

RESEARCH

Open Access

# Correlating changes in lung function with patient outcomes in chronic obstructive pulmonary disease: a pooled analysis

Paul W Jones<sup>1\*</sup>, James F Donohue<sup>2</sup>, Jerry Nedelman<sup>3</sup>, Steve Pascoe<sup>4</sup>, Gregory Pinault<sup>4</sup> and Cheryl Lassen<sup>5</sup>

## Abstract

**Background:** Relationships between improvements in lung function and other clinical outcomes in chronic obstructive pulmonary disease (COPD) are not documented extensively. We examined whether changes in trough forced expiratory volume in 1 second (FEV<sub>1</sub>) are correlated with changes in patient-reported outcomes.

**Methods:** Pooled data from three indacaterol studies (n = 3313) were analysed. Means and responder rates for outcomes including change from baseline in Transition Dyspnoea Index (TDI), St. George's Respiratory Questionnaire (SGRQ) scores (at 12, 26 and 52 weeks), and COPD exacerbation frequency (rate/year) were tabulated across categories of  $\Delta$ FEV<sub>1</sub>. Also, generalised linear modelling was performed adjusting for covariates such as baseline severity and inhaled corticosteroid use.

**Results:** With increasing positive  $\Delta$ FEV<sub>1</sub>, TDI and  $\Delta$ SGRQ improved at all timepoints, exacerbation rate over the study duration declined (P < 0.001). Individual-level correlations were 0.03-0.18, but cohort-level correlations were 0.79-0.95. At 26 weeks, a 100 ml increase in FEV<sub>1</sub> was associated with improved TDI (0.46 units),  $\Delta$ SGRQ (1.3-1.9 points) and exacerbation rate (12% decrease). Overall, adjustments for baseline covariates had little impact on the relationship between  $\Delta$ FEV<sub>1</sub> and outcomes.

**Conclusions:** These results suggest that larger improvements in FEV<sub>1</sub> are likely to be associated with larger patient-reported benefits across a range of clinical outcomes.

**Trial Registration:** ClinicalTrials.gov NCT00393458, NCT00463567, and NCT00624286

**Keywords:** COPD, spirometry, FEV<sub>1</sub>, health status, dyspnoea

## Introduction

In the absence of other widely accepted and validated markers for chronic obstructive pulmonary disease (COPD), lung function measurement, specifically forced expiratory volume in 1 second (FEV<sub>1</sub>), has been used as a global marker for pathophysiological changes [1] and by regulators in the drug approval process. Consequently, clinical trials for new products in COPD are typically powered to demonstrate significant improvements in FEV<sub>1</sub>. However, healthcare professionals are more likely to be interested in improvements in patient-reported outcomes such as symptoms and health status, which may better reflect treatment impact on the

patient. Decision-makers also require evidence to assess trends across large cohorts of patients.

Several studies have demonstrated a significant relationship between poor lung function and worsened health and economic outcomes in patients with COPD [2-13], but few have investigated whether changes in lung function associated with an intervention are correlated with changes in such endpoints [12-15]. There is good evidence that declining lung function leads to worsened patient outcomes, but a surprising lack of evidence that improvements in lung function are correlated with improvements in symptomatic outcomes.

Indacaterol is a novel, inhaled, ultra-long-acting  $\beta_2$ -agonist. Initial Phase III trials included over 3000 patients, providing a large pooled dataset. We analysed this dataset in order to examine the relationships

\* Correspondence: [pjones@sgul.ac.uk](mailto:pjones@sgul.ac.uk)

<sup>1</sup>Division of Clinical Science, St George's, University of London, London, UK  
Full list of author information is available at the end of the article

between change in FEV<sub>1</sub> and outcomes including dyspnoea, health status, exacerbations and rescue medication use.

## Methods

### Study design and treatments

This investigation was a pooled analysis of patient-level data from three Phase III, randomised studies: Study 1 (INVOLVE [INdacaterol: Value in COPD: Longer term Validation of Efficacy and safety]) was a double-blind comparison of indacaterol 300 µg or 600 µg once daily with formoterol 12 µg twice daily and placebo for 52 weeks; Study 2 (INHANCE [INdacaterol versus tiotropium to Help Achieve New COPD treatment Excellence]) compared double-blind indacaterol 150 µg or 300 µg once daily with placebo and open-label tiotropium 18 µg once daily for 26 weeks; Study 3 (INLIGHT 1 [INdacaterol: efficacy evaluation using 150 µg doses with COPD Patients]) was a 12-week study comparing double-blind indacaterol 150 µg once daily with placebo for 12 weeks. Patients were permitted to continue inhaled corticosteroid (ICS) monotherapy if the dose and regimen were stable for 1 month before screening, and were to remain stable throughout the study; patients were also permitted rescue salbutamol as needed. Full details have been reported elsewhere [16-18]. All studies were conducted in accordance with the Declaration of Helsinki (1989) and local applicable laws and regulations. Approval was obtained from the Institutional Review Board or Independent Ethics Committee of each participating study centre. All patients provided written informed consent prior to participating in each study included in the pooled analysis. All patient data was anonymised.

### Patients

Patients were male or female, aged ≥ 40 years, with a smoking history of ≥ 20 pack years and a diagnosis of moderate-to-severe COPD [19]. All patients in whom trough FEV<sub>1</sub> measurements were available both at baseline and at 12 weeks were included. Patients with extreme changes from baseline in trough FEV<sub>1</sub> (> +500 or < -500 ml) were excluded, as these values were considered erroneous.

### Endpoints

The primary endpoint in all three studies was trough FEV<sub>1</sub> (average of the 23 h 10 min and 23 h 45 min post-dose value) after 12 weeks of treatment. Trough FEV<sub>1</sub> at baseline was defined as the average of the FEV<sub>1</sub> values 50 and 15 min prior to the first dose of study drug. Trough FEV<sub>1</sub> was assessed at the end of Weeks 4, 8 and 12 in all studies, Weeks 16, 20, 24, 28, 36, 44 and 52 in Study 1, and Weeks 16, 21 and 26 in Study 2. In

all studies, spirometry equipment and performance of spirometric testing was required to be in accordance with ATS/ERS standards [20].

Secondary endpoints included health status (using the St George's Respiratory Questionnaire [SGRQ] [21]), and dyspnoea (using the Transition Dyspnoea Index [TDI] [22]; Studies 1 and 2 only). The SGRQ provides scores between 0 and 100, with higher values indicating greater impairment. The TDI is inherently a change from baseline and provides values between -9 and +9, with positive values indicating improvement. Rescue medication use (number of puffs of salbutamol) was recorded by patients in diaries. COPD exacerbations were defined as the onset or worsening of > 1 respiratory symptom for > 3 consecutive days, requiring intensified treatment (e.g. systemic steroids, antibiotics, oxygen) and/or hospitalisation or emergency room visit. Severe exacerbations were those requiring hospitalisation.

### Statistical methods

The primary objective was to examine relationships between patient-reported outcomes and change from baseline in trough FEV<sub>1</sub> ( $\Delta$ FEV<sub>1</sub>) using data summarisation and model-based analysis. Outcome variables for both analysis approaches were TDI, change from baseline in SGRQ ( $\Delta$ SGRQ), rescue medication use and exacerbation rates.

For TDI and  $\Delta$ SGRQ, relationships were examined with the average of each patient's  $\Delta$ FEV<sub>1</sub> through the corresponding week of observation. For rescue medication use and exacerbations, the average  $\Delta$ FEV<sub>1</sub> over time on treatment was used.

### Data summaries and related inferences

TDI and  $\Delta$ SGRQ were handled as outcome variables at 12, 24/26 and 52 weeks. Responders were patients who achieved at least the minimal clinically important difference (MCID) from baseline (one and four units for TDI and SGRQ, respectively [21,22]). Daily rescue medication use was the number of puffs during treatment divided by the number of days on treatment. Rate of exacerbations was the number of exacerbations on treatment, normalised to 1 year (365 × number of exacerbations while on treatment/days on treatment).

For each of the timepoints, outcomes and responder rates for  $\Delta$ SGRQ and TDI were tabulated across five categories of  $\Delta$ FEV<sub>1</sub> that were chosen to distribute patients approximately equally across categories, and bounded above and below by ± 500 ml. The hypothesis of equality across categories was tested by the Kruskal-Wallis test. Correlation coefficients were computed between observed individual values of  $\Delta$ FEV<sub>1</sub> and the outcome, and between the category midpoint values of  $\Delta$ FEV<sub>1</sub> (-275 ml, 0 ml, 100 ml, 200 ml and 375 ml) and the category mean response of the outcome.

### Model-based analyses

In line with established statistical procedures, generalised linear modelling [23,24] was performed to examine the relationship between  $\Delta FEV_1$  and each outcome variable. For TDI and  $\Delta SGRQ$ , observations at all timepoints were modelled together using repeated-measures multiple regression analyses, assuming constant variance and an unstructured correlation matrix. Time was included both as a main effect and in an interaction with  $\Delta FEV_1$ .

Rescue medication use and exacerbations were modelled as number of puffs and number of exacerbations, respectively, during time on treatment. Rescue medication use was modelled using the zero-inflated negative binomial distribution for likelihood-based model building, and then the final model was refitted using quasi-likelihood to report parameter estimates. Exacerbations were modelled using the negative binomial distribution. For both, in order to ensure positivity of the modelled mean response, the logarithm of the mean was represented as linear in the covariates, and then the mean was found by taking antilogs.

Other predictor variables were baseline trough  $FEV_1$  (continuous), age (continuous), gender (binary), ICS use (binary: yes or no), treatment (indacaterol, formoterol, tiotropium or placebo), screening  $FEV_1$  measured to assess reversibility before and after a short-acting  $\beta_2$ -agonist, and before and after a short-acting anticholinergic, world region (Western Europe and the USA, Eastern Europe and Turkey, Rest-of-World), and time at risk for exacerbations and rescue medication use. Disease severity was included as a binary variable, based on the Global initiative for chronic Obstructive Lung Disease (GOLD) stages [19]; predominantly GOLD 2 (moderate or less severity including 91% moderate, referred to subsequently as GOLD 2) versus predominantly GOLD 3 (severe or greater severity including 98% severe, referred to as GOLD 3), as measured by per cent predicted  $FEV_1$  at screening after short-acting  $\beta_2$ -agonist. The default condition for all models was: baseline  $FEV_1$  of 1.3 l, age 65 years, gender male, GOLD 2, no ICS use, indacaterol treatment, screening  $FEV_1$  before (and after) reversibility testing of 1.3 l (1.5 l) and Western Europe/USA region. All statistical comparisons were made relative to this combination of covariates, and unless otherwise stated, these were the values of the parameters used for predictions by the models.

Model-based inference steps were performed to test for interactions between  $\Delta FEV_1$  and the covariates treatment, disease severity, ICS use and world region. For this purpose, disease severity was represented jointly by baseline  $FEV_1$ , the binary severity indicator defined above, Baseline Dyspnoea Index (BDI) for TDI and baseline SGRQ for  $\Delta SGRQ$ . To allow for the possibility of

differing relationships for negative versus positive values of  $\Delta FEV_1$ , a possible breakpoint at  $\Delta FEV_1 = 0$  was tested in each model. The main effects of covariates were tested for significance according to Wald P values in the final model, with  $P < 0.01$  judged significant without any adjustments for multiplicity.

For each outcome variable, the improvement in expected response for an increase in  $FEV_1$  from 0 to 100 ml was also computed, based on a model that excluded treatment effects, to allow for variation in  $\Delta FEV_1$  between, as well as within, treatments.

### Results

In total, 3313 patients were included in the analysis. Patient demographic and clinical characteristics are presented in Table 1. Age, pre- and post-bronchodilator  $FEV_1$  and body mass index were well balanced across studies. Study 1 included more males, patients taking ICS and patients with slightly lower per cent predicted  $FEV_1$  and reversibility.

#### Data summaries and related inferences

The distribution of average  $\Delta FEV_1$  responses by timepoint is shown in Table 2, both as frequencies within  $\Delta FEV_1$  categories and as percentiles of distributions. Median values ranged from 75 to 94 ml. Approximately 5% of observations were excluded due to extreme  $\Delta FEV_1$  ( $\pm 500$  ml) at any timepoint; 0.7% observations (24/3313) were less than -500 ml and 4.1% (137/3313) were greater than 500 ml at Week 12.

All relationships between  $\Delta FEV_1$  and outcomes were statistically significant, except for severe exacerbations (Table 3). Individual-level correlations were weak (0.03-0.18), reflecting the large variability in outcomes; however, cohort-level correlations were stronger (0.79-0.95). When outcome means were plotted versus  $\Delta FEV_1$  midpoints, there were clear trends towards greater improvement in outcomes with increasing  $\Delta FEV_1$ , particularly for positive  $\Delta FEV_1$  (Figure 1). Responder rates, in terms of TDI and  $\Delta SGRQ$ , followed a similar pattern to the mean outcomes (Table 4).

#### Model-based results

The plots of curves fitted from the model-based analysis for each outcome variable versus  $\Delta FEV_1$  are presented in Figures 2a-d. For TDI and  $\Delta SGRQ$ , the significant breakpoints at zero are evident in the changes of slope in the fitted lines. For rescue medication and exacerbations the breakpoints were not significant. The fitted curves for rescue medication and exacerbations are linear on logarithmic scales, so appear nonlinear on the scales of these plots.

$\Delta FEV_1$  was significantly correlated with TDI score ( $P < 0.0001$ ). A significant breakpoint in the fitted lines is

**Table 1 Baseline characteristics of study participants included in the analysis**

	Study 1 INVOLVE [16]	Study 2 INHANCE [17]	Study 3 INLIGHT 1 [18]	Total
n	1377	1575	361	3313
Age, years	64 (8)	64 (9)	63 (10)	64 (9)
Male/female, %	78/22	63/37	52/48	69/31
Body mass index, kg/m <sup>2</sup>	27 (5)	27 (6)	28 (7)	27 (6)
FEV <sub>1</sub> , % predicted*	53 (14)	56 (14)	55 (14)	55 (14)
FEV <sub>1</sub> /FVC, %*	51 (10)	53 (10)	53 (10)	52 (10)
Pre-bronchodilator FEV <sub>1</sub> , l	1.35 (0.43)	1.33 (0.49)	1.34 (0.51)	1.34 (0.47)
Post-bronchodilator FEV <sub>1</sub> , l*	1.52 (0.47)	1.50 (0.50)	1.51 (0.52)	1.51 (0.49)
Reversibility, %*	13.2 (13.4)	15.5 (15.9)	16.0 (18.7)	14.6 (15.3)
ICS use yes/no, %	55/45	38/62	32/68	45/55
Smoker/ex-smoker, %	40/60	44/56	52/48	43/57
BDI score	6.6 (2.2)	6.5 (2.3)	NA	6.5 (2.2)
SGRQ total score	44 (18)	45 (18)	49 (19)	45 (18)
Treatments				
Placebo, n	322	311	176	809
Indacaterol 75 µg, n	0	67	0	67
Indacaterol 150 µg, n	0	346	185	531
Indacaterol 300 µg, n	363	357	0	720
Indacaterol 600 µg, n	344	68	0	412
Formoterol 12 µg, n	348	75	0	423
Tiotropium 18 µg, n	0	351	0	351

Data are mean (standard deviation) unless otherwise stated. \*Measured 30 min after salbutamol 400 µg inhalation. Reversibility was calculated as the difference between the pre- and post-bronchodilator values of FEV<sub>1</sub> (in l) as a percentage of the pre-bronchodilator value; BDI, Baseline Dyspnoea Index; FEV<sub>1</sub>, forced expiratory volume in 1 second; FVC, forced vital capacity; ICS, inhaled corticosteroid; INVOLVE, INDacaterol: Value in COPD: Longer term Validation of Efficacy and safety; INHANCE, INDacaterol versus tiotropium to Help Achieve New COPD treatment Excellence; INLIGHT, INDacaterol: efficacy evaluation using 150 µg doses with COPD Patients; NA, not available; SGRQ, St George's Respiratory Questionnaire.

seen at zero; the slope was significantly shallower for negative  $\Delta$ FEV<sub>1</sub> compared with positive  $\Delta$ FEV<sub>1</sub> ( $P = 0.003$  for the difference between slopes). The slope of the relationship (determining the magnitude of change in outcome for a given improvement in  $\Delta$ FEV<sub>1</sub>) was not significantly affected by treatment, baseline severity, ICS use or world region. Hence, the overall model-predicted increase in TDI for a 100 ml increase in  $\Delta$ FEV<sub>1</sub> was the same for all combinations of covariates, and estimated to be 0.46 at Week 24/26. Although the slope of the relationship with  $\Delta$ FEV<sub>1</sub> was the same for all covariates, the intercept, that is, the TDI corresponding to zero change in FEV<sub>1</sub>, was not. For a given  $\Delta$ FEV<sub>1</sub>, patients with lower baseline FEV<sub>1</sub>, lower BDI, using ICS or on

placebo, had significantly lower values of TDI, while those from Eastern Europe/Turkey and Rest-of-World regions had significantly higher values. When covariates representative of patients who were less severe were inputted into the model (i.e., GOLD 2, no ICS and baseline FEV<sub>1</sub> of 1.595 l), the model-predicted TDI for a zero and +100 ml change in FEV<sub>1</sub> was 1.98 and 2.44, respectively. For more severe patients (i.e., GOLD 3, ICS and baseline FEV<sub>1</sub> of 0.95 l), the model-predicted TDI for patients with zero and +100 ml change in FEV<sub>1</sub> was -0.20 and 0.26, respectively.

There was a significant correlation between  $\Delta$ SGRQ and  $\Delta$ FEV<sub>1</sub> ( $P < 0.0001$ ). As with TDI, the slope of the relationship between  $\Delta$ SGRQ and  $\Delta$ FEV<sub>1</sub> was

**Table 2 Summary information for averages of  $\Delta$ FEV<sub>1</sub>**

Average of $\Delta$ FEV <sub>1</sub>	n	Number of observations in intervals defined by ml ranges							Percentiles* of observations						
		< -500	-500, -50	-50, 50	50, 150	150, 250	250, 500	> 500	Minimum	5%	25%	50%	75%	95%	Maximum
Week 4-12	3313	24	623	695	717	563	554	137	-1180	-198	-20	94	220	466	1966
Week 4-24/26	2389	16	478	476	550	388	377	104	-1148	-203	-30	88	214	474	1782
Week 4-52	1169	6	292	218	273	165	168	47	-755	-223	-52	75	201	464	1607
On treatment	3313	24	708	662	751	536	501	131	-1180	-213	-32	82	208	467	1966

\*For example, when averaging over time on treatment, the minimum average  $\Delta$ FEV<sub>1</sub> was -1180 ml and the maximum was 1966 ml; between these points, 5% observations were less than or equal to -213 ml and half were less than or equal to 82 ml;  $\Delta$ FEV<sub>1</sub>, change from baseline in trough forced expiratory volume in 1 second.

**Table 3 Outcome means by average  $\Delta FEV_1$  category, P values for associations between average  $\Delta FEV_1$  and outcome, and correlations at individual and cohort levels**

Average $\Delta FEV_1$ (ml)	Category midpoint value of $\Delta FEV_1$ (ml)	Withdrawal rate* (% patients)	TDI at 12 weeks (n = 2781)	TDI at 24/26 weeks (n = 2208)	TDI at 52 weeks (n = 1099)	$\Delta SGRQ$ at 12 weeks (n = 3141)	$\Delta SGRQ$ at 24/26 weeks (n = 2215)	$\Delta SGRQ$ at 52 weeks (n = 1115)	Rescue medication mean puffs per day (over study duration) (n = 3158)	Exacerbation rate (per year) (n = 3158)	Severe exacerbation rate (per year) (n = 3158)
-500, -50	-275	11.2	1.44	1.57	1.24	-3.15	-4.70	-2.21	2.46	0.63	0.059
-50, 50	0	9.0	1.31	1.39	1.92	-3.17	-3.81	-3.03	2.57	0.58	0.065
50, 150	100	10.1	1.79	1.97	1.65	-3.84	-4.74	-4.22	2.10	0.61	0.057
150, 250	200	10.2	2.12	2.23	2.23	-5.84	-6.34	-6.70	1.80	0.51	0.048
250, 500	375	6.7	2.68	3.03	3.27	-7.38	-7.29	-9.06	1.66	0.38	0.021
P value (Kruskal-Wallis)			< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	0.1
Correlation, individual level			0.15	0.14	0.18	-0.12	-0.07	-0.16	-0.11	-0.06	-0.03
Correlation, cohort level			0.90	0.88	0.92	-0.90	-0.79	-0.95	-0.88	-0.89	-0.81

\*Non-completers in all studies categorised by  $\Delta FEV_1$  at Day 2;  $\Delta FEV_1$ , change from baseline in trough forced expiratory volume in 1 second; TDI, Transition Dyspnoea Index;  $\Delta SGRQ$ , change from baseline in St George's Respiratory Questionnaire

significantly shallower for negative  $\Delta FEV_1$  ( $P = 0.002$  for the difference between slopes). The slope of the relationship with improvement in  $FEV_1$  was not significantly affected by treatment, ICS use, or world region, but it was steeper for patients in GOLD 3, and with baseline  $FEV_1$  0.95 l compared with GOLD 2 and baseline  $FEV_1$  1.595 l ( $P = 0.004$ ). For an increase of  $\Delta FEV_1$  of 100 ml, the model predicted a change in SGRQ of -1.3 for GOLD 2 and -1.9 for GOLD 3 patients at Week 24/26. Patients with worse baseline  $FEV_1$ , with worse baseline SGRQ, using ICS or on placebo, had significantly higher  $\Delta SGRQ$ , whereas patients from Eastern Europe/Turkey and Rest-of-World regions had significantly lower  $\Delta SGRQ$  at Week 24/26. For GOLD 2 patients, who had used no ICS and had baseline SGRQ of < 31, the model-predicted improvement in SGRQ at Week 26 for a zero and +100 ml change in  $FEV_1$  was -1.6 and -2.9, respectively. Similarly, for GOLD 3 patients who had used ICS and had baseline SGRQ of > 58, the model-predicted improvement in SGRQ at Week 24/26 was -0.9 and -2.8, respectively.

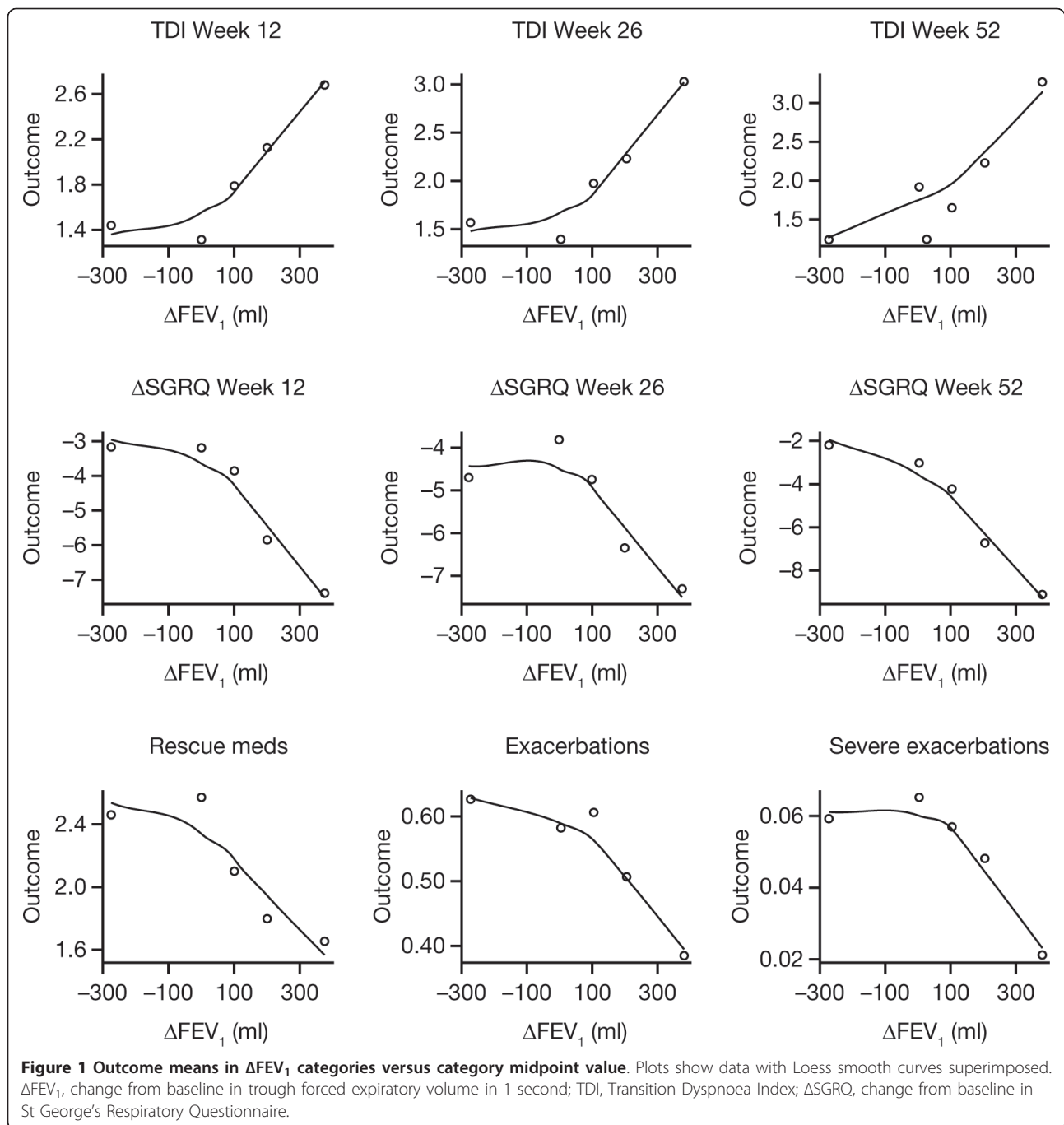
$\Delta FEV_1$  was significantly correlated with rescue medication use ( $P < 0.0001$ ). Treatment, baseline severity, ICS use or world region, did not significantly affect the slope of the relationship, and the slope did not change significantly between negative and positive  $\Delta FEV_1$ . Hence, for all combinations of covariates, an increase of 100 ml in  $\Delta FEV_1$  is predicted to yield the same 10% reduction in rescue medication use. Patients with lower baseline  $FEV_1$ , male patients, those with higher baseline medication usage or more severe disease, using ICS or on placebo or tiotropium, had significantly higher rates

of rescue medication usage. Younger patients (< 65 years) had almost significantly higher rates ( $P = 0.012$ , versus the defined significance level of  $P < 0.01$ ). For GOLD 2 patients not receiving ICS, the predicted daily number of puffs of rescue medication for a zero and +100 ml  $\Delta FEV_1$  was 0.89 and 0.80, respectively, and 1.83 and 1.64 for those in GOLD 3 and using ICS.

$\Delta FEV_1$  was significantly correlated with exacerbations ( $P = 0.002$ ). Treatment, baseline severity, ICS use or world region, did not significantly affect the slope of the relationship. Furthermore, the slope did not change significantly between negative and positive  $\Delta FEV_1$ . Hence, for all combinations of covariates, an increase of 100 ml in  $\Delta FEV_1$  is predicted to yield the same 12% decrease in exacerbations. Patients with lower baseline  $FEV_1$  and patients using ICS had significantly higher rates of exacerbations. Patients from the Eastern Europe/Turkey region had significantly lower rates of exacerbations. The model estimate for the annual rate of exacerbations for patients with a zero and +100 ml  $\Delta FEV_1$  were 0.29 and 0.25, respectively, for GOLD 2 patients not using ICS; and 1.28 and 1.12, respectively, for patients in GOLD 3 and using ICS. As in the summary analysis, the rate of severe exacerbations requiring hospitalisation was not significantly correlated with  $\Delta FEV_1$  ( $P = 0.3$ ).

## Discussion

Our analyses show that improvement in  $FEV_1$  is significantly related to changes in the patient-reported outcomes TDI, SGRQ, exacerbation rate and rescue medication use over 12-52 weeks of treatment. These relationships were significant at both an individual and



population level, although correlations were much stronger in the population-based analyses.

Few studies have examined the relationship between change in  $FEV_1$  and change in outcomes. However, our results are consistent with analyses of patients from the 3-year EUROSCOP (The European Respiratory Society Study on Chronic Obstructive Pulmonary disease) study, in which an improvement of 100 ml in  $FEV_1$  was associated with a 4% reduction in dyspnoea in males [13],

and a 16-week clinical study, in which a significant, but weak correlation between change in  $FEV_1$  and change in SGRQ score was demonstrated ( $r = 0.33$ ,  $P = 0.001$ ) [14]. Further, a recent systematic review of 22 studies found that 100 ml increase in  $FEV_1$  was associated with a statistically significant reduction in SGRQ of 2.5 [15]. However, to our knowledge, the current analysis is the largest and most comprehensive to investigate the correlation between change in  $FEV_1$  and change in outcomes

**Table 4 Responder rates\* for TDI<sup>†</sup> and ΔSGRQ<sup>†</sup> by average ΔFEV<sub>1</sub> category**

Average ΔFEV <sub>1</sub> (ml)	Category midpoint value of ΔFEV <sub>1</sub> (ml)	TDI at 12 weeks % responders (n = 2781)	TDI at 24/26 weeks % responders (n = 2208)	TDI at 52 weeks % responders (n = 1099)	ΔSGRQ at 12 weeks % responders (n = 3141)	ΔSGRQ at 24/26 weeks % responders (n = 2215)	ΔSGRQ at 52 weeks % responders (n = 1115)
-500, -50	-275	50	51	45	42	49	41
-50, 50	0	48	49	53	46	45	45
50, 150	100	54	57	50	48	48	49
150, 250	200	59	60	58	53	59	54
250, 500	375	66	69	69	56	57	65

\*Responders were patients who achieved at least the minimal clinically important difference (MCID; one and four units for TDI and ΔSGRQ, respectively); TDI, Transition Dyspnoea Index; ΔSGRQ, change from baseline in St George's Respiratory Questionnaire; ΔFEV<sub>1</sub>, change from baseline in trough forced expiratory volume in 1 second.

<sup>†</sup>For the outcomes, each row contains approximately 20% of the column's indicated n, as in the left half of Table 2

using individual patient data from studies of similar design. This provides a relatively homogeneous population for analysis, compared with study-level meta-analyses.

We demonstrated that a 100 ml increase in trough FEV<sub>1</sub> (a magnitude of change with perceptible effects [25]) was associated with a 0.46-unit increase in TDI and a 1.3- to 1.9-unit improvement in SGRQ after a 24/26-week treatment period and, over 12-52 weeks of treatment, a 10% decrease in daily rescue medication use and a 12% decrease in the annual exacerbation rate. In general, we found that treatment, baseline severity, concomitant ICS use and world region, did not affect the slope of the relationship between outcome and change in FEV<sub>1</sub>, except for ΔSGRQ where more severe COPD, as characterised by a lower FEV<sub>1</sub> and a higher SGRQ at baseline, was associated with a steeper slope, compared with less severe COPD. This is consistent with results from the 3-year TORCH (TOwards a Revolution in COPD Health) study, in which trends to greater improvement in SGRQ with worsening GOLD severity were noted with active treatments [26].

Although severe exacerbations showed a trend toward greater reductions with increasing ΔFEV<sub>1</sub>, the relationship was not statistically significant. While the observed 12% reduction in overall exacerbation rate for an improvement of 100 ml in FEV<sub>1</sub> was comparable with previously published data [11], the studies included in our analysis were not powered to show an effect on exacerbations, and did not specifically recruit patients at risk of exacerbations.

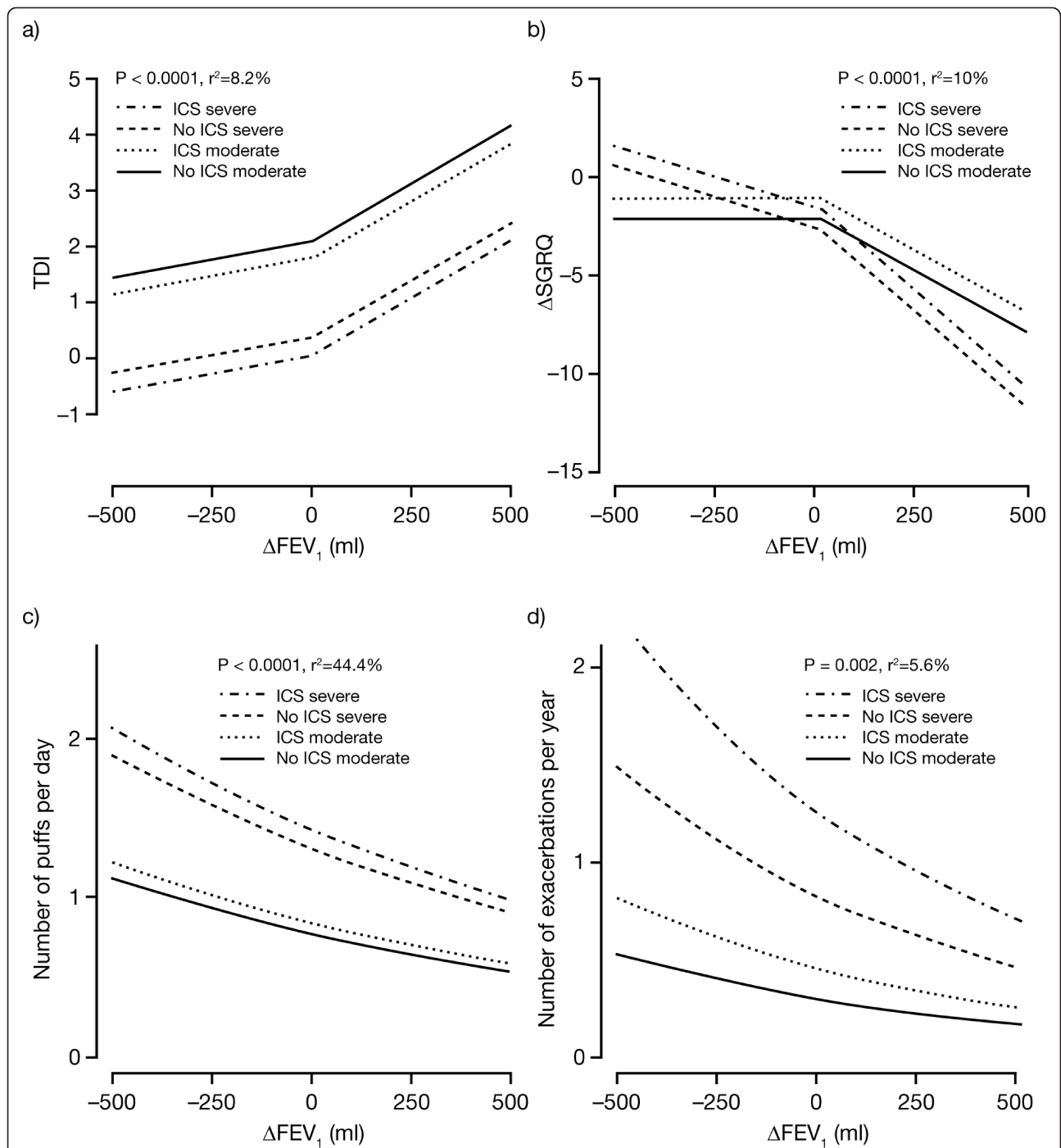
We found inconsistent effects of different treatments across individual outcomes, perhaps due to patient numbers in sub-categories being too low to demonstrate consistent differences for individual treatments across all outcomes. However, our analysis did demonstrate that the relationship between ΔFEV<sub>1</sub> and outcome appeared to be the same, regardless of treatment arm. Similarly, baseline severity, ICS use and world region were assessed as main effects, as well as for their

potential influence on the effect of ΔFEV<sub>1</sub>. Although numbers of patients in GOLD 4 (as well as GOLD 1) were too small to make any inferences, patients predominantly in GOLD 3 at baseline, and those using ICS, consistently exhibited significantly worse outcomes. Indeed, the variability in baseline severity and ICS use are likely to have been major contributors to the large variability in observed outcomes.

The relationships between outcomes and ΔFEV<sub>1</sub> may differ between negative and positive ΔFEV<sub>1</sub>, and for this reason, the models included a possible breakpoint at zero in the relationship slope. The inclusion of this breakpoint was found to be significant for TDI and ΔSGRQ, suggesting that baseline severity and other included covariates could not explain the observed behaviour fully. These results may have been influenced by differences in withdrawal rates between categories [27], since the highest withdrawal rate was in those with a negative change in FEV<sub>1</sub>, although differences between groups were minimal. The inclusion of a breakpoint was not significant for rescue medication and exacerbations, even though Figure 1 may have anticipated its importance, especially for exacerbations. The large variability and count nature of the data for rescue medication and exacerbations may have caused 'Type-2' statistical errors, i.e., failure to find the true breakpoints to be significant.

We found that zero change in FEV<sub>1</sub> was associated with significant positive improvements in TDI and SGRQ. Additionally, while a greater proportion of patients achieved the MCID for TDI and SGRQ as ΔFEV<sub>1</sub> increased, our results indicated that as many as 50% patients responded, irrespective of ΔFEV<sub>1</sub>, possibly an effect of clinical trial participation seen consistently in the placebo limb of clinical trials [28-30].

We constructed the models in our analysis using ΔFEV<sub>1</sub> as a predictor, and the other outcome measures as the response variables, based on the results of a carefully-controlled series of clinical trials. However, ΔFEV<sub>1</sub> was as much a response as was the outcome, so ΔFEV<sub>1</sub> was not an 'independent' variable controlled as part of



**Figure 2 Outcomes versus  $\Delta FEV_1$ : curves fitted from model-based analysis.** SGRQ and TDI data is for Week 24/26. Exacerbations are reported counts normalised to 1 year. Rescue medication use is reported numbers of puffs normalised to 1 day. For plotting predicted curves, 'moderate' refers to baseline  $FEV_1$  1.595 l (third quartile of observed values) and GOLD 2 (moderate or less [19]); 'severe' refers to baseline  $FEV_1$  0.95 l (first quartile of observed values) and GOLD 3 (severe or greater [19]). For TDI, 'moderate' and 'severe' refer to BDI at quartiles 2 and 1, respectively. For  $\Delta SGRQ$ , 'moderate' and 'severe' refer to baseline SGRQ at quartiles 31.3 and 58.13, respectively;  $\Delta FEV_1$ , change from baseline in trough forced expiratory volume in 1 second;  $\Delta SGRQ$ , change from baseline in St George's Respiratory Questionnaire; TDI, Transition Dyspnoea Index; GOLD, Global initiative for chronic Obstructive Lung Disease; BDI, Baseline Dyspnoea Index; ICS, inhaled corticosteroid.



the experimental design. There may have been further confounders that simultaneously affected how both  $\Delta FEV_1$  and the outcome responded to treatment. The fitted models therefore describe the observed relationships under the conditions of a clinical trial, but do not provide a definitive answer as to whether there is a causal relationship between  $\Delta FEV_1$  and the outcomes.

The studies included in our analysis were powered on the spirometric endpoint  $FEV_1$ , which is required by regulatory agencies for the approval of bronchodilators, and is included in the majority of treatment guidelines. For this reason we made  $FEV_1$  the focus of our analysis. Other physiological parameters such as inspiratory capacity may have stronger correlations with dyspnoea [31]. However data for these parameters were not available from our dataset and further research is needed to investigate such correlations in large numbers of patients.

## Conclusions

It is commonly stated that spirometry does not fully capture the impact of COPD on a patient's health [32]. Our analysis of a large cohort of patients has demonstrated that in individual subjects, change in  $FEV_1$  is a significant, albeit relatively weak predictor of improvement in patient-reported outcomes. However, the current analysis also shows that, at a population level, improvements in  $FEV_1$  with long-acting bronchodilator therapy are strongly correlated with improvements in dyspnoea, health status and exacerbations. This suggests that interventions which significantly improve  $FEV_1$  are also likely to be associated with improved clinical and patient-reported outcomes.

## List of abbreviations

BDI: Baseline Dyspnoea Index; COPD: Chronic Obstructive Pulmonary Disease; EUROSCOP: The European Respiratory Society Study on Chronic Obstructive Pulmonary disease;  $FEV_1$ : Forced Expiratory Volume in 1 second; FVC: Forced Vital Capacity; GOLD: Global initiative for chronic Obstructive Lung Disease; ICS: Inhaled Corticosteroid; INHANCE: INdacaterol versus tiotropium to Help Achieve New COPD treatment Excellence; INLIGHT: INdacaterol: efficacy evaluation usInG 150  $\mu$ g doses with COPD Patients; INVOLVE: INdacaterol: Value in COPD: Longer term Validation of Efficacy and safety; MCID: Minimal Clinically Important Difference; SGRQ: St. George's Respiratory Questionnaire; TDI: Transition Dyspnoea Index; TORCH: TOwards a Revolution in COPD Health.

## Acknowledgements

M. Sayers (CircleScience, Tytherington, UK), a professional medical writer funded by Novartis assisted in the preparation of the manuscript. Role of funding source This analysis was sponsored by Novartis Pharma AG (Basel, Switzerland), who were involved in the study design, the collection, analysis and interpretation of data, writing of the study report, and the decision to submit the manuscript for publication.

## Author details

<sup>1</sup>Division of Clinical Science, St George's, University of London, London, UK. <sup>2</sup>Division of Pulmonary & Critical Care Medicine, University of North Carolina,

School of Medicine, Chapel Hill, NC, USA. <sup>3</sup>Novartis Pharmaceuticals Corporation, East Hanover, USA. <sup>4</sup>Novartis Pharma AG, Basel, Switzerland. <sup>5</sup>Novartis Horsham Research Centre, Horsham, West Sussex, UK.

## Authors' contributions

PWJ participated in the design and analysis planning and advised on the interpretation of the study. JFD participated in the design and analysis planning and advised on the interpretation of the study. JN developed the design, concept of the study and analysis, and carried out the statistical analysis. SP conceived of the study, participated in its design and analysis planning and contributed to its interpretation. GP programmed the analysis data set. CL conceived of the study, participated in its design and analysis planning and contributed to its interpretation. All authors had full access to the data and were involved in drafting the manuscript. All authors read and approved the final manuscript.

## Competing interests

PWJ has received consultancy fees and honoraria from AstraZeneca, Almirall, Bayer, Boehringer Ingelheim, Chiesi, GlaxoSmithKline, Novartis, Pfizer, Roche and Spiration; and grants for his institution from GlaxoSmithKline and Novartis.

JFD has received consultancy fees and honoraria from Novartis, Almirall, Forest laboratories, Boehringer Ingelheim, GlaxoSmithKline, Pfizer, Sunovion Dey and Talecris.

JN, GP and CL are employees of Novartis. SP is an ex-employee of Novartis.

Received: 6 October 2011 Accepted: 29 December 2011

Published: 29 December 2011

## References

1. Jones PW, Agusti AGN: **Outcomes and markers in the assessment of chronic obstructive pulmonary disease.** *Eur Respir J* 2006, **27**:822-832.
2. Álvarez-Gutiérrez FJ, Miravittles M, Calle M, Gobartt E, López F, Martín A, Grupo de Estudio EIME: **Impact of chronic obstructive pulmonary disease on activities of daily living: results of the EIME multicenter study.** *Arch Bronconeumol* 2007, **43**:64-72.
3. Anzueto A, Leimer I, Kesten S: **Impact of frequency of COPD exacerbations on pulmonary function, health status and clinical outcomes.** *Int J Chron Obstruct Pulmon Dis* 2009, **4**:245-251.
4. de Torres JP, Cote CG, Lopez MV, Casanova C, Díaz O, Marin JM, Pinto-Plata V, de Oca MM, Nekach H, Dordelly LJ, Aguirre-Jaime A, Celli BR: **Sex differences in mortality in patients with COPD.** *Eur Respir J* 2009, **33**:528-535.
5. Groenewegen KH, Postma DS, Hop WC, Wielders PL, Schlösser NJ, Wouters EF, COSMIC Study Group: **Increased systemic inflammation is a risk factor for COPD exacerbations.** *Chest* 2008, **133**:350-357.
6. Gudmundsson G, Gislason T, Janson C, Lindberg E, Hallin R, Ulrik CS, Brøndum E, Nieminen MM, Aine T, Bakke P: **Risk factors for rehospitalisation in COPD: role of health status, anxiety and depression.** *Eur Respir J* 2005, **26**:414-419.
7. Gudmundsson G, Gislason T, Lindberg E, Hallin R, Ulrik CS, Brøndum E, Nieminen MM, Aine T, Bakke P, Janson C: **Mortality in COPD patients discharged from hospital: the role of treatment and co-morbidity.** *Respir Res* 2006, **7**:109.
8. Mannino DM, Davis KJ: **Lung function decline and outcomes in an elderly population.** *Thorax* 2006, **61**:472-477.
9. McGlone S, Venn A, Walters EH, Wood-Baker R: **Physical activity, spirometry and quality-of-life in chronic obstructive pulmonary disease.** *COPD* 2006, **3**:83-88.
10. Niewoehner DE, Lokhygina Y, Rice K, Kuschner WG, Sharafkhaneh A, Sarosi GA, Krumpe P, Pieper K, Kesten S: **Risk indexes for exacerbations and hospitalizations due to COPD.** *Chest* 2007, **131**:20-28.
11. Niewoehner DE: **Relation of chronic obstructive pulmonary disease exacerbations to  $FEV_1$  - an intricate tango.** *Respiration* 2009, **77**:229-235.
12. Omata M, Wakabayashi R, Kudoh S, Kida K: **Correlation between bronchodilator responsiveness and quality of life in chronic obstructive pulmonary disease.** *Allergol Int* 2007, **56**:15-22.
13. Watson L, Schouten JP, Lofdahl CG, Pride NB, Laitinen LA, Postma DS, European Respiratory Society Study on Chronic Obstructive Pulmonary Disease: **Predictors of COPD symptoms: does the sex of the patient matter?** *Eur Respir J* 2006, **28**:311-318.

14. Jones PW: **Health status measurement in chronic obstructive pulmonary disease.** *Thorax* 2001, **56**:880-887.
15. Westwood M, Bourbeau J, Jones PW, Cerulli A, Capkun-Niggli G, Worthy G: **Relationship between FEV<sub>1</sub> change and patient-reported outcomes in randomised trials of inhaled bronchodilators for stable COPD: a systematic review.** *Respir Res* 2011, **12**:40.
16. Dahl R, Chung KF, Buhl R, Magnussen H, Nonikov V, Jack D, Bleasdale P, Owen R, Higgins M, Kramer B, on behalf of the INVOLVE (INdacaterol: Value in COPD, Longer Term Validation of Efficacy and Safety) Study Investigators: **Efficacy of a new once-daily long-acting inhaled  $\beta_2$ -agonist indacaterol versus twice-daily formoterol in COPD.** *Thorax* 2010, **65**:473-479.
17. Donohue JF, Fogarty C, Lötvalld J, Mahler DA, Worth H, Yorgancıoğlu A, Iqbal A, Swales J, Owen R, Higgins M, Kramer B, for the INHANCE Study Investigators: **Once-daily bronchodilators for chronic obstructive pulmonary disease: indacaterol versus tiotropium.** *Am J Respir Crit Care Med* 2010, **182**:155-162.
18. Feldman G, Siler T, Prasad N, Jack D, Piggott S, Owen R, Higgins M, Kramer B, the INLIGHT 1 study group: **Efficacy and safety of indacaterol 150  $\mu$ g once-daily in COPD: a double-blind, randomised, 12-week study.** *BMC Pulm Med* 2010, **10**:11.
19. Global initiative for chronic Obstructive Lung Disease (GOLD): **Global Strategy for the Diagnosis, Management, and Prevention of chronic obstructive pulmonary disease.**, Date last updated: 2005 [<http://www.goldcopd.org/Guidelines/guidelines-global-strategy-for-diagnosis-management-2005.html>].
20. Miller MR, Hankinson J, Brusasco V, Burgos F, Casaburi R, Coates A, Crapo R, Enright P, van der Grinten CPM, Gustafsson P, Jensen R, Johnson DC, MacIntyre N, McKay R, Navajas D, Pedersen OF, Pellegrino R, Viegi G, Wanger J: **Standardisation of spirometry.** *Eur Respir J* 2005, **26**:319-338.
21. Jones PW: **St. George's Respiratory Questionnaire: MCID.** *COPD* 2005, **2**:75-79.
22. Mahler DA, Witek TJ Jr: **The MCID of the transition dyspnea index is a total score of one unit.** *COPD* 2005, **2**:99-103.
23. Harrell FE: *Regression modeling strategies* New York: Springer-Verlag; 2001.
24. Venables WN, Ripley BD: *Modern applied statistics* New York: Springer-Verlag; 2002.
25. Donohue J: **Minimal clinically important differences in COPD lung function.** *COPD* 2005, **2**:111-124.
26. Jenkins CR, Jones PW, Calverley PMA, Celli B, Anderson JA, Ferguson GT, Yates JC, Willits LR, Vestbo J: **Efficacy of salmeterol/fluticasone propionate by GOLD stage of chronic obstructive pulmonary disease: analysis from the randomised, placebo-controlled TORCH study.** *Respir Res* 2009, **10**:59.
27. Vestbo J, Anderson JA, Calverley PMA, Celli B, Ferguson GT, Jenkins C, Yates JC, Jones PW: **Bias due to withdrawal in long-term randomised trials in COPD: evidence from the TORCH study.** *Clin Respir J* 2011, **5**:44-49.
28. Calverley PMA, Pauwels R, Vestbo J, Jones PW, Pride N, Gulsvick A: **Combined salmeterol and fluticasone in the treatment of chronic obstructive pulmonary disease.** *Lancet* 2003, **361**:449-456.
29. Calverley PM, Anderson JA, Celli BR, Ferguson GT, Jenkins C, Jones PW, Yates JC, Vestbo J: **Salmeterol and fluticasone propionate and survival in chronic obstructive pulmonary disease.** *N Engl J Med* 2007, **356**:775-789.
30. Tashkin DP, Celli B, Senn S, Burkhart D, Kesten S, Menjoge S, Decramer M, UPLIFT Study Investigators: **A 4-year trial of tiotropium in chronic obstructive pulmonary disease.** *N Engl J Med* 2008, **359**:1543-1554.
31. Di Marco F, Milic-Emili J, Boveri B, Carlucci P, Santus P, Casanova F, Cazzola M, Centanni S: **Effect of inhaled bronchodilators on inspiratory capacity and dyspnoea at rest in COPD.** *Eur Respir J* 2003, **21**:86-94.
32. Global initiative for chronic Obstructive Lung Disease (GOLD): **Global Strategy for the Diagnosis, Management, and Prevention of chronic obstructive pulmonary disease.**, Date last updated: 2010 [<http://www.goldcopd.com>].

doi:10.1186/1465-9921-12-161

**Cite this article as:** Jones *et al.*: Correlating changes in lung function with patient outcomes in chronic obstructive pulmonary disease: a pooled analysis. *Respiratory Research* 2011 **12**:161.

**Submit your next manuscript to BioMed Central and take full advantage of:**

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at  
[www.biomedcentral.com/submit](http://www.biomedcentral.com/submit)

