Chronic Obstructive Pulmonary Diseases:
Journal of the COPD Foundation

Original Research

Responder Analyses for Treatment Effects in COPD Using the St George's Respiratory Questionnaire

Paul W. Jones, PhD, FRCP, FERS[1,2] Heather Gelhorn, PhD[3] Hilary Wilson, PhD[3] Niklas Karlsson, PhD[4]
Shailendra Menjoge, PhD[5] Hana Müllerova, PhD[2] Stephen I. Rennard, MD[6,7] Ruth Tal-Singer, PhD[8]
Debora Merrill, MBA[9] Maggie Tabberer, MSc[2]

Abstract

Background: Patient-reported outcomes data in clinical trials are usually reported as mean values, interpreted in comparison to a minimum clinically important difference (MCID) and ignoring the possibility of a sizable proportion of patients experiencing a worthwhile benefit when the majority did not. This analysis tested the reliability of calculated responder rates (from chronic obstructive pulmonary disease [COPD] patients) with the St George's Respiratory Questionnaire (SGRQ) using a range of responder cut-points above and below the MCID (4 units).

Methods: Individual patient data (i.e., data from long-acting bronchodilator [LAB] and inhaled corticosteroids [ICS]/long-acting beta2-agonist [LABA] randomized clinical studies) in the COPD Biomarker Qualification Consortium database were used: short-term (≤1-year duration; 14,814 patients,) and medium-term (2-4 years; 12,043 patients). Responder rates versus placebo across SGRQ score change thresholds ranging from -1.5 to -8.0 were tested; differences were expressed as the odds ratio (OR) of a patient exceeding the threshold versus no change or deterioration.

Results: The ORs measuring benefit of active treatment were similar across thresholds in short-term studies (LAB, ORs 1.40-1.42; LABA/ICS, 1.50-1.56) and medium-term LAB studies (ORs 1.34-1.43), whereas ORs in medium-term studies with LABA/ICS intervention showed a trend for higher response rates at higher values of threshold cut-points (1.64-1.79). In short-term studies, different thresholds had little effect on the OR between active drugs versus a trend for lower ORs with lower thresholds in medium-term studies.

Conclusions: The OR for a treatment effect compared with placebo appears consistent across a range of responder cut-points. In medium-term trials, the treatment difference between active drugs suggests that use of a lower threshold would not increase the odds of observing a measured treatment difference.

1 Division of Clinical Science and Institute of Infection and Immunity, St George's University of London, United Kingdom

2 Research and Development, GlaxoSmithKline, Uxbridge, United Kingdom

3 Evidera, Bethesda, Maryland

4 Research and Development, AstraZeneca, Gothenburg, Sweden

5 Boehringer Ingelheim Pharmaceuticals, Inc., Ridgefield, Connecticut

6 Division of Pulmonary, Critical Care, Sleep and Allergy, Nebraska Medical Center, Omaha

7 Research and Development, AstraZeneca, Cambridge, United Kingdom

8 Research and Development, GlaxoSmithKline, Upper Merion, Pennsylvania

9 COPD Foundation, Washington, D.C.

Address correspondence to:

Paul W. Jones, PhD, FRCP, FERS
St George's University of London
London, United Kingdom
Email: pjones@sgul.ac.uk.
Phone: +44 (0)20 8990 9000

Introduction

Chronic obstructive pulmonary disease (COPD) is a complex, multi-component disease which imposes a significant burden on both patients and health care systems.[1] The measurement of decline in lung function, specifically forced expiratory volume in 1 second (FEV_1), has traditionally been the most frequently used parameter to assess the effectiveness of treatments in clinical trials.[2] However, the poor relationship between FEV_1 and other markers of COPD outcomes has highlighted the importance of including patient-reported outcomes (PROs) in clinical trials to measure the effects of treatment on components of the disease that impact patients' activities of daily living and quality of life.[1,3,4,5]

In the clinical trial setting, significant changes in PRO measures are usually reported as mean values, and the difference in mean values between treatments is then compared with reference to a previously derived value for the minimum clinically important difference (MCID).[6] If the average improvement is less than the MCID, this is often interpreted to mean that the treatment difference is not clinically significant. This conclusion ignores the fact that the difference between the means indicates a shift of the distribution of responses and that a sizable proportion of patients may have improved by more than the MCID. An alternative approach is to perform a responder analysis to examine the odds or probability of achieving a clinically significant improvement in each group.

The St George's Respiratory Questionnaire (SGRQ) is a self-administered health status questionnaire which has an empirically determined MCID of -4 units.[7] The objective of this analysis was to test the reliability of responder rates estimated with a range of threshold values for a SGRQ response, above and below the MCID, using pooled patient data from the COPD Biomarkers Qualification Consortium (CBQC) database.[8]

Methods

Pooled individual patient data for this analysis were obtained from 17 randomized controlled trials (RCTs) identified in the CBQC database, fulfilling the criteria of: (1) treatment comparison of placebo with either long-acting bronchodilators (LAB) (long-acting beta2-agonists [LABA] or long-acting anti-muscarinic agents) or LABAs + inhaled corticosteroid treatment (LABA+ICS), and (2) having the SGRQ as a study outcome. The RCTs comprised 14 short term studies of less than 1 year in duration (N=14,814, of which 13,131 were taking 1 of the 3 relevant treatments and were included in this analysis as evaluable patients) and 3 medium-term studies of 2-4 years' duration (N=12,043, providing 10,797 evaluable patients). Further details on the RCTs are available elsewhere.[9] The objectives and description of the CBQC initiative has been reported in detail previously.[8]

Outcome Measure and Analysis

The outcome of interest was the SGRQ total score.[7] The objective of this analysis was to evaluate responder rates across a range of SGRQ thresholds to determine the effects of choice of responder threshold on the size of treatment difference between placebo and active drug. An additional analysis explored the presence of *super-responders,* i.e., patients showing an improvement in SGRQ that exceeded the 4-unit MCID by 3 times.

All analyses were conducted using the short-term and medium-term databases, but were limited to analysis of outcomes at 6 months in the short-term studies since this is a common duration for such trials and in the medium-term trials the responder status was established at 1-year post-baseline, to minimize bias due to the progressive nature of changes in the SGRQ score in COPD,[10,11] and differential patient drop-out rates.[12]

Responder analyses were used to assess SGRQ change from baseline by treatment class, based on classification of response (responders/ non-responders) using thresholds from -1.5 to -8.0 and -12.0. Logistic regression models were used to examine the odds ratios (ORs) of response in the 2 treatment groups compared with placebo. Adjustment for the following covariates was included: age, income, World Health Organization region, sex, year of study start, COPD duration, body mass index, smoking status, and $FEV_1\%$ predicted. Patients with any missing data were omitted from the analyses.

Results

Patient Population
The demographic and clinical characteristics of patients with COPD included in the CBQC database[8] have been described in the baseline descriptive paper that has been published as part of this series.

Mean Changes in SGRQ Score
The mean changes in SGRQ score in the 3 treatment groups are shown in Table 1. In short term studies the mean differences from placebo at 6 months for LAB and LABA/ICS were -1.99 and -1.74 respectively. In the medium-term studies the mean differences from placebo at 12 months for LAB and LABA/ICS were -2.31 and -2.45 respectively.

Responder Rate
Cumulative frequency distributions for placebo and active treatment are shown in Figure 1. Responder rates

Table 1. Change in SGRQ Total Score Over Time by Treatment Class in Short-term Studies (at 6 months) and Medium-term Studies (at 12 Months)

Study Type	Treatment Class		
	Placebo	LAB only	LABA + ICS
Short-term Studies LS mean (SE)[a]	n=2108 -3.87 (0.30)	n=3722 -5.86 (0.22)	n=1752 -5.61 (0.32)
Medium-term Studies LS mean (SE)[a]	n=2194 -1.99 (0.24)	n=3396 -4.30 (0.22)	n=1018 -4.44 (0.41)

[a]Least squares (LS) mean difference and standard error (SE) determined from analysis of variance models which included treatment class.
SGRQ=St George's Respiratory Questionnaire; LS=least squares; SE=standard error; LAB=long-acting bronchodilator; LABA=long-acting beta2-agonist; ICS=inhaled corticosteroid

for placebo and active treatment (LAB and LABA/ICS) were lower with higher responder thresholds in both the short-term and medium-term studies (Figures 2A and 2B). In both short- and medium-term studies, ≈40% of placebo-treated patients achieved the MCID threshold of 4 points and ≈30% were defined as responders, even using a threshold of 8 units (twice the MCID) (Figure 2A and 2B).

Odds Ratios for Responder Rates: Active Treatment Versus Placebo
In the short-term studies, at 6 months, the OR for active treatment versus placebo remained almost constant across the range of threshold values tested (LAB alone 1.40-1.42; LABA/ICS 1.50-1.56) (Figure 3A). In the medium term trials, evaluated at 12 months, the OR measuring benefit with LAB alone remained almost constant (1.34-1.43), whereas with LAB/ICS there was a discernible trend for higher ORs at higher values of threshold cut-points in the range 1.64-1.79 (Figure 3B). The OR ratio for the treatment effect (active treatment versus placebo) was very similar, whether the MCID or half the MCID (2 units) was used as the threshold (Figure 3A and B).

Odds Ratios for Responder Rates: ICS/LABA Versus Active Treatment
When comparing 2 active drugs (i.e. ICS/LABA versus one or other monocomponents), the choice of threshold had no consistent effect on the OR in the short-term trials, (Figure 4A), but in the medium-term trials there was a clear trend for higher ORs with higher threshold values (Figure 4B), whether comparing ICS/LABA with ICS alone or LABA alone.

Super Responders
In both the short- and medium-term studies, approximately a quarter of patients on active treatment were *super-responders,* i.e., showed improvements in SGRQ total score of 2 or 3 times the 4 unit MCID (Table 2). In both short-and medium-term studies, there was a general trend with the OR for benefit versus placebo, whether with LAB alone or ICS/LABA, at least as large as using a 12-point cut-off as the MCID.

Discussion
The main finding from this analysis was that the OR for benefit of active drug over placebo or active drug versus another active drug was generally very consistent over

Figure 1. Cumulative Frequency Distribution for Short-term Trials and Medium-term Trials
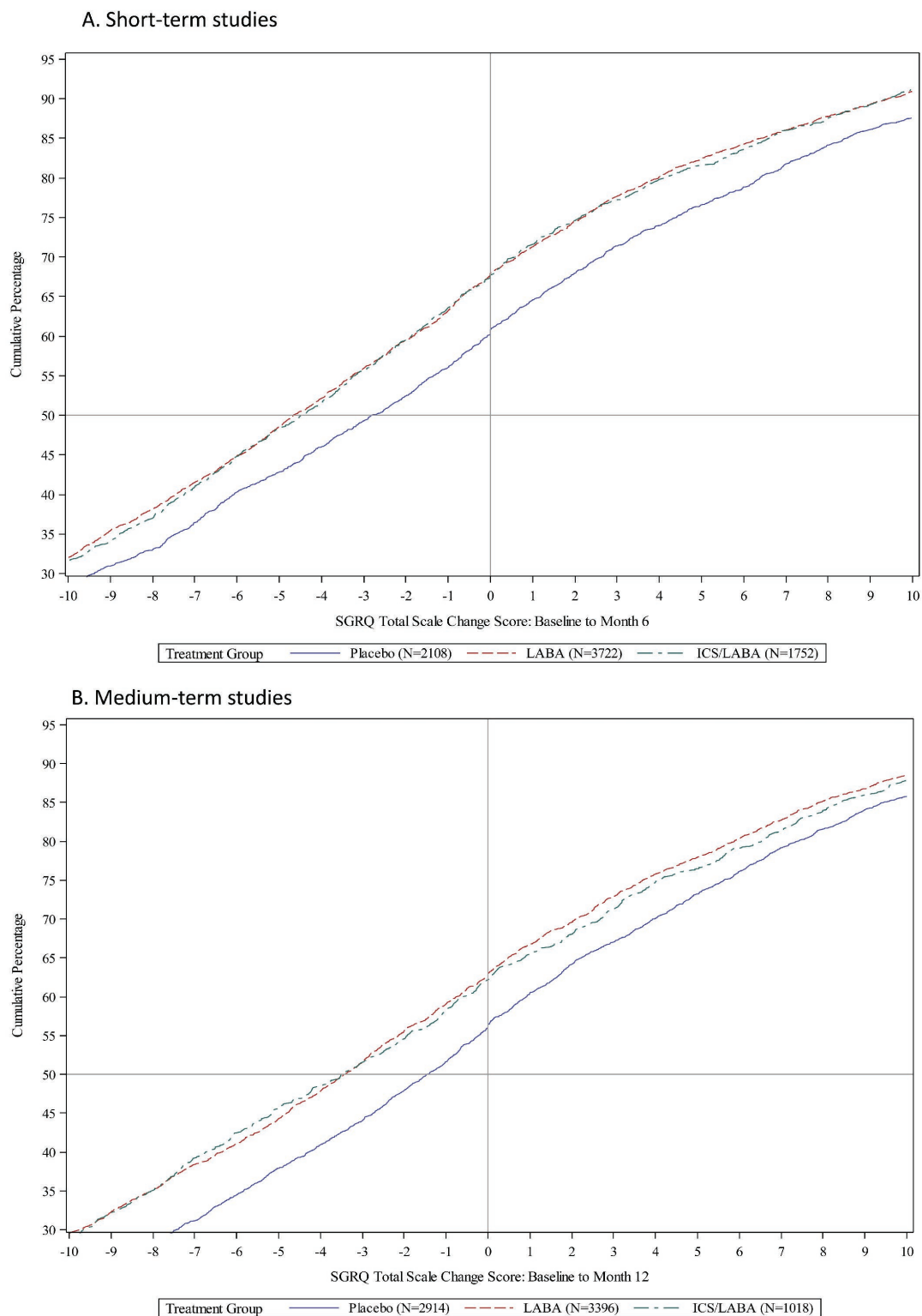
Panel A: Short-term trials; Panel B: Medium-term trials
Note that the tails of the distributions have been curtailed to amplify the central region of the curves

Figure 2. Responder Rates Calculated Using Different Thresholds for Response
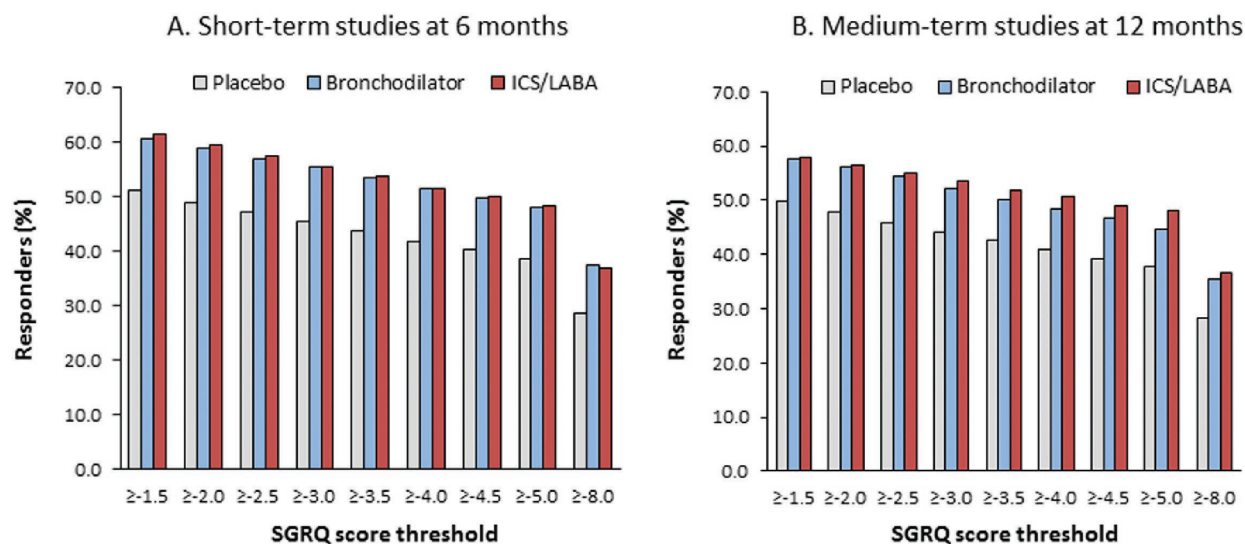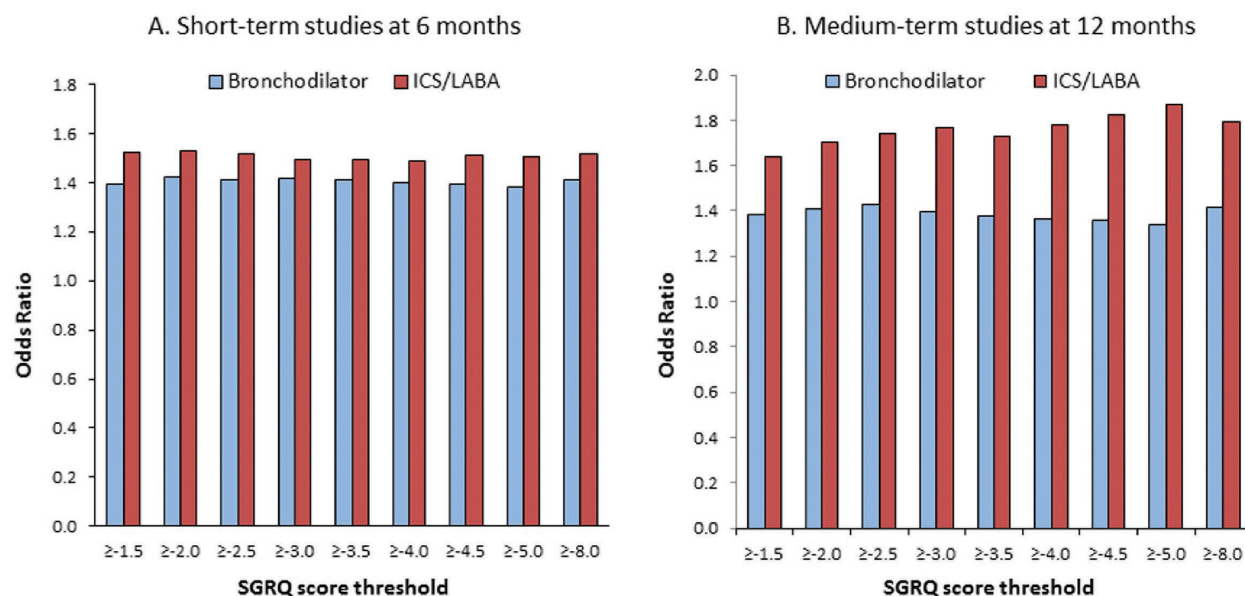

Figure 3. Odds Ratio for Responder Rates Versus Placebo: Calculated Using Different Thresholds for Response


a wide range of threshold values for response, from a little above the MCID (-5 units) to well below it (-1.5 units). The only exception to this general rule was in the medium-term trials where the OR for benefit, particularly for the active versus active comparisons was slightly lower (≈5%) with the threshold set at -1.5 compared to the MCID. Inspection of the cumulative distribution curves shows why the OR is so consistent – in the region of interest, the slope of the placebo and

active treatment curves are almost parallel.

These observations have 2 important implications for the analysis of SGRQ data; first, it shows that the precise value of the MCID, when used in a responder analysis will have little or no impact on the estimated treatment effect; second it shows that there is no reason to postulate that a separate MCID should be created for active versus active comparisons.

The reason for the apparently greater odds of benefit

Figure 4. Odds Ratio for Responder Rates: Inhaled Corticosteroids/ Long-acting Beta2-Agonists Versus Inhaled Corticosteroids Alone or Bronchodilator Alone[a]
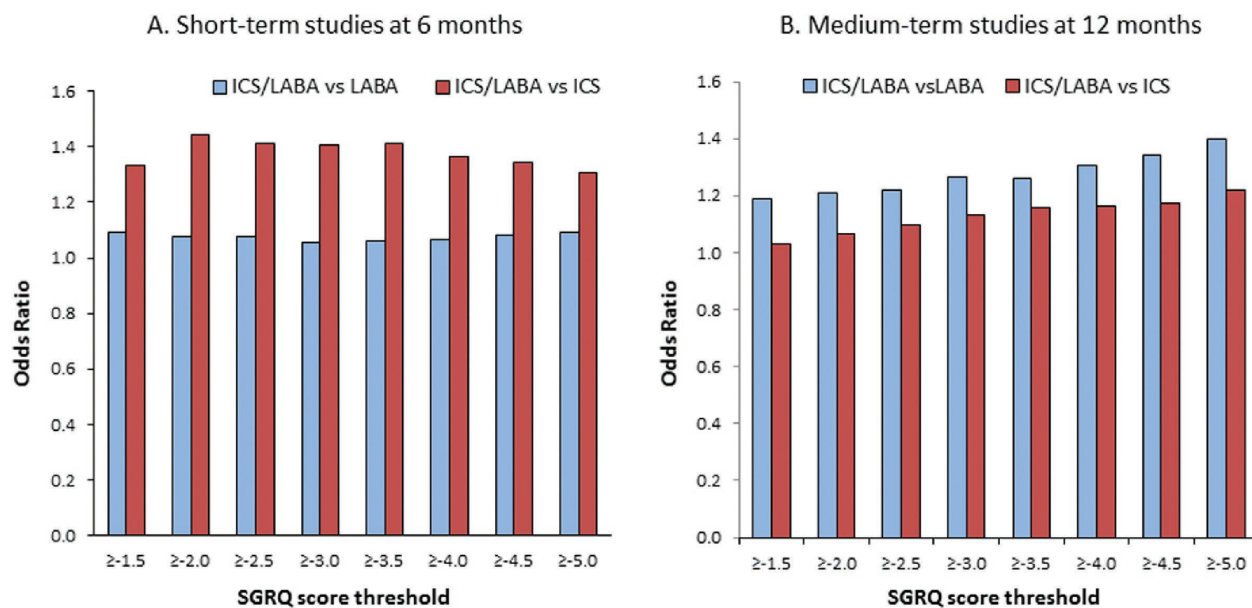
ᵃCalculated using different thresholds for response

Table 2. Responder Analysis of SGRQ Total Score Improvement from Baseline[a]

	Responder Threshold		
	SGRQ ≥4	SGRQ ≥8	SGRQ ≥12
Short-term			
LAB Alone vs Placebo,	1.40	1.41	1.35
OR 95% CI	1.23, 1.59	1.23, 1.62	1.16, 1.58
ICS/LABA vs Placebo,	1.49	1.52	1.51
OR 95% CI	1.28, 1.73	1.29, 1.78	1.27, 1.81
Medium-term			
LAB alone vs Placebo,	1.36	1.42	1.48
OR 95% CI	1.23, 1.51	1.27, 1.58	1.31, 1.68
ICS/LABA vs Placebo,	1.78	1.79	1.89
OR 95% CI	1.47, 2.15	1.46, 2.20	1.50, 2.38

ᵃExpressed as odds ratios and 95% CI with responder thresholds at ≥ 4 (MCID); ≥ 8 (2 x MCID) or ≥ 12 (3 x MCID) by treatment class in short- and medium-term studies.

OR=odds ratio; LAB=long-acting bronchodilator; CI=confidence interval; LABA=long-acting beta2-agonist; ICS=inhaled corticosteroid

for the *super-responders* may be largely mathematical; since the curves are almost parallel in this region, the difference between treatments will be the same, but the responder rate in the reference group (i.e., the denominator) will be lower, so the ratio will be greater. This mathematical point should not detract from the observation that within a population of responders,

there is a greater likelihood of some patients getting very large benefits from treatment rather than placebo.

This analysis also illustrates the value of performing responder analyses. The mean difference from placebo in the short-term trials with either treatment group was <2 units and with the medium-term trials it was <2.5 units, both of which are well below the 4 unit MCID. However, it would be incorrect to conclude that there was no clinically relevant benefit, since with both treatments and in both trials the OR compared with placebo was >1.30 (using the MCID as the cut point), i.e., a >30% greater odds of a clinically significant benefit with treatment compared with placebo.

The strength of this analysis is that it is based on a very large body of pooled individual patient data from several trials. In fact, due to the collaborative nature of the program, the database is made up of a majority of studies performed with these agents in COPD over this time-period, so the results should be generalizable to other trials in COPD recruiting patients with a similar demographic profile. The chief weakness is that these are not within-study comparisons, but the results show a broadly consistent pattern across studies of different duration and between different treatments and different active comparators, so the findings are likely to be reliable. It is important not to make conclusions about

relative treatment efficacy between treatment classes using these data, because they are indirect comparisons.

In conclusion, this analysis has shown that a responder analysis provides very similar estimates of treatment efficacy over a wide range of cut points. This applies whether the comparator is a placebo or an active drug. Use of responder analysis removes the need for any discussion about using different MCIDs for active versus placebo comparisons. Responder analysis does not obviate the use of mean differences in response rate, but it does provide a reliable and clinical method of reporting treatment effects on SGRQ scores that are meaningful in routine clinical practice.

Declaration of Interest
PWJ, NK, SM, HM, SIR, RTS and MT are employees of the pharma companies who funded this analysis. KR is a former employee of Pfizer. HG and HW participated in this project as employees of Evidera, a company which performs work for hire for multiple pharmaceutical and device companies in outcomes research. DM has nothing to declare.

References

1. Global Initiative for Chronic Obstructive Lung Disease (GOLD). Global Strategy for the Diagnosis, Management and Prevention of COPD 2016. GOLD website. http://www.goldcopd.org/. Published 2016. Accessed February 2017.

2. Jones PW, Agusti AGN. Outcomes and markers in the assessment of chronic obstructive pulmonary disease. *Eur Respir J*. 2006; 27: 822–832.
doi: https://doi.org/10.1183/09031936.06.00145104

3. Jones P, Miravitlles M, van der Molen T, Kulich K. Beyond FEV_1 in COPD: a review of patient-reported outcomes and their measurement. *Int J Chron Obstruct Pulmon Dis*. 2012; 7: 697–709. doi: https://doi.org/10.2147/COPD.S32675

4. Cazzola M, Hanania NA, MacNee W, Rüdell K, Hackford C, Tamimi N. A review of the most common patient-reported outcomes in COPD – revisiting current knowledge and estimating future challenges. *Int J Chron Obstruct Pulmon Dis*. 2015; 10(1) 725–738. doi: https://doi.org/10.2147/COPD.S77368

5. Food and Drug Administration (FDA). Chronic obstructive pulmonary disease: Developing drugs for treatment. Guidance for industry. FDA website. http://www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm071575.pdf . Published May 2016. Accessed February 2017.

6. Jones PW, Beeh KM, Chapman KR, Decramer M, Mahler DA, Wedzicha JA. Minimal clinically important differences in pharmacological trials. *Am J Respir Crit Care Med*. 2014;189 (3):250–255. doi: https://doi.org/10.1164/rccm.201310-1863PP

7. Jones PW. St. George's Respiratory Questionnaire: MCID. *COPD*. 2005; 2(1):75-79.
doi: https://doi.org/10.1081/COPD-200050513

8. Casaburi R, Celli B, Crapo J, et al. The COPD Biomarker Qualification Consortium (CBQC). *COPD*. 2013;10(3):367-377.
doi: https://doi.org/10.3109/15412555.2012.752807

9. Tabberer M, Benson VS, Gelhorn H, et al, for the COPD Biomarkers Qualification Consortium. The COPD Biomarkers Qualification Consortium Database: Baseline characteristics of the St George's Respiratory Questionnaire dataset. *Chronic Obstr Pulm Dis*. 2017; 4(2): 112-123.
doi: http://doi.org/10.15326/jcopdf.4.2.2017.0128

10. Calverley PM, Anderson JA, Celli B, et al and the TORCH investigators. Salmeterol and fluticasone propionate and survival in chronic obstructive pulmonary disease. *N Engl J Med*. 2007; 356:775–789. doi: https://doi.org/10.1056/NEJMoa063070

11. Tashkin DP, Celli B, Senn S, et al and the UPLIFT Study Investigators: A 4-year trial of tiotropium in chronic obstructive pulmonary disease. *N Engl J Med*. 2008; 359:1543–1544.
doi: https://doi.org/10.1056/NEJMoa0805800

12. Vestbo J, Anderson JA, Calverley PM, et al. Bias due to withdrawal in long-term randomised trials in COPD: Evidence from the TORCH study. *Clin Respir J*. 2011; 5(1):44-49.
doi: https://doi.org/10.1111/j.1752-699X.2010.00198.x