Composite endpoints in COPD: clinically important deterioration in the UPLIFT trial

Klaus F. Rabe1,2*, David M. G. Halpin3, MeiLan K. Han4, Marc Miravitlles5, Dave Singh6, Lars Grönke7, Florian Voß8 and Fernando J. Martinez9

Abstract

Background: Assessments of lung function, exacerbations and health status are common measures of chronic obstructive pulmonary disease (COPD) progression and treatment response in clinical trials. We hypothesised that a composite endpoint could more holistically assess clinically important deterioration (CID) in a COPD clinical trial setting.

Methods: A composite endpoint was tested in a post hoc analysis of 5652 patients with Global Initiative for Chronic Obstructive Lung Disease (GOLD) 2–4 COPD from the 4-year UPLIFT study. Patients received tiotropium 18 μg or placebo.

Results: The composite endpoint included time to first confirmed decrease in trough forced expiratory volume in 1 s (FEV₁) ≥ 100 mL, confirmed increase in St. George’s Respiratory Questionnaire (SGRQ) total score ≥ 4 units, or moderate/severe exacerbation. Most patients (> 80%) experienced CID, with similar incidence among GOLD subgroups. Most confirmed trough FEV₁ (74.6–81.6%) and SGRQ (72.3–78.1%) deteriorations were sustained across the study and in all GOLD subgroups. Patients with CID more frequently experienced subsequent exacerbation (hazard ratio [HR] 1.79; 95% confidence interval [CI] 1.67, 1.92) or death (HR 1.21; 95% CI 1.06, 1.39) by Month 6. CID was responsive to bronchodilator treatment.

Conclusions: Composite endpoints provide additional information on COPD progression and treatment effects in clinical trials.

Trial registration: ClinicalTrials.gov NCT00144339.

Keywords: Tiotropium, Lung function, Exacerbations

Background

Chronic obstructive pulmonary disease (COPD) is a chronic, progressive, heterogeneous disease. The manifestation of disease progression varies over time and between patients. Despite this, most clinical trials conducted in COPD typically focus on a single primary outcome such as forced expiratory volume in 1 s (FEV₁), exacerbation frequency or, less frequently, mortality.

The impact of interventions on disease progression has been measured by the annual rate of decline in FEV₁ over several years [1]. However, there are limitations to this approach. Individuals with a slower rate of decline may dilute any observable treatment benefit in rapidly progressing subgroups. Moreover, individuals with a rapid decline may discontinue studies early, underestimating the true mean rate of decline in the control arm [2].

Primary analyses of clinical trials typically report group mean results, which can be insufficient to detect
clinically important changes at the individual patient level. Furthermore, the focus on only one dimension of COPD may misrepresent real improvements that are meaningful to patients.

Measuring clinically important deterioration (CID) in terms of the most impactful events at the individual patient level might provide a significant benefit in studying the progression and effects of COPD in clinical trials. The three events included in this composite endpoint – trough FEV$_1$, St. George’s Respiratory Questionnaire (SGRQ) score and moderate/severe exacerbation – have been previously used by Singh et al. [3], Anzueto et al. [4] and Greulich et al. [5], and were selected because they are commonly used in clinical trials and are known to have an impact on patients with COPD.

To explore the composite endpoint of CID further, we used a post hoc analysis of the 4-year UPLIFT study. The objectives of this analysis were to test the validity of CID when only including FEV$_1$ and SGRQ events that were confirmed at a subsequent visit, to prove that CID predicts future outcomes, and to explore other elements of CID.

**Methods**

This post hoc analysis assessed time to first CID as time to the first occurrence of at least one of the following: decrease in trough FEV$_1$ from baseline $\geq$100 mL, increase in SGRQ total score from baseline $\geq$4 units or moderate/severe exacerbation (the same components as suggested by Singh et al. [3]). Changes in FEV$_1$ and SGRQ score were always calculated from baseline. Changes in FEV$_1$ were assessed using pre-bronchodilation values, in line with previous studies assessing CID [6–8] and reflecting real-world clinical practice for FEV$_1$ monitoring. A decrease in trough FEV$_1$ $\geq$ 100 mL is considered to be the minimum clinically important change perceived by patients [9, 10] and is within the defined range suggested by the American Thoracic Society/European Respiratory Society task force [11], whereas an increase in SGRQ total score $\geq$ 4 units is considered the minimum clinically important change in quality of life [12].

Unlike for the composite endpoint published by Singh et al. [3], we only included confirmed FEV$_1$ and SGRQ deteriorations, i.e. events that were present during at least two consecutive assessments (5 or 6 months apart). This excluded short-term fluctuations in the disease, which could provide an unreliable indication of CID. If no further assessment was available, but the patient discontinued study medication or died, the event was also considered as confirmed. Confirmed events were not required for exacerbations of COPD.

We have used the term “sustained” to refer to deteriorations that were then maintained at almost every subsequent visit.

**Study design**

Study design details have been previously reported [13] and are briefly summarised below. UPLIFT (ClinicalTrials.gov: NCT00144339) was a 4-year, randomised, double-blind, parallel-group study comparing tiotropium 18 μg, administered once daily via the HandiHaler®, with matching placebo [14]. The UPLIFT study was conducted in 37 countries [14]. Patients were aged $\geq$40 years, with a smoking history of $\geq$10 pack-years and moderate-to-very severe COPD (Global Initiative for Chronic Obstructive Lung Disease [GOLD] 2–4 [15]). For further details, see the Supplementary Methods. The protocol was approved by the ethics committee at each centre, and all patients provided written, informed consent.

Spirometric testing was performed at randomisation, at the Day 30 visit and at visits every 6 months up to Month 48. SGRQ was assessed at randomisation and every 6 months up to Month 48. Exacerbations and associated hospital admissions were recorded on case report forms at every visit. The two primary endpoints were pre- and post-bronchodilation yearly rate of decline in mean FEV$_1$.

**Statistical analysis**

For time-to-event endpoints, hazard ratios (HRs), 95% confidence intervals (CIs) and $P$ values were calculated using a Cox proportional hazards model. Patients without CID events were censored at the treatment stop date.

To assess the association of CID with future outcomes, patients experiencing a CID event within the first 6 months were compared with those not experiencing the event. For this analysis, the time to first moderate/severe exacerbation was calculated from Month 6 (180 days) to the first subsequent event or treatment discontinuation. Time to death was calculated from Month 6 (180 days) to the date of death or the end of the vital status follow-up (Day 1470).

**Results**

Patient dispositions have been reported previously [14]. Overall, 5652 patients received treatment (2811 tiotropium; 2841 placebo) and had baseline measurements for both FEV$_1$ and SGRQ. GOLD stage at baseline (based on post-bronchodilator FEV$_1$) was available for 5589 patients (GOLD 2: 1293 placebo, 1310 tiotropium; GOLD 3: 1266 placebo, 1239 tiotropium; GOLD 4: 250 placebo, 231 tiotropium).

**Incidence of CID**

Most patients in the total population (83.9%) experienced at least one CID during the study (Table 1). Exacerbations were more frequent than FEV$_1$ or SGRQ decline (Table 1).
The contribution of exacerbations to the composite endpoint became more pronounced whereas the contribution of FEV₁ became less pronounced as COPD severity (GOLD stage) increased in the total population (Table 1).

Time to first event for each component is shown in e-Figure 1.

Overall, about half of patients experienced at least two of the three events qualifying as CID, whereas fewer patients experienced all three events (Fig. 1a). A similar proportion of patients in each GOLD group experienced at least two CID events (Fig. 1b–d). The incidence of all three CID events was also similar for GOLD 2 and 3 patients, whereas few GOLD 4 patients experienced all three CID events (Fig. 1b–d).

Overall, most confirmed events were sustained at subsequent visits. Confirmed trough FEV₁ decline was sustained at 12–48 months after the initial event in 74.6–81.6% of patients (Table 2). Confirmed SGRQ deterioration was also sustained at 12–42 months after the initial event in 72.3–78.1% of patients (Table 2). This pattern was comparable with the GOLD subgroups (e-Table 1), although patient numbers were low for the GOLD 4 subgroup.

For unconfirmed events (reported at one timepoint), the proportion of patients whose FEV₁ decline or SGRQ deterioration was sustained was lower: 51.6–71.9% of patients still had the FEV₁ decline 6–48 months after first decline, and 52.5–65.5% still had SGRQ deterioration (e-Table 2).

In addition, in patients who had confirmed events, mean FEV₁ remained at least 193 mL worse than baseline for the rest of the trial (Table 2). For unconfirmed events, mean FEV₁ in patients with an event ranged from 95 mL worse than baseline at Month 6 to 142 mL worse than baseline at Month 24 and 213 mL at Month 48. In patients with SGRQ deterioration, mean increase was > 10 units for the rest of the trial for confirmed events, but ranged from 4.7 to 8.3 units for unconfirmed events.

### Table 1 Incidence of CID and risk of first CID occurrence in total population and by GOLD stage

| Total UPLIFT population (placebo and tiotropium combined) | Overall population | GOLD 2 | GOLD 3 | GOLD 4 |
|----------------------------------------------------------|-------------------|--------|--------|--------|
| Number of patients, n (%)                                 | 5652 (100.0)      | 2603 (100.0) | 2505 (100.0) | 481 (100.0) |
| At least one of SGRQ deterioration (decrease of ≥4 units), trough FEV₁ decline ≥100 mL or moderate/severe exacerbation | 4741 (83.9)       | 2175 (83.6) | 2127 (84.9) | 388 (80.7) |
| Moderate/severe exacerbation                              | 3814 (67.5)       | 1615 (62.0) | 1796 (71.7) | 357 (74.2) |
| Trough FEV₁ decline ≥100 mL                              | 2503 (44.3)       | 1344 (51.6) | 1031 (41.2) | 100 (20.8) |
| SGRQ score increase ≥4 units                             | 2339 (41.4)       | 1081 (41.5) | 1054 (42.1) | 178 (37.0) |

CID clinically important deterioration, FEV₁, forced expiratory volume in 1 s, GOLD Global Initiative for Chronic Obstructive Lung Disease, SGRQ St. George’s Respiratory Questionnaire

### Relative timing of events

The pattern and timing of clinically relevant events was highly variable for individual patients. Of patients who experienced both confirmed FEV₁ decline and SGRQ deterioration, it was unusual to experience both events at the same assessment (Table 3). The time from FEV₁ decline to subsequent SGRQ deterioration was slightly longer than the time from SGRQ deterioration to subsequent FEV₁ decline (Fig. 2).

For patients who experienced both FEV₁ decline and SGRQ deterioration, those with less spirometric obstruction appeared more likely to experience confirmed FEV₁ decline prior to confirmed SGRQ deterioration (GOLD 2: 50.1%; GOLD 3: 41.6%; GOLD 4: 37.0%) (Table 3).

Exacerbations demonstrated a greater contribution to the composite endpoint in more severe patients. Patients with more severe COPD were more likely to experience an exacerbation prior to experiencing FEV₁ decline or SGRQ deterioration (Table 3).

### Response to treatment

The time to first CID event, and time to first occurrence of the individual components, was sensitive to therapeutic intervention (Table 4). Time to first CID, two CID events and all three CID events was longer with tiotropium than with placebo (Table 4 and Fig. 1a). This trend was observed in GOLD 2 and 3 subgroups, but less so with GOLD 4 patients (Fig. 1b–d).

### Risk of future exacerbations and death

Patients who had CID events by Month 6 were more likely to experience a moderate or severe exacerbation (HR 1.79; 95% CI 1.67, 1.92), a severe exacerbation (HR 1.67; 95% CI 1.49, 1.86) or death (HR 1.21; 95% CI 1.06, 1.39) (Table 5). The increase in the risk of exacerbations was qualitatively similar for GOLD 2–4 subgroups (Table 5).

When the composite endpoint was broken down into its component events, the HRs for future exacerbations were smaller for FEV₁ decline and SGRQ deterioration...
by Month 6 than for the composite endpoint in the overall population, and among GOLD 2 and GOLD 3 COPD patients (Table 5). Exacerbations within 6 months had higher HRs for any exacerbation and for severe exacerbations than the composite endpoint.

For unconfirmed events, the HRs for long-term outcomes were lower than for the sustained events (Table 5 and e-Table 3).

Investigating future events by CID status at Month 12 showed similar results (e-Table 4).

Mortality analysis with CID
Additional analyses using time to composite event or time to one of the component events as a time-varying covariate were performed. The HR for death for patients with a CID event versus patients without an event was 1.69 (95% CI 1.42, 2.01) (e-Table 5).

Using a stepwise Cox regression model to adjust for important baseline predictors of mortality had little effect on the predictive performance of the composite (e-Table 5). When all three components were included as separate predictors, all were associated with increased mortality risk (e-Table 5).

To validate these findings, the results in e-Tables 4, 6 and 7 are presented for the placebo and tiotropium arms separately. The HRs are slightly higher in the tiotropium arm, which may be related to the larger number of events in the placebo arm before Month 6. The results in e-Tables 6 and 7 are similar between arms and confirm the results in the total population.
Composite endpoints have only recently been introduced in post hoc analyses of COPD clinical trials [3–6, 16]. Here, we conducted a post hoc analysis of the UPLIFT study. This analysis demonstrated the importance of using confirmed events in CID analysis and that CID predicts future outcomes. It also confirmed that the components of this composite endpoint behaved differently based on the baseline FEV\textsubscript{1} of the individual patient. These data suggest that sustained decline in trough FEV\textsubscript{1}, sustained deterioration in SGRQ score of ≥4 units and a moderate/severe exacerbation are appropriate components of a composite endpoint for the assessment of CID in COPD.

### Table 2 Patients with FEV\textsubscript{1} decline or SGRQ deterioration in the total population

| Patients with confirmed FEV\textsubscript{1} decline or SGRQ deterioration at time points after the initial decline (available assessments) | Months after confirmed FEV\textsubscript{1} decline or SGRQ increase |
|---|---|
| 6 | 12 | 18 | 24 | 30 | 36 | 42 | 48 |
| **FEV\textsubscript{1} decline** | | | | | | | |
| No. of patients with decline, n (%) | 1924/1924 (100) | 1239/1661 (74.6) | 1091/1447 (75.4) | 895/1199 (74.6) | 754/966 (78.1) | 604/755 (80.0) | 419/549 (76.3) | 280/343 (81.6) |
| Mean FEV\textsubscript{1} change from baseline, mL (SD) | –223 (114) | –193 (176) | –202 (187) | –214 (195) | –237 (212) | –256 (224) | –258 (224) | –258 (224) |
| Median FEV\textsubscript{1} change from baseline mL (min, max) | –190 (–1090, –40) | –180 (–1240, 850) | –200 (–1000, 0.630) | –210 (–1070, 0.740) | –240 (–1190, 650) | –240 (–1200, 490) | –250 (–980, 460) | –270 (–920, 670) |
| **SGRQ deterioration** | | | | | | | |
| No. of patients with deterioration, n (%) | 1785/1785 (100) | 1140/1510 (75.5) | 941/1301 (72.3) | 776/1067 (72.7) | 629/843 (74.6) | 462/617 (74.9) | 303/388 (78.1) |
| Mean SGRQ change from baseline (SD) | 12.5 (7.5) | 10.6 (10.7) | 10.2 (11.4) | 10.7 (12.3) | 11.3 (12.6) | 11.8 (12.8) | 12.3 (12.8) |
| Median SGRQ score change from baseline (min, max) | 10.5 (4.0, 53.1) | 9.5 (48.4, 53.9) | 9.3 (44.2, 67.8) | 10.2 (44.5, 69.6) | 10.7 (36.6, 54.6) | 10.9 (33.8, 50.1) | 11.9 (29.8, 47.3) |
| For patients with confirmed decline at Month 1, their assessments at Months 6, 12, 18, 24, 30, 36, 42 and 48 are used as time from first confirmed deterioration. Unscheduled visits were excluded for this analysis. Only patients with confirmed FEV\textsubscript{1} decline or confirmed SGRQ deterioration and available assessments at each time point after the deterioration were included in the analysis. FEV\textsubscript{1} forced expiratory volume in 1 s, SD standard deviation, SGRQ St. George’s Respiratory Questionnaire. |

### Discussion

Composite endpoints have only recently been introduced in post hoc analyses of COPD clinical trials [3–6, 16]. Here, we conducted a post hoc analysis of the UPLIFT study. This analysis demonstrated the importance of using confirmed events in CID analysis and that CID predicts future outcomes. It also confirmed that the components of this composite endpoint behaved differently based on the baseline FEV\textsubscript{1} of the individual patient. These data suggest that sustained decline in trough FEV\textsubscript{1}, sustained deterioration in SGRQ score of ≥4 units and a moderate/severe exacerbation are appropriate components of a composite endpoint for the assessment of CID in COPD.

### Table 3 Timing of FEV\textsubscript{1} decline and SGRQ deterioration in the overall population and GOLD subgroups

| Patients with both confirmed FEV\textsubscript{1} decline and confirmed SGRQ deterioration, n (%) | Overall | GOLD 2 | GOLD 3 | GOLD 4 |
|---|---|---|---|---|
| No. of patients with both confirmed FEV\textsubscript{1} decline and confirmed SGRQ deterioration, n (%) | 1344 (100.0) | 698 (100.0) | 575 (100.0) | 54 (100.0) |
| FEV\textsubscript{1} decline before SGRQ deterioration | 620 (46.1) | 350 (50.1) | 239 (41.6) | 20 (37.0) |
| SGRQ deterioration before FEV\textsubscript{1} decline | 484 (36.0) | 227 (32.5) | 228 (39.7) | 25 (46.3) |
| Moderate/severe exacerbation before FEV\textsubscript{1} and SGRQ deterioration | 546 (40.6) | 246 (35.2) | 269 (46.8) | 24 (44.4) |
| Patients with confirmed FEV\textsubscript{1} decline and no confirmed SGRQ deterioration, n (%) | 1160 (100.0) | 644 (100.0) | 459 (100.0) | 46 (100.0) |
| Moderate/severe exacerbation before FEV\textsubscript{1} decline | 680 (59.0) | 378 (58.5) | 228 (49.4) | 23 (41.3) |
| Unconfirmed SGRQ deterioration | 302 (26.0) | 180 (28.0) | 109 (23.7) | 10 (21.7) |
| On same assessment as confirmed FEV\textsubscript{1} decline | 34 (2.9) | 20 (3.1) | 12 (2.6) | 1 (2.2) |
| Before confirmed FEV\textsubscript{1} decline | 116 (10.0) | 72 (11.2) | 36 (7.8) | 7 (15.2) |
| Patients with confirmed SGRQ and no confirmed FEV\textsubscript{1} deterioration, n (%) | 995 (100.0) | 383 (100.0) | 479 (100.0) | 124 (100.0) |
| Moderate/severe exacerbation before SGRQ deterioration | 574 (57.7) | 184 (48.0) | 305 (63.7) | 80 (64.5) |
| Unconfirmed FEV\textsubscript{1} deterioration | 284 (28.5) | 141 (36.8) | 126 (26.3) | 15 (12.1) |
| On same assessment as confirmed SGRQ deterioration | 50 (5.0) | 23 (6.0) | 24 (5.0) | 3 (2.4) |
| Before confirmed SGRQ deterioration | 151 (15.2) | 80 (20.9) | 64 (13.4) | 5 (4.0) |

For patients with confirmed decline at Month 1, their assessments at Months 6, 12, 18, 24, 30, 36, 42 and 48 are used as time from first confirmed deterioration. Unscheduled visits were excluded for this analysis. Only patients with confirmed FEV\textsubscript{1} decline or confirmed SGRQ deterioration and available assessments at each time point after the deterioration were included in the analysis. FEV\textsubscript{1} forced expiratory volume in 1 s, GOLD Global Initiative for Chronic Obstructive Lung Disease, SGRQ St. George’s Respiratory Questionnaire.
patients enrolled in COPD clinical trials. Earlier analyses of the UPLIFT trial have focused on exacerbations or a composite endpoint of the more severe events (exacerbations, respiratory failure, death and trial withdrawal due to worsening COPD), which do not provide an in-depth view of the impact of COPD on patient symptoms or quality of life [16, 17]. In the current analysis we focus on a composite endpoint of validated clinically important criteria (FEV1, SGRQ and exacerbations) to provide a more complete assessment of the impact on patients.

The individual components of the composite endpoint comprise characteristics of COPD that impact patient well-being, are clinically relevant events for the patient and predict future outcomes [15]. Although there are other parameters that could be included in such an endpoint, the components included are relatively easy to include in clinical trials and have established minimum clinically important differences.

Most deteriorations in FEV1 and SGRQ that were confirmed at a second visit were maintained for the rest of the 4-year UPLIFT study. Some publications of composite endpoints in COPD have not required confirmation at a subsequent visit [3, 16]. We believe that counting only confirmed FEV1 and SGRQ deteriorations improves the reliability of the composite endpoint, as it excludes short-term variation and inconsistent measurements. This is supported by the low proportion of patients with unconfirmed events whose FEV1 or SGRQ deterioration is sustained at subsequent timepoints, and by the lower HRs for long-term outcomes with unconfirmed events compared with confirmed events.

Our analysis demonstrated that the components of the composite endpoint rarely occur at the same time in an individual patient. Most patients experience decline of trough FEV1, deterioration of SGRQ score and moderate/severe exacerbations on an individualised time scale. This supports the value of individual components in a
composite endpoint. The stepwise regression data also show that each component independently contributes to increased mortality risk. The composite endpoint is also sensitive to pharmacological treatment, and is similar to the findings of Singh et al., who observed a reduction in first CID with umeclidinium/vilanterol versus placebo in a post hoc study of the same composite endpoint [3].

Other post hoc analyses have used slightly different composite endpoints: FEV1, SGRQ and Transition Dyspnea Index focal score [6]; FEV1 or Transition Dyspnea Index; an increase in SGRQ; and a moderate-to-severe COPD exacerbation [4].

In all the publications that included FEV1, the strongest driver of CID in each of the analysis populations was lung function [3–6]. In contrast to these previous results, the most commonly reported endpoint in our study was exacerbations, perhaps because the UPLIFT study was 4 years long compared with the shorter (maximum 26 weeks) duration of the previous studies [3]. Our analysis showed a high overall frequency of CID for both treatment arms, which is expected due to the long study duration.

Lastly, we have shown that patients considered to have a CID early in the UPLIFT study (within the first 6 months) had worse outcomes for the 42-month remainder of the study; this was also confirmed in an analysis using CID as a time-varying covariate. These outcomes support results from previous analyses of the shorter TORCH and ECLIPSE studies. The 4-year length of our study provided valuable information on sustained CID and the relationship between clinically important events that could not be ascertained in clinical trials of shorter duration.

The study had limitations. In addition, relatively few patients with GOLD 4 lung function impairment were enrolled. Additionally, GOLD 4 patients have a lower baseline FEV1 than GOLD 2 or 3 patients, and as such, declines in FEV1 of ≥100 mL were less common, and would be expected to be more debilitating, in these patients. This should be considered in future studies, where percentage declines may be considered as an alternative clinically significant decline. The composite index considers the parameters SGRQ and moderate/severe exacerbations, which could be seen as subjective;
therefore, it is possible that this could introduce some variability in the results. Also, this was a post hoc analysis, although the large population and long follow-up time allowed for a satisfactory number of events to be observed.

**Conclusions**

We believe these results indicate that a composite endpoint of CID is a promising endpoint to assess disease activity in COPD clinical trials and may be a useful outcome that helps clinicians interpret the implications of trial results for individual patient management. Development of prospective studies is required to determine whether patients who experience disease progression (i.e. those who experience CID) at an increased rate can be identified earlier. By stratifying patients based on time to CID in a clinical trial database, it may be possible to identify characteristics that are associated with longer-term poor outcomes that could be useful for identifying which patients require further treatment earlier. Moreover, the composite endpoint may also serve to reduce patient numbers in clinical trials, as large numbers of patients are required to generate enough statistical power to detect a single outcome within patients with moderate COPD [18]. The length of trials may also be reduced, thereby limiting challenges such as patient discontinuation and cost that are prohibitive in trials of increased duration. Prospective studies are needed on the use of this concept to understand the sensitivity and efficacy of current and potential therapies.

| Table 5 Risk of exacerbation or death from Month 6 onwards by CID status at Month 6 in the overall population and by GOLD |
| --- |
| **Outcome at Month 6** | Patients with any CID event vs. patients without, HR (95% CI) | Patients with confirmed FEV₁ decline vs. patients without, HR (95% CI) | Patients with confirmed SGRQ deterioration vs. patients without, HR (95% CI) | Patients with moderate/severe exacerbation vs. patients without, HR (95% CI) |
| Overall population | 1.79 (1.67, 1.92) | 1.11 (1.02, 1.22) | 1.30 (1.18, 1.43) | 2.36 (2.20, 2.53) |
| Moderate/severe exacerbation | 1.67 (1.49, 1.86) | 1.06 (0.92, 1.23) | 1.66 (1.44, 1.91) | 1.88 (1.68, 2.11) |
| Severe exacerbation | 1.21 (1.06, 1.39) | 1.09 (0.92, 1.31) | 1.27 (1.05, 1.54) | 1.22 (1.05, 1.41) |
| Death up to Day 1470 | 1.21 (1.06, 1.39) | 1.09 (0.92, 1.31) | 1.27 (1.05, 1.54) | 1.22 (1.05, 1.41) |
| GOLD 2 | 1.73 (1.56, 1.92) | 1.15 (1.01, 1.30) | 1.24 (1.07, 1.44) | 2.49 (2.23, 2.78) |
| Moderate/severe exacerbation | 1.58 (1.30, 1.92) | 1.23 (0.97, 1.55) | 1.77 (1.37, 2.27) | 1.79 (1.46, 2.21) |
| Severe exacerbation | 1.21 (0.95, 1.55) | 1.26 (0.94, 1.68) | 1.36 (0.98, 1.90) | 1.01 (0.75, 1.35) |
| Death up to Day 1470 | 1.21 (0.95, 1.55) | 1.26 (0.94, 1.68) | 1.36 (0.98, 1.90) | 1.01 (0.75, 1.35) |
| GOLD 3 | 1.84 (1.67, 2.04) | 1.18 (1.03, 1.35) | 1.32 (1.15, 1.52) | 2.21 (1.99, 2.45) |
| Moderate/severe exacerbation | 1.69 (1.45, 1.96) | 1.08 (0.88, 1.33) | 1.57 (1.29, 1.92) | 1.79 (1.54, 2.09) |
| Severe exacerbation | 1.20 (0.99, 1.45) | 1.17 (0.91, 1.51) | 1.30 (1.00, 1.70) | 1.15 (0.94, 1.41) |
| Death up to Day 1470 | 1.20 (0.99, 1.45) | 1.17 (0.91, 1.51) | 1.30 (1.00, 1.70) | 1.15 (0.94, 1.41) |
| GOLD 4 | 1.84 (1.47, 2.32) | 1.35 (0.85, 2.15) | 1.57 (1.14, 2.16) | 1.97 (1.55, 2.50) |
| Moderate/severe exacerbation | 1.69 (1.26, 2.25) | 2.38 (1.46, 3.88) | 1.81 (1.23, 2.67) | 1.46 (1.09, 1.97) |
| Severe exacerbation | 1.16 (0.84, 1.60) | 1.52 (0.88, 2.63) | 1.00 (0.60, 1.66) | 1.22 (0.87, 1.70) |

Time to death was calculated from Month 6 (180 days) to the date of death or the end of the vital status follow-up (Day 1470).

CI: confidence interval, CID: clinically important deterioration, GOLD: Global Initiative for Chronic Obstructive Lung Disease, HR: hazard ratio, SGRQ: St. George's Respiratory Questionnaire.
Supplementary information

Supplementary information accompanies this paper at https://doi.org/10.1186/s12931-020-01431-y.

Additional file 1: Supplementary Table 1. Patients with FEV1 decline or SGRQ deterioration and at least two confirmed events or unconfirmed clinically important deterioration (CID) event between baseline and Month 12 in the tiotropium and placebo arms and in the total population. For patients with FEV1 decline or SGRQ deterioration, Risk of exacerbation or death by confirmed clinically important deterioration status calculated using clinically important deterioration event as a time-varying covariate.

Supplementary Table 2. Patients with FEV1 decline or SGRQ deterioration in the total UPLIFT population 6, 12, 18, 24, 30, 36, and 42 months after the initial confirmed FEV1 decline or SGRQ deterioration. Confirmed FEV1 decline or SGRQ deterioration.

Supplementary Table 3. Risk of exacerbation or death by confirmed clinically important deterioration status at Months 6 and 12.

Supplementary Table 4. Risk of exacerbation or death by confirmed clinically important deterioration status at Month 6 in the tiotropium and placebo arms, and at Month 12 in the tiotropium and placebo arms and total population.

Supplementary Table 5. Risk of exacerbation or death by confirmed clinically important deterioration status calculated using clinically important deterioration event as a time-varying covariate. Patients with FEV1 decline or SGRQ deterioration and at least two confirmed events or unconfirmed clinically important deterioration (CID) event between baseline and Month 12 in the tiotropium and placebo arms.

Supplementary Table 6. Timing of FEV1 decline and SGRQ deterioration relative to each other in the tiotropium and placebo arms.

Supplementary Table 7. Timing of FEV1 decline and SGRQ deterioration relative to each other in the tiotropium and placebo arms.

Supplementary Figure 1. Time to first event for (A) trough FEV1 decline ≥100 ml, (B) SGRQ increase ≥4 units and (C) moderate/severe exacerbation.

Abbreviations
CI: Confidence interval; CID: Clinically important deterioration; COPD: Chronic obstructive pulmonary disease; FEV1: Forced expiratory volume in 1 s; GOLD: Global Initiative for Chronic Obstructive Lung Disease; HR: Hazard ratio; SGRQ: St. George's Respiratory Questionnaire

Acknowledgements
The authors meet the criteria for authorship as recommended by the International Committee of Medical Journal Editors. They take full responsibility for the scope, direction, content of, and editorial decisions relating to the manuscript. All authors provided final approval of the manuscript. The authors received no compensation related to the development of the manuscript. Medical writing assistance was provided by Laura Badtke of Complete HealthVision and Claire Scofield of MediTech Media, which was contracted and compensated by Boehringer Ingelheim Pharma GmbH & Co. KG. Dave Singh is supported by the National Institute for Health Research (NIHR) Manchester Biomedical Research Centre (BRC).

Authors’ contributions
All the authors have made substantial contributions to the study design, data acquisition, analysis or interpretation, drafting the article, or critically revising the content. All authors provided final approval of the submitted version.

Funding
Supported by Boehringer Ingelheim Pharma GmbH & Co. KG. Medical writing assistance was contracted and compensated by Boehringer Ingelheim Pharma GmbH & Co. KG.

Availability of data and materials
Data are available from the corresponding author upon reasonable request.

Ethics approval and consent to participate
All studies included in this analysis were performed in accordance with the provisions of the Declaration of Helsinki (1996 version), in accordance with the International Conference on Harmonization Tripartite Guideline for Good Clinical Practice, and in accordance with applicable regulatory requirements and Boehringer Ingelheim Standard Operating Procedures. All patients provided written informed consent. This article does not report individual patient data; all data presented here are anonymised. The clinical trial protocols and the informed consent and patient information forms were reviewed and received approval/favourable opinion from a constituted local institutional review board or an independent ethics committee at each centre prior to the start of the study.

Competing interests
KF.R. reports personal fees from Boehringer Ingelheim, AstraZeneca, Novartis and Chiesi, and grants from Boehringer Ingelheim, AstraZeneca and Takeda, outside the submitted work. D.M.G.H. reports personal fees from AstraZeneca, Chiesi and Pfizer, and grants and personal fees from Boehringer Ingelheim, GlaxoSmithKline and Novartis, outside the submitted work. M.M. reports speaker fees from AstraZeneca, Boehringer Ingelheim, Chiesi, Cipla, Menarini, Rovi, Bial, Zambon, CSL Behring, Grifols and Novartis, and consulting fees from AstraZeneca, Boehringer Ingelheim, Chiesi, GlaxoSmithKline, Bial, Gebro Pharma, CSL Behring, Laboratorios Esteve, Fermer, Mereos Biopharma, Verona Pharma, TEVA, pH Pharma, Novartis and Grifols and research grants from GlaxoSmithKline and Grifols. MK.H. reports consultancy fees from Boehringer Ingelheim, GlaxoSmithKline, AstraZeneca and Mylan, speaker fees from Boehringer Ingelheim, and research support from Novartis Senvion, outside the submitted work. F.J.M. reports grants from NHLBI during the conduct of the study, grants from the National Institutes of Health, personal fees from Continuing Education, Forest Laboratories, GlaxoSmithKline, Nycomed/Takeda, AstraZeneca, Boehringer Ingelheim, Bellerophon Therapeutics (formerly Ikaria), Genentech, Novartis, Pearl, Roche, Senvion, Theracon, OME Incite, the Annenberg Center for Health Sciences at Eisenhower, IntThught, the National Association for Continuing Education, Paradigm Medical Communications, LLC, PeerVoice, UpToDate, Haymarket Communications, the Western Society of Allergy and Immunology, ProterixBio (formerly Bioscale), Unity Biotechnology, Concert Pharmaceuticals, Lucid, Methodist Hospital, Columbia University, Prime Healthcare Ltd., WebMD, PeerView Network, the California Society of Allergy and Immunology, Chiesi and the Puerto Rico Thoracic Society, and advisory board participation for Janssen, outside the submitted work. D.S reports personal fees from Apellis, Cipla, Genentech, Peptinnovate and Skyepharma, and grants and personal fees from AstraZeneca, Boehringer Ingelheim, Chiesi, Glenmark, Merck, Mundipharma, Novartis, Pfizer, Pulmatrix, Teva, Theravance and Verona, outside the submitted work. L.G. was an employee of Boehringer Ingelheim at the time of the study and is now employed by CSL Behring. F.V. is an employee of Boehringer Ingelheim.

Author details
1Member of the German Center for Lung Research (DZL), LungClinic Grossharsdorf, Wöhrendamm 80, 22927 Grossharsdorf, Germany. 2Member of the German Center for Lung Research (DZL), Christian Albrechts University Kiel, Kiel, Germany. 3University of Exeter Medical School, College of Medicine and Health, University of Exeter, Exeter, UK. 4Division of Pulmonary and Critical Care, University of Michigan Health System, Ann Arbor, MI, USA. 5Pneumology Department, Hospital Universitari Vall d’Hebron, Vall d’Hebron Institut de Recerca (VHIR), Vall d’Hebron Barcelona Hospital Campus, Ciber de Enfermedades Respiratorias (CIBERSERES), Barcelona, Spain. 6Medicines Evaluation Unit (MEU), University of Manchester, Manchester University NHS Foundation Trust, Manchester, UK. 7Clinical Development, CSL Behring GmbH, Marburg, Germany. 8Boehringer Ingelheim Pharma GmbH & Co. KG, Ingelheim am Rhein, Germany. 9Department of Internal Medicine, Weill Cornell School of Medicine, New York, NY, USA.

Received: 7 November 2019 Accepted: 21 June 2020
Published online: 09 July 2020

References
1. Halm EA, Tashkin DP. Defining disease modification in chronic obstructive pulmonary disease. COPD. 2009;6:211–25.
2. Tantucci C, Modina D. Lung function decline in COPD. Int J Chron Obstruct Pulmon Dis. 2012;7:95–9.
3. Singh D, Maleki-Yazdi MR, Tombs L, Iklab A, Fahy WA, Nay A. Prevention of clinically important deteriorations in COPD with umeclidinium/vilanterol. Int J Chron Obstruct Pulmon Dis. 2016;11:1413–24.
4. Anzueto AR, Vogelmeier CF, Kostikas K, Mezzi K, Fucile S, Bader G, Shen S, Banerji D, Fogel R. The effect of indacaterol/glycopyrronium versus tiotropium or salmeterol/fluticasone on the prevention of clinically important deterioration in COPD. Int J Chron Obstruct Pulmon Dis. 2017;12:1325–37.

Institutional Review Board or an Independent Ethics Committee at each centre prior to the start of the study.
5. Greulich T, Kostikas K, Gaga M, Alamian-Mattheis M, Lossi NS, Patalano F, Nunez X, Pagano VA, Fogel R, Vogelmeier CF, Clemens A. Indacaterol/glycopyrronium reduces the risk of clinically important deterioration after direct switch from baseline therapies in patients with moderate COPD: a post hoc analysis of the CRYSTAL study. Int J Chron Obstruct Pulmon Dis. 2018;13:1229–37.

6. Singh D, D’Urzo AD, Chuecos F, Munoz A, Garcia GE. Reduction in clinically important deterioration in chronic obstructive pulmonary disease with aclidinium/formoterol. Respir Res. 2017;18:106.

7. Martinez FJ, Lipworth BJ, Rabe KF, Collier DJ, Ferguson GT, Sethi S, Feldman GI, O’Brien G, Jenkins M, Reinsen C. Benefits of glycopyrrolate/formoterol fumarate metered dose inhaler (GFF MDI) in improving lung function and reducing exacerbations in patients with moderate-to-very severe COPD: a pooled analysis of the PINNACLE studies. Respir Res. 2020;21:128.

8. D’Urzo A, Bader G, Shen S, Goyal P, Altman P. Comparison of glycopyrronium versus tiotropium on the time to clinically important deteriorations in patients with COPD: a post-hoc analysis of randomized trials. NPJ Prim Care Respir Med. 2018;28:18.

9. Donohue JF. Minimal clinically important differences in COPD lung function. COPD. 2005;2:111–24.

10. Westwood M, Bourbeau J, Jones PW, Gerulli A, Capkun-Niggli G, Worthy G. 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.

11. Cazzola M, MacNee W, Martinez FJ, Rabe KF, Francis LG, Barnes PJ, Brusasco V, Burge PS, Calverley PM, Celli BR, et al. Outcomes for COPD pharmacological trials: from lung function to biomarkers. Eur Respir J. 2008;31:416–69.

12. Jones PW. Estimation and application of the minimum clinically important difference in COPD. Lancet Respir Med. 2014;2:167–9.

13. Decramer M, Celli B, Tashkin DP, Pauwels RA, Burkhart D, Cassino C, Kesten S. Clinical trial design considerations in assessing long-term functional impacts of tiotropium in COPD: the UPLIFT trial. COPD. 2004;1:303–12.

14. 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;359:1543–54.

15. Global Initiative for Chronic Obstructive Lung Disease. Global strategy for the diagnosis, management and prevention of chronic obstructive pulmonary disease (2019 report). 2019. https://goldcopd.org/wp-content/uploads/2018/11/GOLD-2019-v1.7-FINAL-14Nov2018-WMS.pdf. Accessed 7 Nov 2019.

16. Celli BR, Decramer M, Liu D, Metzdorf N, Asijee GM, Tashkin DP. Defining a COPD composite safety endpoint for demonstrating efficacy in clinical trials: results from the randomized, placebo-controlled UPLIFT® trial. Respir Res. 2016;17:48.

17. Celli BR, Decramer M, Asijee GM, Kupas K, Tashkin DP. Effects of tiotropium on exacerbations in patients with COPD with low or high risk of exacerbations: a post-hoc analysis from the 4-year UPLIFT® trial. Chronic Obstr Pulm Dis. 2015;2:122–30.

18. Freemantle N, Calvert M, Wood J, Eastaugh J, Griffin C. Composite outcomes in randomized trials: greater precision but with greater uncertainty? JAMA. 2003;289:2554–9.

Publisher’s Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.