.029), and cough score in the less severe population (p=0.001). Overall, there was no clear evidence of treatment differences between the two severity subgroups.

Table 3 Effect of 52 weeks' treatment on exacerbation rate according to baseline severity and therapy group

	Placebo	Sal/FP	Sal	FP
	All exacerbations			
FEV_1 <50% predicted				
Patients (n)[a]	239	231	244	235
Exacerbation rate based on Poisson model[b]	1.42	0.99	1.09	1.07
Treatment difference vs placebo (95% CI)[c]	-	30% (17, 41)	23% (9, 35)	24% (10, 36)
Treatment difference vs Sal/FP (95% CI)[d]	-	-	9% (–8, 24)	8% (–10, 22)
p-value vs placebo	-	< 0.001	0.002	0.001
p-value vs Sal/FP	-	-	0.266	0.378
FEV_1 ≥50% predicted				
Patients (n)	122	125	127	139
Exacerbation rate based on Poisson model[b]	1.09	0.98	0.98	1.05
Treatment difference vs placebo (95% CI)[c]	-	10% (–22, 33)	10% (–21, 33)	3% (–29, 37)
Treatment difference vs Sal/FP (95% CI)[d]	-	-	–1% (–37, 26)	7% (–25, 30)
p-value vs placebo	-	0.511	0.488	0.831
p-value vs Sal/FP	-	-	0.975	0.633
	Exacerbations requiring oral corticosteroids			
FEV_1 <50% predicted				
Patients (n)	239	231	244	235
Exacerbation rate based on Poisson model[b]	0.81	0.47	0.58	0.52
Treatment difference vs placebo (95% CI)[c]	-	43% (29, 54)	29% (13, 42)	35% (20, 48)
Treatment difference vs Sal/FP (95% CI)[d]	-	-	19% (–1, 35)	11% (–12, 29)
p-value vs placebo	-	<0.001	<0.001	<0.001
p-value vs Sal/FP	-	-	0.059	0.323
FEV_1 ≥50% predicted				
Patients (n)	122	127	139	125
Exacerbation rate based on Poisson model[b]	0.66	0.50	0.47	0.47
Treatment difference vs placebo (95% CI)[c]	-	24% (–17, 51)	29% (–9, 54)	29% (–6, 53)
Treatment difference vs Sal/FP (95% CI)[d]	-	-	–7% (–68, 33)	–7% (–65, 31)
p-value vs placebo	-	0.116	0.094	0.205
p-value vs Sal/FP	-	-	0.785	0.759

[a]n=number of patients in the active groups.
[b]Exacerbation rate defined as the mean number of exacerbations per year from Poisson model.
[c]Treatment difference represents the percentage reduction in exacerbation rate vs placebo.
[d]Treatment difference represents the percentage change for Sal/FP vs individual components (negative value represents a reduction in exacerbations for component).
Abbreviations: FP, fluticasone propionate; Sal/FP, salmeterol and fluticasone propionate combination; Sal, salmeterol.

Safety

All treatments were well tolerated. The frequency of adverse events was comparable between the severity subgroups (74%–81% and 67%–81%, across <50% and ≥50% subgroups, respectively). Morning serum cortisol concentrations were lower in patients who had received active treatment compared with those on placebo, and this was statistically significant in patients (at 24 weeks) who had received Sal/FP (p=0.04) and FP (p=0.01; at 52 weeks). There was no increase in the incidence of bruise counts in either of the subgroups following active treatment compared with placebo, with more than 98% of patients in both subgroups and across treatments reporting no spontaneous bruising. The frequency of patients in both subgroups and across treatment groups who experienced hoarseness and cough, predictable side effects associated with ICS use, was found to be 1%–4% and 1%–5%, respectively.

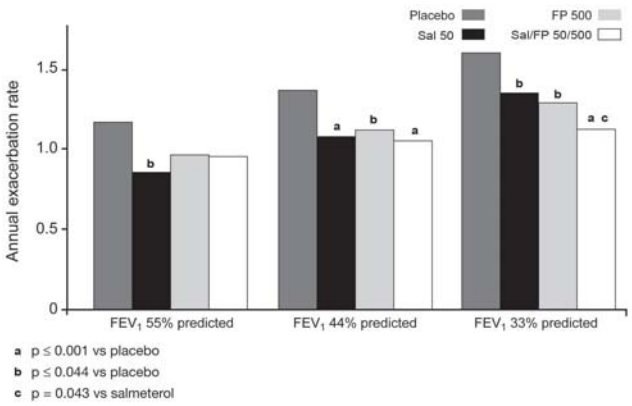

Figure 2 Estimated annual rates of moderate–severe exacerbations by baseline severity and therapy group.
Abbreviations: FP 500, fluticasone propionate 500 μg; Sal 50, salmeterol 50 μg; Sal/FP 50/500, salmeterol/fluticasone propionate 50/500 μg.

Discussion

Although randomized controlled trials remain our primary source of evidence when choosing treatment, diseases like COPD where there is pathological heterogeneity among patients and concerns about the influence of disease severity on response to treatment, are well suited to a subgroup analysis of large data sets (Rothwell 2005). To be valid such an analysis should consider a limited number of subgroups, be pre-defined, and report comparisons primarily in terms of treatment by subgroup interaction as was the case in our study (Rothwell 2005). Current treatment recommendations in COPD use consensus-based thresholds of spirometric severity (BTS 1997; GOLD 2005) and subsequent studies have either reported subgroup analyses based on these (Jones et al 2003) or restricted recruitment of patients according to these criteria (Calverley et al 2003a; Szafranski et al 2003). Unsurprisingly in the light of previous data suggesting ICS treatment has a greater effect in more severe COPD (Burge et al 2000), subgroup analysis to identify potentially responsive patients was requested by regulators following the ITT report of the TRISTAN data. In this new analysis, we found that most treatment effects could not be predicted

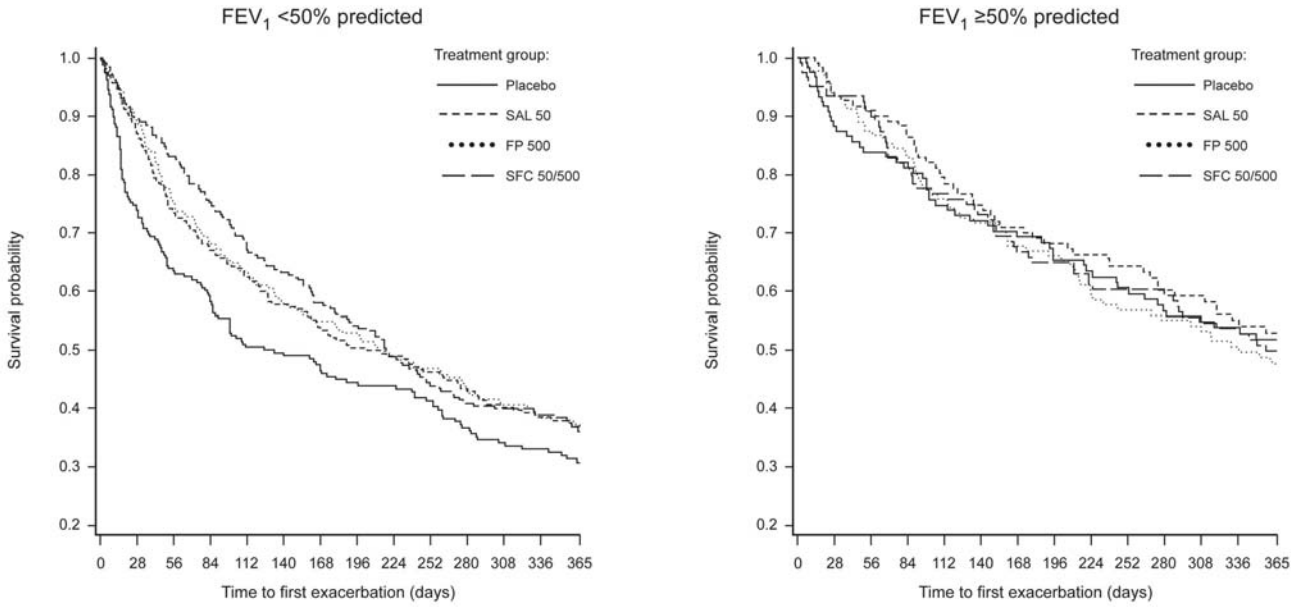

Figure 3 Time to first exacerbation by baseline severity and therapy group.
Abbreviations: FP 500, fluticasone propionate 500 μg; Sal 50, salmeterol 50 μg; Sal/FP 50/500, salmeterol/fluticasone propionate 50/500 μg.

Table 4 Effect of 52 weeks' treatment on SGRQ total score according to baseline severity and therapy group

Severity subgroup	Placebo	Sal/FP	Sal	FP
FEV_1 <50% predicted				
Patients (n)	211	205	212	219
Mean baseline SGRQ	49.8	50.2	51.3	51.1
Adjusted mean change in SGRQ	–1.3	–3.6	–2.6	–1.7
Treatment difference vs placebo (95% CI)[a]	-	–2.3 (–3.7, –0.8)	–1.3 (–2.8, 0.1)	–0.5 (–1.9, 1.0)
Treatment difference vs Sal/FP (95% CI)[b]	-	-	–0.9 (–2.4, 0.5)	–1.8 (–3.2, –0.4)
p-value vs placebo	-	0.002	0.069	0.527
p-value vs Sal/FP	-	-	0.201	0.014
FEV_1 ≥50% predicted				
Patients (n)	107	115	108	121
Mean baseline SGRQ	41.0	41.3	43.0	47.4
Adjusted mean change in SGRQ	–2.1	–4.0	–2.7	–3.5
Treatment difference vs placebo (95% CI)[a]	-	–1.9 (–3.9, 0.1)	–0.6 (–2.6, 1.4)	–1.4 (–3.3, 0.6)
Treatment difference vs Sal/FP (95% CI)[b]	-	-	–1.3 (–3.3, 0.7)	–0.6 (–2.5, 1.4)
p-value vs placebo	-	0.058	0.544	0.175
p-value vs Sal/FP	-	-	0.199	0.572

[a]Treatment difference – a negative value indicates superiority of active group (improvement in SGRQ).
[b]Treatment difference – a negative value indicates superiority of Sal/FP over individual components.
(A clinically relevant reduction in SGRQ total score is a decline of at least 4 points.)
Abbreviations: FP, fluticasone propionate; Sal/FP, salmeterol and fluticasone propionate combination; Sal, salmeterol; SGRQ, St George's Respiratory Questionnaire.

reliably using an arbitrary split of baseline FEV_1 <50% and ≥50%, although the effect of treatment on exacerbation frequency was most evident in those with worst initial lung function. Thus rigid adherence to specific thresholds may not be the best way to determine treatment in COPD, and we propose from this exploratory analysis that due consideration is given to define more tightly which groups of patients are likely to benefit from combination treatment, whichever agents are used.

While the TRISTAN trial was a relatively large study it was powered statistically only to show a difference between combination therapy and placebo treatment, making additional comparisons between the combination and its individual components statistically less robust. For this reason we have reported the modeled data analysis which was pre-specified but considered to be exploratory and hypothesis-generating in nature. It is important to note, however, that not all of the clinical outcomes were suited to this approach – for example, rescue medication was not; in this case, analysis based on categorization by the threshold approach was employed.

The spirometric changes with therapy were small, as would be expected in patients already selected for their limited bronchodilator reversibility. However, such changes are a reproducible marker of the effect of treatment on lung function and are a qualitative indicator of the likely effect of treatment on more complex physiological outcomes (O'Donnell et al 1999). Subgroup analysis revealed that treatment improved lung function irrespective of baseline FEV_1, with the largest changes observed when the bronchodilator Sal and the ICS FP were combined. On analyzing treatment responses based on FEV_1 as either a categorical or continuous variable, in patients with less severe disease, the use of combination treatment improved FEV_1 to a degree equivalent to the sum of the individual components. As severity of disease increased and baseline FEV_1 declined, the addition of the ICS to Sal appeared to be more than simply additive. These data may explain why combining treatment is relatively more effective in patients with the worst initial lung function.

The relationship seen between exacerbation rate and baseline lung function was in keeping with other studies (Jones et al 2003; Szafranski et al 2003). The lower exacerbation rate in patients with an FEV_1 ≥50% predicted following combination therapy was not statistically different from that observed with placebo, although when analyzing FEV_1 as a continuous variable, Sal alone appeared to be effective in this less severe population. However, it should be noted that despite the entry criteria, there were substantially fewer exacerbations than expected, with 46% of the total population not experiencing an exacerbation during the study. This significantly reduced the power of the study to detect differences, as the smaller number of patients in the FEV_1 ≥50% subgroup reduced the potential for the analysis to identify a statistically significant effect.

Even in the more severe subgroup, the exacerbation rate in the placebo arm (1.4 per patient per year) was lower than that seen in apparently similar patients receiving placebo in other 1-year studies (1.8–1.9 events per year) (Szafranski et al 2003; Calverley et al 2003a). This may reflect differences in patient recruitment between centers and the effect of increased patient supervision in reducing medical contacts due to the repeated clinic visits (Bourbeau et al 2003). Nonetheless, if the number of patients with an FEV_1 <50% predicted who need to be treated to prevent one exacerbation (NNT) is calculated, then the NNT is 2.4 per year of therapy if combination and placebo are compared using the observed exacerbation rate in the placebo group. If the exacerbation rate is used from the year prior to study entry (two exacerbations per patient), the NNT is 1.7 per year. Even if combination therapy is compared with Sal using the observed exacerbation rates, the NNT is 10 per year. These numbers are well within the range considered beneficial in most clinical specialties.

The relationship between health status and FEV_1 is much weaker than that between health status and exacerbations (Spencer et al 2004). This may explain why the health status effects seen in this study were unrelated to the initial spirometry, with a similar ranking order of effect between monotherapy and combination treatment in both severe and milder patients. Methodological factors might have reduced the power of the present study to show a clinically significant difference but there is no suggestion that baseline lung function would predict such a change.

Breathlessness (as measured by the Transitional Dyspnoea Index [Mahler et al 1984]) has been shown to improve with combination treatment but was not specifically assessed here (Mahler et al 2002). The changes in the daily diary card would be compatible with such an effect but there was no evidence that initial lung function predicted benefit in most of the symptoms, where this was reported. Encouragingly, there was no evidence of any worse adverse event profile in relation to baseline lung function.

COPD management guidelines try to offer balanced advice about the relative efficacy and hazards of current treatment, often having to use limited evidence to shape their recommendations. Their proposal to introduce ICS therapy in those with an FEV_1 <50% predicted is a reasonable one given the uncertainties about the longer-term safety of ICS in COPD, a point which may be resolved when prospective mortality data become available from studies such as TORCH (Vestbo 2004). However, this should not be confused with treatment having no effect in patients above this arbitrary threshold value. As our subgroup analyses demonstrated, all the therapies were effective, irrespective of the initial FEV_1, with Sal/FP producing the greatest benefit in a number of outcomes. These data support the approach outlined in the UK NICE recommendations (NICE 2004), where treatment is offered according to the presence of symptoms rather than at a specific FEV_1 value and may provide preliminary evidence for revision of baseline severity stratifications within current guidelines. While accepting the limitations of this study, this preliminary evidence may provide the basis for considering a revision of baseline severity stratifications within current guidelines. Thus, the potential benefits of assessing baseline severity as a continuous variable and the impact this may have on management decisions should be studied and validated in future, large clinical studies.

Acknowledgments

Funding for this study (protocol number SFCB3024) was provided by GlaxoSmithKline R&D.

References

American Thoracic Society/European Respiratory Society Task Force. 2004. Standards for the Diagnosis and Management of Patients with COPD [online]. Version 1.2. New York: American Thoracic Society. Accessed 30 April 2006.URL: http://www.thoracic.org/copd/.

Bourbeau J, Julien M, Maltais F, et al. 2003. Reduction of hospital utilization in patients with chronic obstructive pulmonary disease: a disease-specific self-management intervention. *Arch Intern Med*, 163:585–91.

[BTS] British Thoracic Society. 1997. British Thoracic Society COPD guidelines summary. *Thorax*, 52(Suppl 5):1–32.

Brown H, Prescot R. 1999. Applied mixed models in medicine. In Barnett V (ed). Repeated measures data. New York: J Wiley and Sons. p 199–259.

Burge PS, Calverley PM, Jones PW, et al. 2000. Randomised, double blind, placebo controlled study of fluticasone propionate in patients with moderate to severe chronic obstructive pulmonary disease: the ISOLDE trial. *Br Med J*, 320:1297–303.

Calverley PM, Boonsawat W, Cseke Z, et al. 2003a. Maintenance therapy with budesonide and formoterol in chronic obstructive pulmonary disease. *Eur Respir J*, 22:912–19.

Calverley P, Pauwels R, Vestbo J, et al. 2003b. Combined salmeterol and fluticasone in the treatment of chronic obstructive pulmonary disease: a randomised controlled trial. *Lancet*, 361:449–56.

Calverley PM, Walker P. 2003. Chronic obstructive pulmonary disease. *Lancet*, 362:1053–61.

Cox DR. 1972. Regression models and life-tables (with discussion). *J Royal Stat Soc Ser B*, 34:187–220.

[GOLD] Global Initiative for Chronic Obstructive Lung Disease. 2005. Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease - updated 2005. Modena: Global Initiative for Chronic Obstructive Lung Disease [online]. Accessed 30 April 2006. URL: www.goldcopd.com.

Hollander M & Wolfe DA. The two-sample location problem. In Nonparametric statistical methods, 1973; p 67–82. New York: John Wiley and Sons.

Jones PW, Quirk FH, Baveystock CM, et al. 1992. A self-complete measure of health status for chronic airflow limitation. The St. George's Respiratory Questionnaire. *Am Rev Respir Dis*, 145:1321–7.

Jones PW, Willits LR, Burge PS, et al. 2003. Disease severity and the effect of fluticasone propionate on chronic obstructive pulmonary disease exacerbations. *Eur Respir J*, 21:68–73.

Lehmann E. 1975. Nonparametrics: statistical methods based on ranks. San Francisco: Holden-Day. p 132–7.

Mahler DA, Weinberg DH, Wells CK, et al. 1984. The measurement of dyspnea. Contents, interobserver agreement, and physiologic correlates of two new clinical indexes. *Chest*, 85:751–8.

Mahler DA, Wire P, Horstman D, et al. 2002. Effectiveness of fluticasone propionate and salmeterol combination delivered via the Diskus device in the treatment of chronic obstructive pulmonary disease. *Am J Respir Crit Care Med*, 166:1084–91.

[NICE] National Collaborating Centre for Chronic Conditions. 2004. Chronic Obstructive Pulmonary Disease (NICE Guideline No. 12). *Thorax*, 59(Suppl 1):1–232.

O'Donnell DE, Lam M, Webb KA. 1999. Spirometric correlates of improvement in exercise performance after anticholinergic therapy in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med*, 160:542–9.

Rothwell PM. 2005. Treating individuals 2: subgroup analysis in randomised controlled trials: importance, indications, and interpretation. *Lancet*, 365:176–86.

Spencer S, Calverley PM, Burge PS, et al. 2004. Impact of preventing exacerbations on deterioration of health status in COPD. *Eur Respir J*, 23:698–702.

Szafranski W, Cukier A, Ramirez A, et al. 2003. Efficacy and safety of budesonide/formoterol in the management of chronic obstructive pulmonary disease. *Eur Respir J*, 21:74–81.

van Elteren P. 1960. On the combination of independent two-sample tests of Wilcoxon. *Bull Int Statist Inst*, 37:351–61.

Vestbo J. 2004. The TORCH (towards a revolution in COPD health) survival study protocol. *Eur Respir J*, 24:206–10.