Minimal Clinically Important Differences in Pharmacological Trials

Paul W. Jones[1], Kai M. Beeh[2], Kenneth R. Chapman[3], Marc Decramer[4], Donald A. Mahler[5], and Jadwiga A. Wedzicha[6]

[1]Division of Clinical Science, St George's University of London, London, United Kingdom; [2]insaf Respiratory Research Institute, Wiesbaden, Germany; [3]Asthma and Airway Centre, University Health Network, Toronto, Ontario, Canada; [4]Respiratory Division, University Hospital, University of Leuven, Leuven, Belgium; [5]Section of Pulmonary and Critical Care Medicine, Geisel School of Medicine at Dartmouth, Hanover, New Hampshire; and [6]Department of Academic Respiratory Medicine, University College London, Royal Free Campus, London, United Kingdom

Abstract

The concept of a minimal clinically important difference (MCID) is well established. Here, we review the evidence base and methods used to define MCIDs as well as their strengths and limitations. Most MCIDs in chronic obstructive pulmonary disease (COPD) are empirically derived estimates applying to populations of patients. Validated MCIDs are available for many commonly used outcomes in COPD, including lung function (100 ml for trough FEV_1), dyspnea (improvement of $\geqslant$ 1 unit in the Transition Dyspnea Index total score or 5 units in the University of California, San Diego Shortness of Breath Questionnaire), health status (reduction of 4 units in the St George's Respiratory Questionnaire total score), and exercise capacity (47.5 m for the incremental shuttle walking test, 45–85 s for the endurance shuttle walking test, and 46–105 s for constant-load cycling endurance tests), but there is currently no validated MCID for exacerbations. In a clinical trial setting, many factors, including study duration, withdrawal rate, baseline severity, and Hawthorne effects, can influence the measured treatment effect and determine whether it reaches the MCID. We also address recent challenges presented by clinical trials that compare active treatments and suggest that MCIDs should be used to identify the additional proportion of patients who benefit, for example, when one drug is replaced by another or when a second drug is added to a first. We propose the term "minimum worthwhile incremental advantage" to describe this parameter.

Keywords: chronic bronchitis; emphysema; chronic obstructive pulmonary disease; minimal clinically important difference; outcomes assessment

A minimal clinically important difference (MCID) provides a guide as to whether an intervention provides a minimum level of perceived benefit and moves beyond the concept of statistical differences. The term was first described in 1989 (1) and is now a well-established concept, viewed by healthcare professionals and regulatory bodies as a reliable method for evaluating an intervention.

Chronic obstructive pulmonary disease (COPD) exemplifies some of the challenges associated with calculating and using MCIDs (2). Here, we review the evidence base and methods used to define MCIDs for the major outcomes in COPD. In response to the advent of multiple-drug therapies for COPD, we propose the introduction of a concept, based around the MCID, termed the "minimum worthwhile incremental advantage," to describe the benefit of one active treatment regimen over another.

MCID Determination

MCIDs are usually derived by one of two approaches: distribution- and anchor-based (3, 4). Distribution-based methods use the frequency distribution of observed events, and 0.5 SD and 1 SEM have been suggested as MCIDs (5). These methods provide a value that indicates a statistical difference, but this may not be perceivable by a patient or clinician. In contrast, anchor-based methods examine the relationship between scores on the assessment instrument and other measures of impaired health (the anchors) (3, 4).

Estimation of an MCID for an instrument should be based on multiple approaches, with the resulting values narrowed down to a single value or small range, based on a process of triangulation (4). When empirically derived triangulated values are not possible, a modified Delphi model has been used to identify a consensus value. The latter illustrates

(Received in original form October 21, 2013; accepted in final form December 23, 2013)

Medical writing support for this article was funded by Novartis Pharma AG, Basel, Switzerland.

Author Contributions: All authors were involved in drafting, reviewing, and critically revising the manuscript and approved the final version of the text.

Correspondence and requests for reprints should be addressed to Paul W. Jones, M.D., Ph.D., Division of Clinical Science, St George's University of London, Cranmer Terrace, London SW17 0RE, UK. E-mail: pjones@sgul.ac.uk

Am J Respir Crit Care Med Vol 189, Iss 3, pp 250–255, Feb 1, 2014

Originally Published in Press as DOI: 10.1164/rccm.201310-1863PP on January 2, 2014
Internet address: www.atsjournals.org

an important underlying component of all MCID estimation—at a key point in its development a value judgment will have been made, even if the data were collected empirically. There is no absolute way of determining an MCID. It is important to recognize that value judgments are not verifiable against external objective standards, only against judgments made by others. It should also be appreciated that MCIDs are average estimates obtained in groups of patients. For an individual patient, a worthwhile perceived benefit may occur below the mean estimated MCID. Until recently, most experience with MCIDs has been in the context of placebo-controlled trials, in which the measured treatment effects may be large, but for trials comparing active treatments, the differences between treatments may be smaller. This does not mean that an MCID should be recalculated for this setting, but it does require a recalibration of the expected size of benefit.

MCIDs in COPD

The following sections discuss how COPD-related MCIDs have been defined and validated, and the values of these MCIDs are summarized in Table 1.

Lung Function (FEV_1)

Improving lung function is not an objective of COPD management (6, 7), but it is the primary endpoint most frequently used by regulatory authorities in interpreting drug efficacy in COPD trials (8).

Opinions on what constitutes an MCID for FEV_1 vary. The American Thoracic Society/European Respiratory Society task force has defined a range of 100 to 140 ml (8). Regulators consider a change of 5 to 10% from baseline as clinically important and a change of less than 3% from baseline as not clinically important (8). An MCID of 100 ml for predose or trough FEV_1 has been proposed, based on clinical anchoring to endpoints such as exacerbations, perception of dyspnea, and decline in lung function, but not survival (9).

Although trough FEV_1 measurements are reproducible, there are issues with repeatability (noise effects can be $>$ 100 ml) (9). Additionally, baseline lung function can affect the potential for improvement, so relative change rather than absolute change may be more meaningful in patients with worse airflow limitation (2).

One of the major problems with determining the MCID for FEV_1 is the poor correlation between it and appropriate anchors. Perhaps the best that can be achieved is recognition that in studies where the treatment achieves the current FEV_1 consensus MCID, other endpoints achieve their respective MCIDs.

Exacerbations

A reduction in frequency of 20% has been suggested as a reasonable MCID for exacerbations, calculated by anchoring

Table 1: Minimal Clinically Important Differences for Commonly Used Outcomes in Chronic Obstructive Pulmonary Disease

Endpoint	MCID (Improvement)	Method of Estimation	Reference
Lung function			
Trough FEV_1	100 ml	Anchor-based (exacerbations, patient perception, 2-yr decline in lung function)	9
Exacerbations	No validated MCID	—	—
Dyspnea			
TDI total score	1 unit	Anchor-based (physician's global evaluation score), distribution-based (SEM, 0.5 SD), expert preference	19
UCSD SOBQ	5 units	Anchor-based (CRQ dyspnea domain, TDI), distribution-based (SEM, Cohen's effect size), estimate by experienced users	20
Health status			
SGRQ total score	4 units	Anchor-based (MRC dyspnea grade, CRQ dyspnea domain, mortality rate), expert and patient preference	23
CRQ domain scores	0.5 units (average)*	Anchor-based (patient perspectives), distribution-based (SEM, Cohen's effect size), expert panel-based	24
Exercise capacity			
6-min walk distance	26 $\pm$ 2 m (patients with severe COPD)	Anchor-based (SGRQ, UCSD SOBQ), distribution-based (SEM, Cohen's effect size, empirical rule effect size)	31
Incremental shuttle walking test	47.5 m	Anchor-based (patient perception)	32
Endurance shuttle walking test	45–85 s	Anchor-based (patient perception), distribution-based (0.5 SD)	33
Constant-load cycling endurance tests	46–105 s	Distribution-based (0.5 SD)	8
Dyspnea during exercise tests			
Modified Borg scale	1 unit	Distribution-based (Cohen's effect size)	39
Visual analog scale	10–20 units	Distribution-based (Cohen's effect size)	39

Definition of abbreviations: COPD = chronic obstructive pulmonary disease; CRQ = Chronic Respiratory Questionnaire; MCID = minimal clinically important difference; MRC = Medical Research Council; SGRQ = St George's Respiratory Questionnaire; TDI = Transition Dyspnea Index; UCSD SOBQ = University of California, San Diego Shortness of Breath Questionnaire.
*The MCIDs for the individual domains differ around this mean estimate.

exacerbation rates to the St George's Respiratory Questionnaire (SGRQ) (10). Even with this 20% value, there appears to be a large range in what is considered an important change. Rates between 4.4 and 42.0%, for example, have been associated with meaningful changes in questionnaire-based instruments (11), and if the studies that have influenced the 2011 Global Initiative for Chronic Obstructive Lung Disease (GOLD) guidelines are considered, then statistically significant differences in exacerbation rates between 9 and 53.5% indicate meaningful clinical benefit (12).

The development of an MCID for exacerbations is complicated by the lack of a uniform definition for exacerbations and severity grading as well as underreporting (12, 13). Moreover, the distribution of exacerbation rates is skewed, with seasonal variation (8) and substantial inter- and intrapatient variability in frequency (10, 14). Thus, there is presently no validated MCID for exacerbations; indeed, as the wide range of possible MCIDs indicates, this might not be possible using clinical anchor-based approaches. Perhaps more patient-centered methods might be used, for example using discrete choice modeling techniques in patients.

Dyspnea

Dyspnea can be measured by different instruments: the Transition Dyspnea Index (TDI) (15), the modified Medical Research Council (mMRC) scale (16), and the University of California, San Diego Shortness of Breath Questionnaire (17).

The mMRC scale is used widely as a discriminative instrument but has poor evaluative properties to assess changes in dyspnea (18). The TDI is widely used to measure treatment effects; an improvement of 1 or more units in the total score has been defined as the MCID (19). An improvement of 5 units is suggested as a reasonable MCID for the University of California, San Diego Shortness of Breath Questionnaire (20).

Health Status Measurements

In COPD trials, the SGRQ (21) and Chronic Respiratory Questionnaire (CRQ) (22) are widely used. The MCID for the SGRQ is a reduction of 4 units in the total score, estimated using triangulation (23). The average MCID for the CRQ domain scores is 0.5 (24), although the MCIDs for the individual domains differ around this mean estimate (25). The MCID has never, to our knowledge, formally been defined for a total CRQ score.

With older therapies, such as salmeterol/fluticasone or tiotropium, average treatment effects with SGRQ compared with placebo have typically been at or below the MCID (26, 27). More recently, agents such as indacaterol (28, 29) and aclidinium (30) have shown mean SGRQ improvements at or above the MCID, compared with placebo.

Exercise Capacity

The 6-minute walk distance and incremental shuttle walking test are commonly used in COPD (31, 32). A recent study identified an MCID of 26 $\pm$ 2 m for the 6-minute walk distance in patients with severe COPD using distribution- and anchor-based estimates (31). The MCID for the incremental shuttle walking test is 47.5 m (32).

Constant work-rate exercise is another measure used to assess exercise capacity in COPD. For the endurance shuttle walking test, an MCID of 45 to 85 seconds has been proposed (33), whereas the American Thoracic Society/European Respiratory Society task force states that for constant-load cycling endurance tests, an improvement of 46 to 105 seconds can be considered clinically important (8). Improvements in endurance time achieved with bronchodilators show large variation (34–36), which may be attributable to variations in patient characteristics and phenotypes, exercise protocols, and study duration (37).This complicates the estimation of a generally acceptable MCID for endurance.

Dyspnea may be measured during exercise tests using the modified Borg scale and the visual analog scale (VAS) (38). A difference of 1 or more units has been proposed as the MCID for the Borg scale, whereas an improvement of 10 to 20 units on the VAS was associated with moderate symptom improvement (39).

Factors Affecting Attainment of MCIDs in COPD

In clinical trials, several factors may influence the treatment effect and determine whether it reaches the MCID, but these are factors that determine measured efficacy; they do not alter the threshold change at which a treatment may be judged to be clinically beneficial.

Study Duration

A minimum study duration is important and will vary by outcome. Studies of 6 to 12 months' duration appear to provide the optimal length to measure patient-reported outcomes. The largest SGRQ effect, compared with placebo, occurs at around 6 months with long-acting bronchodilators (30, 40), although using long-acting β_2-agonist plus inhaled corticosteroid combinations, 2 months appeared to be sufficient (41, 42). A duration of 2 to 3 months may be sufficient for assessment of the TDI (30, 40).

Long-Term Trials

Studies longer than 1 year raise complications due to the progressive nature of COPD and the presence of differential dropout. For example, although treatments may produce an initial improvement in SGRQ, scores may subsequently return to the baseline level (or worse) due to disease progression (26, 27). Healthy survivor effects are also important; in the 3-year TOwards a Revolution in COPD Health (TORCH) trial, early study withdrawal was associated with worse baseline lung function and health status, and more frequent exacerbations regardless of treatment allocation (43). Because there are usually more patient withdrawals in the less effective treatment arm, this may lead to a biased estimate of effectiveness.

Baseline Disease Severity

There is little evidence that MCID attainment is dependent on disease severity. However, treatment effects may vary by baseline disease severity. In a *post hoc* analysis of the TORCH dataset, the greatest improvement in SGRQ score in patients treated with salmeterol/fluticasone versus placebo was seen in those with the most severe GOLD grade at baseline (44). Similarly, a high dose of indacaterol (300 μg) was seen to have greater benefit in patients with an mMRC grade greater than or equal to 2 at baseline (18).

In theory, baseline severity may affect the attainment of MCID due to floor and ceiling effects (45); however, in most COPD studies baseline values of the most widely used measurements lie in the middle of the scaling range, so this is unlikely to be a major confounding issue.

Therapeutic Specificity of MCIDs

MCIDs were not developed in the context of specific treatments—the reference point (at least with anchor-based methods) is patient- or clinician-perceived benefit. For that reason, the same MCID should apply, regardless of therapeutic modality. For example, a dyspnea MCID should be the same, whether benefit is achieved through a bronchodilator or through pulmonary rehabilitation. For health status measures it is potentially more complicated, because these instruments measure a wider range of aspects of the disease and a bronchodilator may produce benefit through a different route than an antiinflammatory agent. However, the principle should still apply if the instrument has been developed properly so that undue weight is not given to one area of the disease over another.

Hawthorne Effects

Hawthorne effects—changes in behavior due to observation of a participant—may occur on entry to a clinical trial. For example, improved compliance or better inhaler technique with concomitant treatments may improve symptoms in patients receiving placebo. A systematic review of randomized controlled trials of inhaled bronchodilators in patients with COPD suggests that a Hawthorne effect influences SGRQ scores in COPD trials (46). Typically this results in an improvement of 2 to 3 points on the SGRQ with placebo.

A few recent studies have reported changes in placebo-treated patients that exceed the MCID (26, 27). The reason for this is not fully understood but may be related to the socioeconomic status of the country. This in turn will require a large treatment effect for the therapy to be judged worthwhile, a problem that will be exacerbated if there is a sizeable difference in dropout rates (47).

Using MCIDs to Assess Clinical Efficacy

Because an MCID is an average estimate, it should be used as an indicative value rather than an absolute cut-off point between benefit and no benefit (23).

It is useful to consider the approach taken on this issue in other therapeutic areas. For example, an empirically validated noninferiority margin, using measurements from a VAS, has been established to compare treatment effects on endometriosis-associated pelvic pain (48). In rheumatology, the proportion of patients reaching MCIDs across a range of patient-reported outcomes has been used to compare active treatments (49).

One suggested approach for COPD has been to use the threshold for clinical significance and the 95% confidence intervals around the mean treatment effect to categorize the size of treatment response. This analysis produces one of five effects, ranging from "no effect" to a "large clinically significant effect" (Figure 1) (25). However, this method still uses the MCID as a critical hurdle. In a chronic disease this is a major challenge, because it requires at least half of the patients to improve by the MCID. Arguably, this is unrealistic in a disease such as COPD.

An alternative approach to using means is to use a responder analysis to identify the proportion of patients who improve by more than the MCID. This is attractive because responder rates may remain relatively stable across a range of thresholds (25, 50). Of course, this approach raises the question: what is a worthwhile responder rate? That is, again, a value judgment. Taking the TDI as an example, in recent studies, the responder rate for bronchodilator monotherapy compared with placebo was approximately 10% (30, 51). Using dual bronchodilator therapy, higher responder rates versus placebo have been reported, but the responder rate was still below 20% (52). This illustrates the challenges associated with modern COPD trials that compare combination therapies with monotherapy. It seems unrealistic to expect the incremental gain from adding a second active agent on top of a first to be as great as the difference between monotherapy and placebo and, as a consequence, to expect the additional drug to produce an improvement that exceeds the MCID. In this context, use of responder analyses may shift the debate from a possibly misdirected attempt to redefine the MCID to a lower value for use in active comparator studies toward forming a consensus view about the additional proportion of patients who benefit and what constitutes a worthwhile incremental gain. We propose the term "minimum worthwhile incremental advantage" to describe the percentage of patients who would experience improvement at or above the MCID on adding one treatment to another, or comparing two active treatments.

Future Directions

There has been little new methodology in MCID estimation over the last decade, and

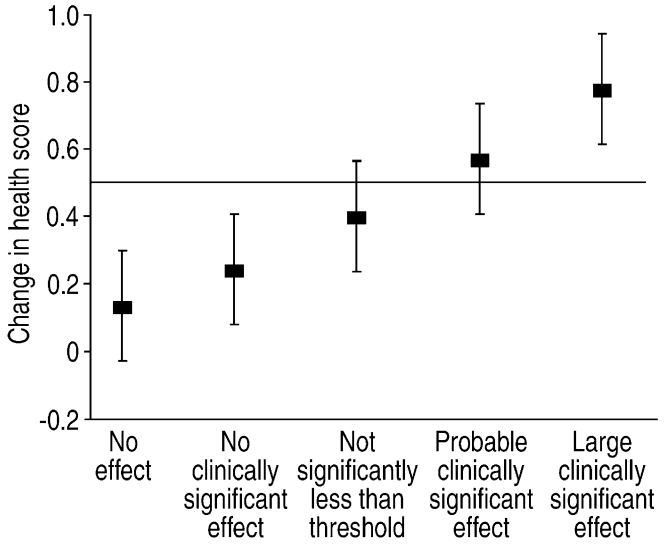

Figure 1. Method of categorizing clinical trial results using the threshold for clinical significance and the confidence intervals around the mean treatment effect. The *solid line* represents the threshold for a minimum clinically significant effect. Reproduced with permission of the European Respiratory Society (*Eur Respir J* March 2002 19:398–404; doi:10.1183/09031936.02.00063702; Reference 25).

few methods are truly patient-centered, but as suggested for exacerbations, new methodologies could be developed. These may also be able to address the question of whether an MCID might vary by severity. A more complex area is linked to the increasing perception that COPD is part of a multisystem disease process. However, the first challenge is the creation of instruments that measure the totality of the disease effect. In theory this is addressed by the use of generic health status instruments, but, although MCIDs are available for these, such instruments are known to be poorly responsive to COPD-specific interventions.

Conclusions

In COPD, validated MCIDs exist for a variety of outcomes, including lung function, dyspnea, health status, and exercise capacity, but there is as yet no validated MCID for exacerbations. Factors such as trial duration, Hawthorne effects, withdrawal rates, and baseline disease severity may affect the size of benefit relative to the MCID in clinical trials, and active comparator trials present additional challenges. MCIDs should be interpreted as indicative, rather than absolute. We suggest that responder analysis is an appropriate method to assess the minimum worthwhile incremental advantage between active treatments. ■

Author disclosures are available with the text of this article at www.atsjournals.org.

Acknowledgment: The authors wrote the manuscript and were assisted in its preparation by Molly Heitz and Maggie Davis, professional medical writers contracted to CircleScience (Macclesfield, UK). The authors thank Donald Banerji (Novartis Pharmaceuticals Corporation, East Hanover, NJ) for reviewing the manuscript.

References

1. Jaeschke R, Singer J, Guyatt GH. Measurement of health status: ascertaining the minimal clinically important difference. *Control Clin Trials* 1989;10:407–415.
2. Kiley JP, Sri Ram J, Croxton TL, Weinmann GG. Challenges associated with estimating minimal clinically important differences in COPD-the NHLBI perspective. *COPD* 2005;2:43–46.
3. Guyatt GH, Osoba D, Wu AW, Wyrwich KW, Norman GR; Clinical Significance Consensus Meeting Group. Methods to explain the clinical significance of health status measures. *Mayo Clin Proc* 2002; 77:371–383.
4. Revicki D, Hays RD, Cella D, Sloan J. Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. *J Clin Epidemiol* 2008;61:102–109.
5. Norman GR, Sloan JA, Wyrwich KW. Interpretation of changes in health-related quality of life: the remarkable universality of half a standard deviation. *Med Care* 2003;41:582–592.
6. Global Initiative for Chronic Obstructive Lung Disease. Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease; 2013 [updated 2013; accessed 14 September 2013]. Available from: http://www.goldcopd.org/uploads/users/files/GOLD_Report_2013_Feb20.pdf
7. American Thoracic Society and European Respiratory Society. Standards for the diagnosis and management of patients with COPD. [2004; accessed 15 October 2013]. Available from: http://www.thoracic.org/clinical/copd-guidelines/resources/copddoc.pdf
8. Cazzola M, MacNee W, Martinez FJ, Rabe KF, Franciosi LG, Barnes PJ, Brusasco V, Burge PS, Calverley PM, Celli BR, *et al*.; American Thoracic Society; European Respiratory Society Task Force on outcomes of COPD. Outcomes for COPD pharmacological trials: from lung function to biomarkers. *Eur Respir J* 2008;31:416–469.
9. Donohue JF. Minimal clinically important differences in COPD lung function. *COPD* 2005;2:111–124.
10. Calverley PM. Minimal clinically important difference—exacerbations of COPD. *COPD* 2005;2:143–148.
11. Anzueto A, Leimer I, Kesten S. Impact of frequency of COPD exacerbations on pulmonary function, health status and clinical outcomes. *Int J Chron Obstruct Pulmon Dis* 2009;4:245–251.
12. Chapman KR, Bergeron C, Bhutani M, Bourbeau J, Grossman RF, Hernandez P, McIvor RA, Mayers I. Do we know the minimal clinically important difference (MCID) for COPD exacerbations? *COPD* 2013; 10:243–249.
13. Seemungal TA, Donaldson GC, Paul EA, Bestall JC, Jeffries DJ, Wedzicha JA. Effect of exacerbation on quality of life in patients with chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 1998;157:1418–1422.
14. Donaldson GC, Goldring JJ, Wedzicha JA. Influence of season on exacerbation characteristics in patients with COPD. *Chest* 2012;141: 94–100.
15. Mahler DA, Weinberg DH, Wells CK, Feinstein AR. The measurement of dyspnea. Contents, interobserver agreement, and physiologic correlates of two new clinical indexes. *Chest* 1984;85:751–758.
16. Bestall JC, Paul EA, Garrod R, Garnham R, Jones PW, Wedzicha JA. 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.
17. Eakin EG, Resnikoff PM, Prewitt LM, Ries AL, Kaplan RM. Validation of a new dyspnea measure: the UCSD Shortness of Breath Questionnaire. University of California, San Diego. *Chest* 1998;113:619–624.
18. Mahler DA, Buhl R, Lawrence D, McBryan D. Efficacy and safety of indacaterol and tiotropium in COPD patients according to dyspnoea severity. *Pulm Pharmacol Ther* 2013;26:348–355.
19. Mahler DA, Witek TJ Jr. The MCID of the transition dyspnea index is a total score of one unit. *COPD* 2005;2:99–103.
20. Kupferberg DH, Kaplan RM, Slymen DJ, Ries AL. Minimal clinically important difference for the UCSD Shortness of Breath Questionnaire. *J Cardiopulm Rehabil* 2005;25:370–377.
21. Jones PW, Quirk FH, Baveystock CM, Littlejohns P. 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.
22. Guyatt GH, Berman LB, Townsend M, Pugsley SO, Chambers LW. A measure of quality of life for clinical trials in chronic lung disease. *Thorax* 1987;42:773–778.
23. Jones PW. St. George's Respiratory Questionnaire: MCID. *COPD* 2005; 2:75–79.
24. Schünemann HJ, Puhan M, Goldstein R, Jaeschke R, Guyatt GH. Measurement properties and interpretability of the Chronic Respiratory Disease Questionnaire (CRQ). *COPD* 2005;2:81–89.
25. Jones PW. Interpreting thresholds for a clinically significant change in health status in asthma and COPD. *Eur Respir J* 2002;19:398–404.
26. Calverley PM, Anderson JA, Celli B, Ferguson GT, Jenkins C, Jones PW, Yates JC, Vestbo J; TORCH investigators. Salmeterol and fluticasone propionate and survival in chronic obstructive pulmonary disease. *N Engl J Med* 2007;356:775–789.
27. Tashkin DP, Celli B, Senn S, Burkhart D, Kesten S, Menjoge S, Decramer M; UPLIFT Study Investigators. A 4-year trial of tiotropium in chronic obstructive pulmonary disease. *N Engl J Med* 2008;359:1543–1554.
28. Buhl R, Dunn LJ, Disdier C, Lassen C, Amos C, Henley M, Kramer B; INTENSITY study investigators. Blinded 12-week comparison of once-daily indacaterol and tiotropium in COPD. *Eur Respir J* 2011; 38:797–803.
29. Dahl R, Chung KF, Buhl R, Magnussen H, Nonikov V, Jack D, Bleasdale P, Owen R, Higgins M, Kramer B; INVOLVE (INdacaterol: Value in COPD: Longer Term Validation of Efficacy and Safety) Study Investigators. Efficacy of a new once-daily long-acting inhaled β_2-agonist indacaterol versus twice-daily formoterol in COPD. *Thorax* 2010;65:473–479.
30. Jones PW, Singh D, Bateman ED, Agusti A, Lamarca R, de Miquel G, Segarra R, Caracta C, Garcia Gil E. Efficacy and safety of twice-daily aclidinium bromide in COPD patients: the ATTAIN study. *Eur Respir J* 2012;40:830–836.

31. Puhan MA, Chandra D, Mosenifar Z, Ries A, Make B, Hansel NN, Wise RA, Sciurba F; National Emphysema Treatment Trial (NETT) Research Group. The minimal important difference of exercise tests in severe COPD. *Eur Respir J* 2011;37:784–790.
32. Singh SJ, Jones PW, Evans R, Morgan MD. Minimum clinically important improvement for the incremental shuttle walking test. *Thorax* 2008;63:775–777.
33. Pepin V, Laviolette L, Brouillard C, Sewell L, Singh SJ, Revill SM, Lacasse Y, Maltais F. Significance of changes in endurance shuttle walking performance. *Thorax* 2011;66:115–120.
34. Man WD, Mustfa N, Nikoletou D, Kaul S, Hart N, Rafferty GF, Donaldson N, Polkey MI, Moxham J. Effect of salmeterol on respiratory muscle activity during exercise in poorly reversible COPD. *Thorax* 2004;59:471–476.
35. Bédard ME, Brouillard C, Pepin V, Provencher S, Milot J, Lacasse Y, Leblanc P, Maltais F. Tiotropium improves walking endurance in COPD. *Eur Respir J* 2012;39:265–271.
36. O'Donnell DE, Casaburi R, Vincken W, Puente-Maestu L, Swales J, Lawrence D, Kramer B; INABLE 1 study group. Effect of indacaterol on exercise endurance and lung hyperinflation in COPD. *Respir Med* 2011;105:1030–1036.
37. Thomas M, Decramer M, O'Donnell DE. No room to breathe: the importance of lung hyperinflation in COPD. *Prim Care Respir J* 2013; 22:101–111.
38. Hareendran A, Leidy NK, Monz BU, Winnette R, Becker K, Mahler DA. Proposing a standardized method for evaluating patient report of the intensity of dyspnea during exercise testing in COPD. *Int J Chron Obstruct Pulmon Dis* 2012;7:345–355.
39. Ries AL. Minimally clinically important difference for the UCSD Shortness of Breath Questionnaire, Borg Scale, and Visual Analog Scale. *COPD* 2005;2:105–110.
40. Vincken W, van Noord JA, Greefhorst AP, Bantje TA, Kesten S, Korducki L, Cornelissen PJ; Dutch/Belgian Tiotropium Study Group. Improved health outcomes in patients with COPD during 1 yr's treatment with tiotropium. *Eur Respir J* 2002;19:209–216.
41. Calverley P, Pauwels R, Vestbo J, Jones P, Pride N, Gulsvik A, Anderson J, Maden C; TRial of Inhaled STeroids ANd long-acting beta2 agonists study group. Combined salmeterol and fluticasone in the treatment of chronic obstructive pulmonary disease: a randomised controlled trial. *Lancet* 2003;361:449–456.
42. Zheng JP, Yang L, Wu YM, Chen P, Wen ZG, Huang WJ, Shi Y, Wang CZ, Huang SG, Sun TY, *et al*. The efficacy and safety of combination salmeterol (50 microg)/fluticasone propionate (500 microg) inhalation twice daily via accuhaler in Chinese patients with COPD. *Chest* 2007; 132:1756–1763.
43. Vestbo J, Anderson JA, Calverley PM, Celli B, Ferguson GT, Jenkins C, Yates JC, Jones PW. Bias due to withdrawal in long-term randomised trials in COPD: evidence from the TORCH study. *Clin Respir J* 2011;5:44–49.
44. Jenkins CR, Jones PW, Calverley PM, Celli B, Anderson JA, Ferguson GT, Yates JC, Willits LR, Vestbo J. Efficacy of salmeterol/fluticasone propionate by GOLD stage of chronic obstructive pulmonary disease: analysis from the randomised, placebo-controlled TORCH study. *Respir Res* 2009;10:59.
45. Copay AG, Subach BR, Glassman SD, Polly DW Jr, Schuler TC. Understanding the minimum clinically important difference: a review of concepts and methods. *Spine J* 2007;7:541–546.
46. Westwood M, Bourbeau J, Jones PW, Cerulli A, Capkun-Niggli G, Worthy G. Relationship between FEV_1 change and patient-reported outcomes in randomised trials of inhaled bronchodilators for stable COPD: a systematic review. *Respir Res* 2011;12:40.
47. Decramer M, Molenberghs G, Liu D, Celli B, Kesten S, Lystig T, Tashkin DP; UPLIFT investigators. Premature discontinuation during the UPLIFT study. *Respir Med* 2011;105:1523–1530.
48. Gerlinger C, Schumacher U, Faustmann T, Colligs A, Schmitz H, Seitz C. Defining a minimal clinically important difference for endometriosis-associated pelvic pain measured on a visual analog scale: analyses of two placebo-controlled, randomized trials. *Health Qual Life Outcomes* 2010;8:138.
49. Kekow J, Moots RJ, Emery P, Durez P, Koenig A, Singh A, Pedersen R, Robertson D, Freundlich B, Sato R. Patient-reported outcomes improve with etanercept plus methotrexate in active early rheumatoid arthritis and the improvement is strongly associated with remission: the COMET trial. *Ann Rheum Dis* 2010;69:222–225.
50. Norman GR, Sridhar FG, Guyatt GH, Walter SD. Relation of distribution- and anchor-based approaches in interpretation of changes in health-related quality of life. *Med Care* 2001;39:1039–1047.
51. Kerwin E, Hébert J, Gallagher N, Martin C, Overend T, Alagappan VK, Lu Y, Banerji D. Efficacy and safety of NVA237 versus placebo and tiotropium in patients with COPD: the GLOW2 study. *Eur Respir J* 2012;40:1106–1114.
52. Mahler DA, Decramer M, D'Urzo A, Worth H, White T, Alagappan VK, Chen H, Gallagher N, Kulich K, Banerji D. Dual bronchodilation with QVA149 reduces patient-reported dyspnoea in COPD: BLAZE study. *Eur Respir J* [online ahead of print] 31 Oct 2013; DOI: 10.1183/09031936.00124013.