ORIGINAL ARTICLE
COPD

Treatment response in COPD: does FEV1 say it all? A *post hoc* analysis of the CRYSTAL study

Konstantinos Kostikas[1], Timm Greulich[2], Alexander J. Mackay[3], Nadine S. Lossi[4], Maryam Aalamian-Mattheis[5], Xavier Nunez[6], Veronica A. Pagano[6], Francesco Patalano[5], Andreas Clemens[5,7,8] and Claus F. Vogelmeier[2,8]

Affiliations: [1]Respiratory Medicine Dept, University of Ioannina, Ioannina, Greece. [2]Dept of Medicine, Pulmonary and Critical Care Medicine, University Medical Center Giessen and Marburg, Philipps-Universität Marburg, Member of the German Center for Lung Research (DZL), Giessen, Germany. [3]Airways Disease Section, National Heart and Lung Institute, Imperial College London, London, UK. [4]Novartis Pharma GmbH, Nuremberg, Germany. [5]Novartis Pharma AG, Basel, Switzerland. [6]TFS Develop, Barcelona, Spain. [7]Dept of Cardiology and Angiology I, Heart Center Freiburg University, Faculty of Medicine, University of Freiburg, Freiburg, Germany. [8]These two authors contributed equally to this work.

Correspondence: Konstantinos Kostikas, Respiratory Medicine Dept, University of Ioannina, Leoforos Stavrou Niarchou, Ioannina 45500, Greece. E-mail: ktkostikas@gmail.com

ABSTRACT The association between clinically relevant changes in patient-reported outcomes (PROs) and forced expiratory volume in 1 s (FEV_1) in patients with chronic obstructive pulmonary disease (COPD) has rarely been investigated.

Using CRYSTAL, a 12-week open-label study in symptomatic, nonfrequently exacerbating patients with moderate COPD, we assessed at baseline the correlations between several PROs (Baseline Dyspnoea Index, modified Medical Research Council dyspnoea scale, COPD Assessment Test (CAT) and Clinical COPD Questionnaire (CCQ)), and between FEV_1 and PROs. Associations between clinically relevant responses in FEV_1, CAT, CCQ and Transition Dyspnoea Index (TDI) at week 12 were also assessed.

Using data from 4324 patients, a strong correlation was observed between CAT and CCQ (r_s=0.793) at baseline, with moderate or weak correlations between other PROs, and no correlation between FEV_1 and any PRO. At week 12, 2774 (64.2%) patients were responders regarding TDI, CAT or CCQ, with 583 (13.5%) responding using all three measures. In comparison, 3235 (74.8%) were responders regarding FEV_1, TDI, CAT or CCQ, with 307 (7.1%) responding concerning all four parameters.

Increases in lung function were accompanied by clinically relevant improvements of PROs in a minority of patients. Our results also suggest that PROs are not interchangeable. Thus, the observed treatment success in a clinical trial may depend on the selected parameters.

@ERSpublications
Assessments of both lung function and various patient-reported outcomes in clinical trials may be necessary for a more complete picture of treatment response in patients with COPD and to guide treatment decisions http://ow.ly/msoz30nmupG

Cite this article as: Kostikas K, Greulich T, Mackay AJ, *et al*. Treatment response in COPD: does FEV_1 say it all? A *post hoc* analysis of the CRYSTAL study. *ERJ Open Res* 2019; 5: 00243-2018 [https://doi.org/10.1183/23120541.00243-2018].

Received: Dec 21 2018 | Accepted: Dec 21 2018

Introduction

Chronic obstructive pulmonary disease (COPD) is a complex, heterogeneous disease usually with a decline in lung function and worsening symptoms; hence, both forced expiratory volume in 1 s (FEV_1; lung function) and validated patient-reported outcomes (PROs) are used in clinical trials to assess disease severity and response to treatment [1–4]. The PROs are different in terms of their scope of assessment and in the information that they capture. PRO questionnaires such as the Baseline Dyspnoea Index (BDI), Transition Dyspnoea Index (TDI) and modified Medical Research Council (mMRC) dyspnoea scale are used to assess dyspnoea, whereas the Clinical COPD Questionnaire (CCQ), COPD Assessment Test (CAT) and St George's Respiratory Questionnaire (SGRQ) are commonly used to assess patients' health status [5–11]. Furthermore, the mMRC scale is unidirectional and minimally responsive to treatment interventions, while the BDI, TDI, CAT, CCQ and SGRQ (approved by the USA Food and Drug Administration) are multidirectional [5–12].

With the availability of numerous PROs, it is important to understand which provide a better evaluation of patients' health status and demonstrate responses to treatment. Even when PROs evaluate the same parameter, *e.g.* dyspnoea, they may not always capture a uniform response [13]. Hence, it would be useful to examine if the PROs correlate with each other and whether any specific PROs better reflect treatment benefit (as expressed by minimal clinically important differences (MCIDs)) than the others. Furthermore, understanding the relationship between the PROs and lung function (FEV_1) may provide insights into whether a change in lung function translates to a change perceptible by the patients (assessed through PROs).

The CRYSTAL study assessed a large number of symptomatic, nonfrequently exacerbating COPD patients with moderate airflow limitation, who were directly switched to glycopyrronium or indacaterol/glycopyrronium from prior treatments in a clinical practice setting, with FEV_1 and TDI evaluated as co-primary end-points [4]. In this *post hoc* analysis, we investigated the associations between the different PROs evaluated in the CRYSTAL trial (mMRC scale, BDI, CAT and CCQ), and between FEV_1 and the PROs at baseline. We also assessed the associations between clinically relevant changes in PROs and FEV_1 during the 12-week study period, regardless of the effects of study interventions.

Methods

Study design and patients

CRYSTAL was a 12-week, randomised, open-label study in patients with moderate COPD and a history of one exacerbation or less in the previous year (ClinicalTrials.gov identifier NCT01985334). The study assessed the efficacy and safety of a direct switch from previous treatments to indacaterol/glycopyrronium 110/50 μg or glycopyrronium 50 μg once daily. Patients were categorised based on mMRC grade and prior medication. The detailed study design and patient characteristics have been described in the primary publication of the study [4].

Patients aged ⩾40 years were included if they had a clinical diagnosis of moderate COPD, were current or ex-smokers with a smoking history of ⩾10 pack-years, mMRC grade ⩾1, FEV_1 ⩾50% and <80% predicted, and FEV_1/forced vital capacity ratio <0.70. Patients were excluded if they had a body mass index >40 $kg \cdot m^{-2}$, history of asthma and more than one COPD exacerbation requiring systemic corticosteroids or antibiotics and/or hospitalisation in the previous 12 months. Baseline treatment included short-acting β_2-agonists and/or short-acting muscarinic antagonists, or long-acting β_2-agonists or long-acting muscarinic antagonists, or long-acting β_2-agonists plus inhaled corticosteroids in free- or fixed-dose combinations.

The CRYSTAL study was conducted in accordance with the guidelines for Good Clinical Practice and as per the ethical principles of the Declaration of Helsinki. The protocol was approved by an independent ethics committee or an institutional review board for each centre in each country (Austria, Belgium, Czech Republic, Denmark, Estonia, France, Germany, Greece, Hungary, Ireland, Italy, Latvia, Lithuania, Norway, Poland, Portugal, Romania, Russian Federation, Slovakia, Slovenia, Spain, Sweden and UK). Further details are presented in the primary publication [4].

Assessments

The present analysis was an exploratory objective of the CRYSTAL study. PROs and FEV_1 were measured at baseline and at week 12. BDI scores range from 0 to 12 [5], TDI from −9 to +9 [5] and mMRC grades from 0 to 5 [6]. CAT scores range from 0 to 40 [8] and CCQ (and its domain scores: symptoms, functional state and mental state) from 0 to 6 [7]. Lower BDI and TDI scores indicate higher disease burden, whereas higher mMRC, CAT and CCQ scores indicate higher disease burden. FEV_1 was measured both as absolute volume (mL) and as percentage predicted.

Baseline correlation analysis

Relationships among the following PROs were assessed at baseline: 1) mMRC *versus* BDI, CAT and CCQ; 2) BDI *versus* CAT and CCQ; and 3) CAT *versus* CCQ. Correlations between lung function (FEV_1 % pred) and CAT, BDI and CCQ (total and domain scores) at baseline were also assessed.

Responder analysis at week 12

Patients were defined as responders if they achieved the MCID from baseline of ⩾100 mL increase in FEV_1 [14], ⩾1 unit increase in TDI [5], ⩾0.4 unit decrease in CCQ [15] or ⩾2 unit decrease in CAT [16]. Distribution of responders was assessed based on MCID in PROs only, and in PROs and FEV_1 together. Although the MCIDs are generally calculated in comparison with placebo, in the CRYSTAL study we have predefined the MCIDs in comparison with the active comparators.

Statistical analysis

The intention-to-treat population of the CRYSTAL study, regardless of the study treatment, was used for the analysis. Spearman's rank-order analysis was performed to evaluate the relationship between the four PROs, and between FEV_1 % pred and PROs. The correlation coefficient (r_s) indicates the strength of correlation. Test for significance was based on the assumption of a correlation coefficient equal to 0.

Results

All 4324 patients from the intention-to-treat population of the CRYSTAL study were included in this analysis. Mean age was 64.6 years, with approximately 67% men, 53% current smokers and a moderate airflow limitation (mean post-bronchodilator FEV_1 64.7% predicted). Demographics and baseline characteristics are presented in table 1.

Baseline correlation analysis

At baseline, there was a strong positive correlation between the CAT and CCQ health status questionnaires (r_s=0.793), while there was a moderate negative correlation between the two dyspnoea questionnaires, *i.e.* BDI and mMRC (r_s=−0.466). There were moderate negative correlations between BDI and CAT or CCQ

TABLE 1 Demographic and baseline characteristics (intention-to-treat population)

Patients	4324
Age years	64.6±8.30
Male	2908 (67.3)
Body mass index kg·m^{-2}	27.6±5.00
Current smokers	2292 (53.0)
Duration of COPD years	6.1±5.50
mMRC grade	
1	2591 (59.9)
⩾2	1693 (39.2)
BDI total score	7.3±1.80
CAT total score	13.2±6.50
CCQ total score	1.8±0.90
Post-bronchodilator FEV_1 L	1.85±0.49
Post-bronchodilator FEV_1 % pred	64.7±8.70
Post-bronchodilator FEV_1/FVC	0.57±0.08
Exacerbations in previous 12 months	
1	867 (20.1)
⩾2	16 (0.4)
Baseline treatment	
SABA, SAMA or SABA+SAMA	496 (11.5)
LABA+ICS (fixed-dose or free combination)	1059 (24.5)
LABA (only monotherapy)	1193 (27.6)
LAMA (only monotherapy)	1512 (35.0)
Others/not applicable/unknown[#]	187 (4.4)

Data are presented as n, mean±SD or n (%). COPD: chronic obstructive pulmonary disease; mMRC: modified Medical Research Council dyspnoea scale; BDI: Baseline Dyspnoea Index; CAT: COPD Assessment Test; CCQ: Clinical COPD Questionnaire; FEV_1: forced expiratory volume in 1 s; FVC: forced vital capacity; SABA: short-acting β_2-agonist; SAMA: short-acting muscarinic antagonist; LABA: long-acting β_2-agonist; ICS: inhaled corticosteroid; LAMA: long-acting muscarinic antagonist. [#]: LAMA+ICS, LABA+LAMA, ICS monotherapy, LABA+LAMA+ICS (triple), systemic corticosteroids, methylxanthines, roflumilast and others.

(r_s=−0.437 and −0.451, respectively). In comparison, mMRC showed a weak positive correlation with CAT and CCQ (r_s=0.356 and 0.380, respectively) (figure 1).

At baseline, there was no correlation between lung function (FEV1 % pred) and the health status questionnaires (CAT r_s=−0.068; CCQ r_s=−0.087) or the dyspnoea questionnaire (BDI r_s=0.121) (figure 2). Moreover, there was no correlation between lung function and the CCQ domains (symptoms r_s=−0.070; functional state r_s=−0.096; mental state r_s=−0.045).

Responder analysis

At week 12, the proportion of patients who were responders was highest for TDI (n=2008 (46.4%)), followed by FEV1 (n=1680 (38.9%)), CAT (n=1585 (36.7%)) and CCQ (n=1173 (27.1%)).

Responders to PROs (TDI, CCQ and CAT)

Overall, 2774 (64.2%) patients were classified as responders based on MCID in at least one of the three PROs (CAT, TDI and CCQ). Of these, 1365 (31.6%), 826 (19.1%) and 583 (13.5%) patients showed an improvement in one, two or all three PROs, respectively. To evaluate the association between the PROs, the proportions of patients showing a response to combinations of two or more PROs were assessed. Among responders to two PROs, more patients showed an improvement in both CAT and TDI (n=373 (8.6%))

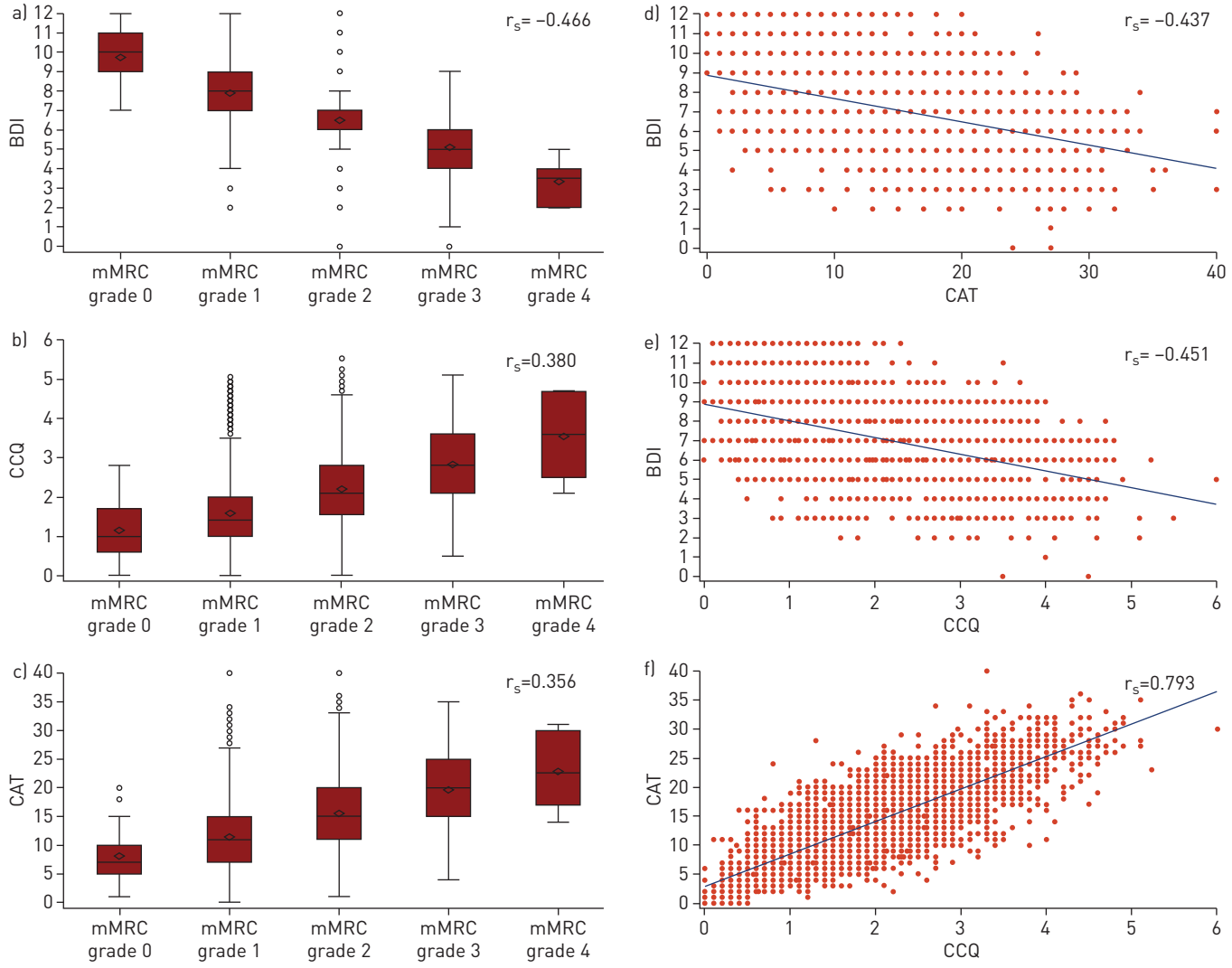

FIGURE 1 Plots showing baseline correlation between the modified Medical Research Council (mMRC) dyspnoea scale, Baseline Dyspnoea Index (BDI), COPD Assessment Test (CAT) and Clinical COPD Questionnaire (CCQ). a–c) Box and whisker plots of mMRC *versus* a) BDI, b) CCQ and c) CAT, where BDI, CCQ and CAT have been represented by mMRC grade. Data are presented as median and interquartile range (box) with minimum and maximum values (whiskers). Diamonds indicate mean values and circles indicate outliers. d–f) Scatterplots of d) BDI *versus* CAT, e) BDI *versus* CCQ and f) CAT *versus* CCQ. Spearman's rank correlation coefficient (r_s) values are indicated. COPD: chronic obstructive pulmonary disease.

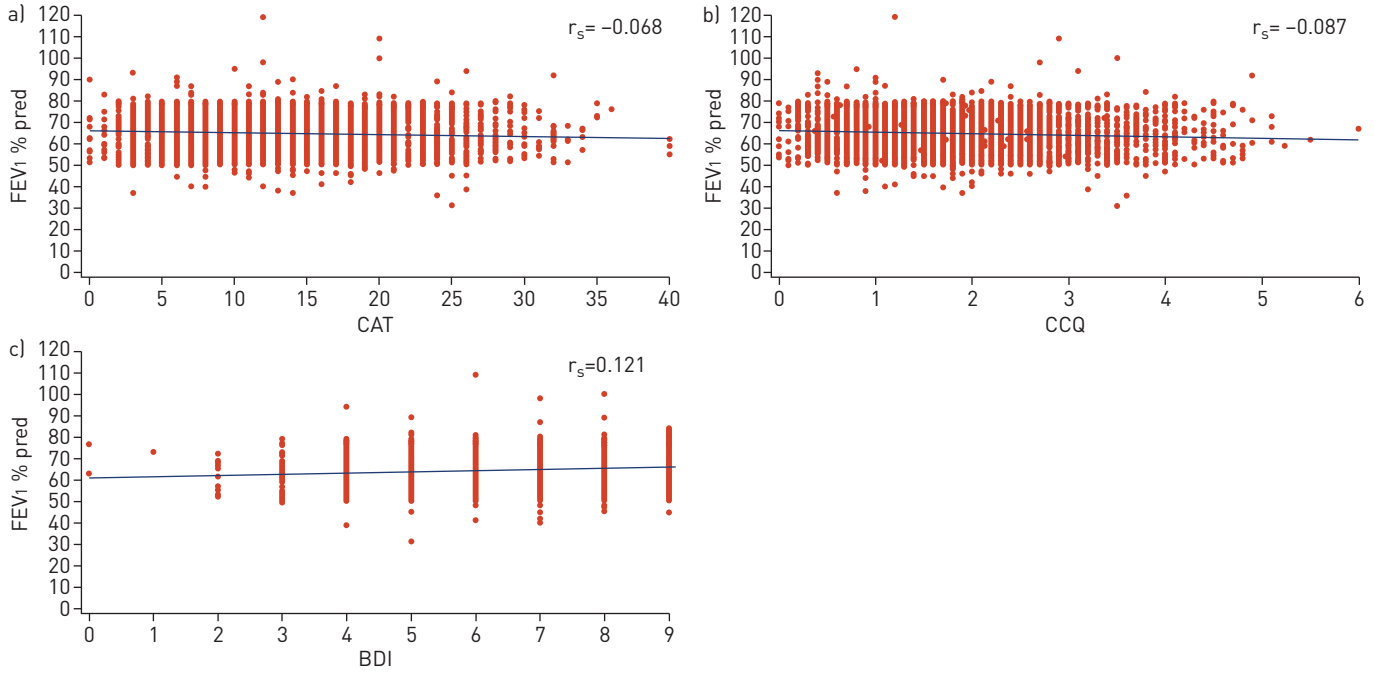

FIGURE 2 Scatterplots showing baseline correlation between forced expiratory volume in 1 s (FEV1) % pred and a) COPD Assessment Test (CAT), b) Clinical COPD Questionnaire (CCQ) and c) Baseline Dyspnoea Index (BDI). Spearman's rank correlation coefficient (r_s) values are indicated. COPD: chronic obstructive pulmonary disease.

compared with CAT and CCQ (n=271 (6.3%)) and TDI and CCQ (n=182 (4.2%)). In terms of responders to a single PRO, more patients showed an improvement in TDI (n=870 (20.1%)) compared with the other PROs (figure 3).

Responders to FEV1, TDI, CCQ and CAT

There were 3235 (74.8%) patients found to be responders based on response to FEV1 and/or PROs. A higher proportion of patients were responders to PROs alone (n=1555 (36.0%)) compared with a

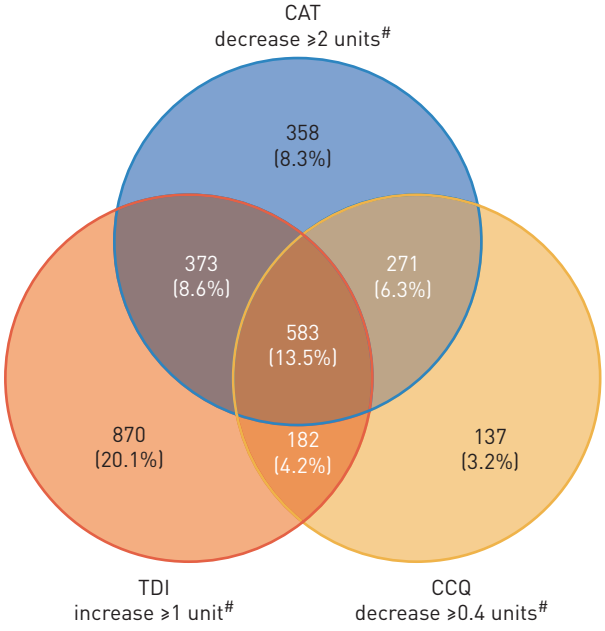

FIGURE 3 Responder analysis based on patient-reported outcomes. COPD: chronic obstructive pulmonary disease; CAT: COPD Assessment Test; TDI: Transition Dyspnoea Index; CCQ: Clinical COPD Questionnaire. Nonresponders n=1550 (35.8%). #: minimal clinically important difference.

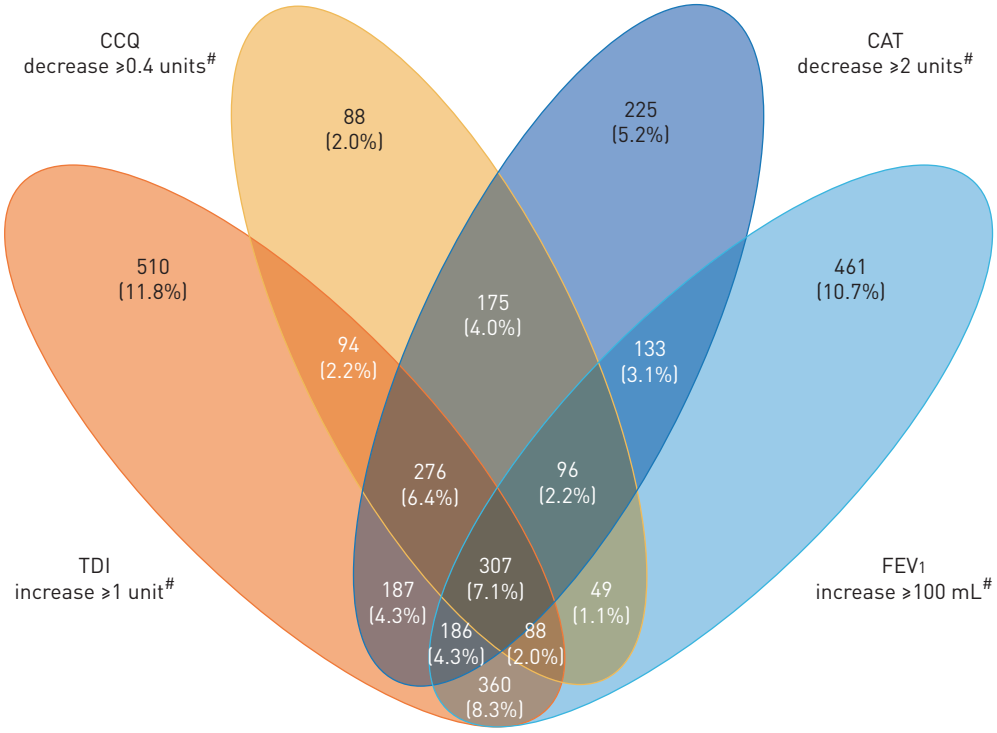

FIGURE 4 Responder analysis based on patient-reported outcomes and forced expiratory volume in 1 s (FEV_1). COPD: chronic obstructive pulmonary disease. CCQ: Clinical COPD Questionnaire; CAT: COPD Assessment Test; TDI: Transition Dyspnoea Index. Nonresponders n=1089 (25.2%). #: minimal clinically important difference.

combination of FEV_1 and one or more PRO responders (n=1219 (28.2%)). Of these, 542 (12.5%), 370 (8.5%) and 307 (7.1%) patients showed an improvement in FEV_1 and one, two or three PROs, respectively. Among FEV_1 and two PRO responders, a higher proportion of patients demonstrated a response to a combination of FEV_1, TDI and CAT (n=186 (4.3%)). Among FEV_1 and single PRO responders, more patients showed a response to a combination of FEV_1 and TDI (n=360 (8.3%)) (figure 4).

Responders by subgroup

Overall, no differences were observed among the subgroups in responders to PROs, with minor differences between sexes, especially for CCQ. Among FEV_1 responders, patients aged <65 years, male and having a bronchodilator reversibility >12% were more likely to be responders. Table 2 shows the descriptive observation of the distribution of patients based on subgroups among responders to PROs and FEV_1.

Discussion

In this analysis of the CRYSTAL study, the PROs CAT and CCQ showed a strong correlation at baseline, while only moderate-to-weak correlations were observed between BDI, CAT, CCQ and mMRC; however, no correlation was observed between lung function (FEV_1) and any of the PROs at baseline. At week 12, there were small overlaps among the patients who responded to PROs (TDI, CCQ and CAT) and those who responded to FEV_1 and PROs. Only a very small proportion of patients presented a clinically relevant response to all four parameters studied. These findings are consistent with the outcomes of previous studies that have also been unable to find a substantial correlation between PROs and FEV_1 [17, 18]. This analysis is the first in a large population of patients with moderate airflow limitation involving three different PROs and lung function responses that had already been pre-specified in the study subgroups.

In clinical trials, CAT and CCQ are routinely used to assess patients' health status and may be considered to be closely related since both assess similar symptoms such as cough and phlegm production. As might be anticipated, these two PROs showed a strong correlation in our statistical analysis, and this is also supported by findings from other *post hoc* studies and meta-analyses [19–21]. However, while the strong correlation was observed at the group level, some variability was observed at the individual patient level. In contrast, we found that at week 12 only a small proportion of patients presented a clinically relevant response to both of these PROs. This suggests that the cross-sectional evaluation of PROs differs from

TABLE 2 Distribution of responders based on patient-reported outcomes (PROs) and forced expiratory volume in 1 s (FEV_1) by subgroup

	Patients[#]	TDI	CCQ	CAT	FEV_1
Age years					
<65	2077	993 (47.8)	610 (29.4)	817 (39.3)	874 (42.1)
⩾65	2247	1015 (45.2)	563 (25.1)	768 (34.2)	806 (35.9)
Sex					
Female	1416	648 (45.8)	416 (29.4)	531 (37.5)	505 (35.7)
Male	2908	1360 (46.8)	757 (26.0)	1054 (36.2)	1175 (40.4)
Smoking status					
Current smoker	2292	1106 (48.3)	664 (29.0)	878 (38.3)	908 (39.6)
Ex-smoker	2028	901 (44.4)	507 (25.0)	705 (34.8)	772 (38.1)
Never-smoker	4	1 (25.0)	2 (50.0)	2 (50.0)	0 (0.0)
Exacerbations in previous 12 months					
0	3441	1602 (46.6)	941 (27.3)	1277 (37.1)	1359 (39.5)
⩾1	883	406 (46.0)	232 (26.3)	308 (34.9)	321 (36.4)
Baseline FEV_1 % pred					
<60%	1367	617 (45.1)	357 (26.1)	476 (34.8)	555 (40.6)
⩾60%	2948	1389 (47.1)	816 (27.7)	1108 (37.6)	1124 (38.1)
Bronchodilator reversibility					
⩽12%	3251	1521 (46.8)	866 (26.6)	1197 (36.8)	1060 (32.6)
>12%	1067	486 (45.5)	307 (28.8)	387 (36.3)	619 (58.0)

Data are presented as n or n (%). TDI: Transition Dyspnoea Index; CCQ: Clinical COPD Questionnaire; CAT: COPD Assessment Test. [#]: total number of patients; some of the responders showed response to more than one PRO and/or FEV_1.

their longitudinal evaluation and their response to treatment interventions. CCQ includes an additional in-depth evaluation of patients' physical and mental activities, whereas CAT focuses more on symptoms such as chest tightness and sleep quality [22]. These may lead to greater differences in sensitivity in assessment of treatment response [4], and may account for the differences between findings from the baseline correlation and week 12 responder analysis.

The unidirectional mMRC scale assesses dyspnoea only in terms of the patient's level of activity [6], while BDI is multidimensional and assesses functional impairment, the extent of tasks performed and the magnitude of effort expended [5]. This may account for the moderate correlation between BDI and mMRC observed in our analysis. Similar findings were reported in a cross-sectional study in French COPD patients, with the authors suggesting that these PROs are not interchangeable for the evaluation of dyspnoea [23].

In the present analysis, the correlations between the dyspnoea (BDI and mMRC) and health status (CAT and CCQ) PROs were weak at baseline, which was also reflected in the responder analysis at week 12 where there was very little overlap between responders to TDI, CAT and CCQ. These findings are not unexpected as the scope of measurements in the different PROs varies considerably. Furthermore, the relative perception of symptomatic burden in patients receiving the same treatment may be different and thus emphasises the distinct response measured by each PRO [5, 7, 8].

We did not find any correlations between FEV_1 and any of the PROs at baseline in this analysis. Furthermore, at week 12, only a small overlap was observed between responders to FEV_1 and the PROs, despite considerable numbers of patients being responders to PROs alone. Other clinical and retrospective studies also report moderate [18, 24, 25] to weak or no correlation between FEV_1 and PROs [26–29]. FEV_1 is purely a metric of airflow limitation, while PROs assess the overall wellbeing of patients. Additionally, COPD is a multicomponent disease and not just restricted to airflow limitation, and therefore lung function (FEV_1) may not accurately reflect disease severity when it is far more complex [30]. This is supported by the revised Global Initiative for Chronic Obstructive Lung Disease strategy, which includes symptoms and exacerbations together with FEV_1 to assess patients and guide therapy [31]. These responder analysis results clearly show that the PROs used to capture different clinically relevant responses in individual patients are complementary to the improvement of airflow limitation.

Interestingly, FEV_1 responders demonstrated a greater response based on combinations of PROs that included TDI and CAT compared with those that included CCQ. This is in agreement with a previous

pooled analysis of 23 clinical trials of long-acting bronchodilators in patients with COPD of variable severity where Donohue *et al.* [32] found weak correlations between ΔFEV1 and PROs (SGRQ and TDI scores). Although the correlation between FEV1 and TDI was weak, these findings suggest the possibility that TDI may more closely reflect the response to treatment in line with improvements in lung function. In our analysis, we found that a greater number of patients responded to TDI alone than the other PROs. This is further supported by the more pronounced improvement in FEV1 and TDI in dual bronchodilator-treated patients reported in the primary publication of the CRYSTAL study [4].

Thus, the findings of our analyses reinforce the need to consider both lung function and PROs when evaluating treatment outcomes in patients and guiding COPD management. This was a robust analysis of the CRYSTAL study involving more than 4000 patients, with the PROs and FEV1 assessed at the beginning and the end of the trial. Importantly for the interpretation of the results, only nonfrequently exacerbating patients with moderate COPD were included in the analysis; hence, the results may not be applicable to patients with more frequent exacerbations or patients with different disease severities. Furthermore, the study only included assessments up to 12 weeks, which may have an impact on the size of the changes observed in the PROs. Moreover, CRYSTAL was an open-label, non-placebo-controlled study and this may have had some influence on the response to PROs. However, the large number of patients and the fact that our analyses were independent of treatment interventions add value to our results. We have used responder analysis with MCID as a measure of clinically relevant response; however, in all these analyses, we need to consider that MCID values are average estimates obtained in groups of patients and may not identify accurately the perceived benefit of each individual patient [33]. The results of the present analysis suggest that further studies are warranted to evaluate the treatment response of PROs and FEV1 in potential composite indexes in other populations of patients with COPD.

Conclusions

Increases in lung function are accompanied by clinically relevant improvements in PROs only in a minority of COPD patients. Our correlation analysis suggests that PROs are not interchangeable; hence, observed treatment success may depend on the parameters selected. Assessments based on both lung function and various PROs may be necessary to obtain a more complete picture of treatment response in patients with COPD and to guide treatment decisions.

Acknowledgements: The authors thank Archana Jayaraman and Santanu Bhadra (Novartis Healthcare Pvt Ltd, Hyderabad, India) for providing writing support in accordance with Good Publication Practice guidelines (www.ismpp.org/gpp3).

Author contributions: All authors have contributed substantially in all aspects of the analysis, including design, conduct, data analysis and interpretation, and were involved in all the stages of manuscript development.

Conflict of interest: K. Kostikas was an employee of Novartis until October 31, 2018. T. Greulich reports receiving support from Novartis for participation as a centre in the current the study; and personal fees for lectures and advisory boards from AstraZeneca, Berlin-Chemie, Boehringer Ingelheim, Chiesi, CSL Behring, GSK and Novartis, grants and personal fees for lectures and advisory boards from Grifols, and grants from the German Centre for Lung Research (DZL), Marburg, Germany, outside the submitted work. A.J. Mackay was a European Respiratory Society Fellow in Industry at Novartis during the preparation of the manuscript. He is a current employee of AstraZeneca and has received speaker fees from Pfizer outside the submitted work. N.S. Lossi is currently an employee of Novartis Pharma GmbH. M. Aalamian-Mattheis is an employee of Novartis Pharma AG. X. Nunez analysed the data in the present study as a statistician employed by TFS. V.A. Pagano analysed the data in the present study as a statistician employed by TFS. F. Patalano is an employee and a shareholder of Novartis Pharma AG. A. Clemens is a full-time employee and shareholder at Novartis Pharma AG. C.F. Vogelmeier reports personal fees from Almirall, Cipla, Berlin-Chemie/Menarini, CSL Behring and Teva, grants and personal fees from AstraZeneca, Ingelheim, Chiesi, GSK, Grifols, Mundipharma, Novartis and Takeda, grants from German Federal Ministry of Education and Research (BMBF) Competence Network Asthma and COPD (ASCONET), Bayer Schering Pharma AG, MSD and Pfizer, outside the submitted work.

Support statement: This analysis has been funded by Novartis Pharma AG, Basel, Switzerland. Funding information for this article has been deposited with the Crossref Funder Registry.

References

1 D'Urzo A, Ferguson GT, van Noord JA, *et al.* Efficacy and safety of once-daily NVA237 in patients with moderate-to-severe COPD: the GLOW1 trial. *Respir Res* 2011; 12: 156.
2 Kerwin EM, Gotfried MH, Lawrence D, *et al.* Efficacy and tolerability of indacaterol 75 µg once daily in patients aged ⩾40 years with chronic obstructive pulmonary disease: results from 2 double-blind, placebo-controlled 12-week studies. *Clin Ther* 2011; 33: 1974–1984.
3 Vogelmeier CF, Bateman ED, Pallante J, *et al.* Efficacy and safety of once-daily QVA149 compared with twice-daily salmeterol–fluticasone in patients with chronic obstructive pulmonary disease (ILLUMINATE): a randomised, double-blind, parallel group study. *Lancet Respir Med* 2013; 1: 51–60.
4 Vogelmeier CF, Gaga M, Aalamian-Mattheis M, *et al.* Efficacy and safety of direct switch to indacaterol/glycopyrronium in patients with moderate COPD: the CRYSTAL open-label randomised trial. *Respir Res* 2017; 18: 140.

5 Mahler DA, Witek TJ Jr. The MCID of the Transition Dyspnea Index is a total score of one unit. *COPD* 2005; 2: 99–103.
6 Bestall JC, Paul EA, Garrod R, *et al.* Usefulness of the Medical Research Council (MRC) dyspnoea scale as a measure of disability in patients with chronic obstructive pulmonary disease. *Thorax* 1999; 54: 581–586.
7 van der Molen T, Willemse BW, Schokker S, *et al.* Development, validity and responsiveness of the Clinical COPD Questionnaire. *Health Qual Life Outcomes* 2003; 1: 13.
8 Jones PW, Harding G, Berry P, *et al.* Development and first validation of the COPD Assessment Test. *Eur Respir J* 2009; 34: 648–654.
9 Jones PW. Health status and the spiral of decline. *COPD* 2009; 6: 59–63.
10 Jones PW, Quirk FH, Baveystock CM. The St George's Respiratory Questionnaire. *Respir Med* 1991; 85: Suppl. B, 25–31.
11 Jones PW, Quirk FH, Baveystock CM, *et al.* A self-complete measure of health status for chronic airflow limitation. The St. George's Respiratory Questionnaire. *Am Rev Respir Dis* 1992; 145: 1321–1327.
12 Tabberer M, Jones PW. The COPD Biomarkers Qualification Consortium St George's Respiratory Questionnaire Manuscripts: output of a consortium to advance drug development. *Chronic Obstr Pulm Dis* 2017; 4: 109–111.
13 Jones P, Miravitlles M, van der Molen T, *et al.* Beyond FEV1 in COPD: a review of patient-reported outcomes and their measurement. *Int J Chron Obstruct Pulmon Dis* 2012; 7: 697–709.
14 Donohue JF. Minimal clinically important differences in COPD lung function. *COPD* 2005; 2: 111–124.
15 Kocks JW, Tuinenga MG, Uil SM, *et al.* Health status measurement in COPD: the minimal clinically important difference of the clinical COPD questionnaire. *Respir Res* 2006; 7: 62.
16 Kon SS, Canavan JL, Jones SE, *et al.* Minimum clinically important difference for the COPD Assessment Test: a prospective analysis. *Lancet Respir Med* 2014; 2: 195–203.
17 Westwood M, Bourbeau J, Jones PW, *et al.* Relationship between FEV1 change and patient-reported outcomes in randomised trials of inhaled bronchodilators for stable COPD: a systematic review. *Respir Res* 2011; 12: 40.
18 Nishimura K, Yasui M, Nishimura T, *et al.* Airflow limitation or static hyperinflation: which is more closely related to dyspnea with activities of daily living in patients with COPD? *Respir Res* 2011; 12: 135.
19 Tsiligianni IG, van der Molen T, Moraitaki D, *et al.* Assessing health status in COPD. A head-to-head comparison between the COPD assessment test (CAT) and the clinical COPD questionnaire (CCQ). *BMC Pulm Med* 2012; 12: 20.
20 Sundh J, Stallberg B, Lisspers K, *et al.* Comparison of the COPD Assessment Test (CAT) and the Clinical COPD Questionnaire (CCQ) in a clinical population. *COPD* 2016; 13: 57–65.
21 Gupta N, Pinto LM, Morogan A, *et al.* The COPD assessment test: a systematic review. *Eur Respir J* 2014; 44: 873–884.
22 Zhou Z, Zhou A, Zhao Y, *et al.* A comparison of the assessment of health status between CCQ and CAT in a Chinese COPD clinical population: a cross-sectional analysis. *Int J Chron Obstruct Pulmon Dis* 2018; 13: 1675–1682.
23 Perez T, Burgel PR, Paillasseur JL, *et al.* Modified Medical Research Council scale vs Baseline Dyspnea Index to evaluate dyspnea in chronic obstructive pulmonary disease. *Int J Chron Obstruct Pulmon Dis* 2015; 10: 1663–1672.
24 Zhou Z, Zhou A, Zhao Y, *et al.* Evaluating the Clinical COPD Questionnaire: a systematic review. *Respirology* 2017; 22: 251–262.
25 Huang WC, Wu MF, Chen HC, *et al.* Features of COPD patients by comparing CAT with mMRC: a retrospective, cross-sectional study. *NPJ Prim Care Respir Med* 2015; 25: 15063.
26 Reda AA, Kotz D, Kocks JW, *et al.* Reliability and validity of the clinical COPD questionnaire and chronic respiratory questionnaire. *Respir Med* 2010; 104: 1675–1682.
27 Jogi J, Ekberg M, Jonson B, *et al.* Ventilation/perfusion SPECT in chronic obstructive pulmonary disease: an evaluation by reference to symptoms, spirometric lung function and emphysema, as assessed with HRCT. *Eur J Nucl Med Mol Imaging* 2011; 38: 1344–1352.
28 Dodd JW, Hogg L, Nolan J, *et al.* The COPD assessment test (CAT): response to pulmonary rehabilitation. A multicentre, prospective study. *Thorax* 2011; 66: 425–429.
29 Camargo LA, Pereira CA. Dyspnea in COPD: beyond the modified Medical Research Council scale. *J Bras Pneumol* 2010; 36: 571–578.
30 Global Initiative for Chronic Obstructive Lung Disease. Global Strategy for the Diagnosis, Management, and Prevention of Chronic Obstructive Pulmonary Disease. 2018. http://goldcopd.org/gold-2018-global-strategy-diagnosis-management-prevention-copd Date last accessed: October 15, 2018.
31 Global Initiative for Chronic Obstructive Lung Disease. Global Strategy for the Diagnosis, Management, and Prevention of Chronic Obstructive Pulmonary Disease. 2017. http://goldcopd.org/gold-2017-global-strategy-diagnosis-management-prevention-copd Date last accessed: October 15, 2018.
32 Donohue JF, Jones PW, Bartels C, *et al.* Correlations between FEV1 and patient-reported outcomes: a pooled analysis of 23 clinical trials in patients with chronic obstructive pulmonary disease. *Pulm Pharmacol Ther* 2018; 49: 11–19.
33 Jones PW, Beeh KM, Chapman KR, *et al.* Minimal clinically important differences in pharmacological trials. *Am J Respir Crit Care Med* 2014; 189: 250–255.