

What are minimal important changes for asthma measures in a clinical trial?

N.C. Santanello*, J. Zhang⁺, B. Seidenberg⁺, T.F. Reiss⁺, B.L. Barber[#]

What are minimal important changes for asthma measures in a clinical trial? N.C. Santanello, J. Zhang, B. Seidenberg, T.F. Reiss, B.L. Barber. ©ERS Journals Ltd 1999.

ABSTRACT: In this study, the perceptions of asthmatics to change in their disease was associated with observed changes in clinical asthma measures, in order to identify the threshold where changes in clinical asthma measures are perceivable by patients.

The study included 281 asthmatic patients, aged 18–63 yrs, in a randomized, placebo-controlled clinical trial of a leukotriene antagonist. Changes were related in: 1) asthma symptom scores; 2) inhaled β -agonist use; 3) forced expiratory volume in one second (FEV₁); and 4) peak expiratory flow (PEF) to a global question that queried overall change in asthma since starting the study drug. Additional analyses examined differences in the group reporting minimal improvement by treatment (active treatment *versus* placebo), sex and age groups.

The average minimal patient perceivable improvement for each measure was: 1) -0.31 points for the symptom score on a scale of 0–6; 2) -0.81 puffs·day⁻¹ for inhaled β -agonist use; 3) 0.23 L for FEV₁; and 4) 18.79 L·min⁻¹ for PEF. In general placebo-treated patients and older patients, who reported minimal improvement, experienced less mean improvement from baseline than active-treated patients and younger patients, who reported minimal improvement.

Determining the minimal patient perceivable improvement value for a measure may be helpful to interpret changes. However, interpretation should be carried out cautiously when reporting a single value as a clinically important change.

Eur Respir J 1999; 14: 23–27.

*Merck Research Laboratories, West Point, PA, USA. ⁺Merck Research Laboratories, Rahway, NJ, USA. [#]Lilly and Company, Inc., Indianapolis, IN, USA.

Correspondence: N.C. Santanello
PO Box 4, BL 1-7
West Point
PA 19486
USA
Fax: 1 6103972992

Keywords: Asthma
epidemiology
methods
questionnaires

Received: October 7 1998

Accepted after revision January 24 1999

Asthma is a chronic respiratory disease characterized histologically by airway inflammation and mucus hypersecretion, and clinically by reversible airway obstruction and specific patient-reported symptoms [1]. In clinical trials, changes in asthma are measured by pulmonary function tests, such as forced expiratory volume in one second (FEV₁) and peak expiratory flow (PEF), as well as by β -agonist inhaler use, and validated symptom scales that capture patient-reported symptoms [2]. While a clinical trial can be powered to find small differences between treatment groups on any one of these asthma measures, the meaning of these changes is often unknown. The question is often asked: "Do differences in asthma measures noted at the end of the study matter to the patients, that is are they clinically meaningful?". As an example, is an average of 1 or 2 puffs·day⁻¹ decrease in β -agonist inhaler use or an increase of 8 or 10% in FEV₁ perceived by patients?

One method of providing clinical meaning is to relate change in one measure to a global measure of patient-rated change. For example, JAESCHKE *et al.* [3] related average changes in scores on a health-related quality-of-life (HRQoL) questionnaire to patients' global ratings of change in their disease within an observational study environment. In this way, they provided meaning to changes in scores from baseline within groups on a newly developed HRQoL questionnaire.

In this study, the theoretical framework of JAESCHKE *et al.* [3] was utilized in order to provide clinical meaning to

changes in asthma measures observed in a clinical trial. Specifically, the average score changes from baseline for: 1) a validated asthma daytime symptom scale; 2) β -agonist inhaler use; 3) FEV₁; and 4) PEF were related to patient global ratings of change in asthma. However, unlike JAESCHKE *et al.* [3], and others [4, 5], this study looked at changes from baseline for active treatment and placebo groups, and sex and age groups, in order to determine whether treatment group assignment, age and sex influence patient perception of change.

Methods

Subjects

Study participants were adult patients, 18–65 yrs of age, with mild-to-moderate asthma symptoms who were evaluated in a clinical trial. Patients were required to have: a history of typical, recurring symptoms of asthma; FEV₁ between 40–80% of predicted; and evidence of reversibility of airway obstruction after inhaled β -agonist (at least 15% increase in FEV₁). Further, all patients were required to be able to complete a daily symptom diary and perform pulmonary function tests (spirometry and PEF) in a consistent manner.

Participants signed informed consents and the study was approved by an Institutional Review Board.

Study design

The study was a blinded, placebo-controlled, dose ranging clinical trial of an investigational compound (three dose levels of the test compound, leukotriene antagonist (montelukast) and matching placebo). It included a 2-week placebo run-in period, a 3-week double-blind treatment period and a 1-week single-blind, wash-out period [6].

Patients completed a validated daytime asthma symptom scale and twice daily peak flow measurements on a diary record. At a weekly clinic visit, patients returned the diary cards, and they performed pulmonary function tests. At the end of the treatment period patients assessed their perception of the change in their asthma on a global question.

Measurements

The validated daytime symptom score [2] consists of four questions to assess frequency of symptoms, the extent to which symptoms were troublesome, frequency of limitation in usual activities and how asthma symptoms affect this limitation in activities. The scale was completed each evening on the diary record. Additionally, patients recorded their PEF and the number of puffs of inhaled β -agonist used (upon arising in the morning and at bedtime) on the diary record. At weekly clinic visits FEV₁ was measured.

The average of the four daily daytime symptom questions was computed as the daily asthma symptom score. Then, weekly averages of the daily asthma scale scores were used as an overall daily score for the week. The range for responses to each question was 0–6, where 0 indicated the least asthma symptomatology and 6 the most. The change in the asthma symptom scale score (possible range: -6 indicating improvement to 6 indicating worsening) was computed as the difference between the average daily score from the final week of the placebo run-in period to the final week of the active treatment period. The change in average daily β -agonist inhaler use and PEF was computed in a similar manner. The change in FEV₁ was computed as the difference from the baseline visit at the end of the placebo run-in period to the last active treatment period study visit.

The global questions of change in asthma (one given to the patient, one given to the investigator) asked the patient or investigator to assess the change in the patient's asthma. These global assessments were completed by the patients and investigators at the end of the treatment-period clinic visit while still blinded to treatment assignment. Response options for the globals were presented as seven-point scales: very much better, moderately better, a little better, unchanged, a little worse, moderately worse, and very much worse.

When completing the global rating, patients were blinded to the results of their pulmonary function test (FEV₁ and forced vital capacity (FVC)), and investigators were blinded to the patient-reported symptoms diary responses and the patients' global assessments of change in their asthma.

Lung function was measured by standardized pulmonary function tests at weekly clinic visits.

Statistical analysis

JAESCHKE *et al.* [3] related changes in score on an HRQoL questionnaire to changes in a global rating of disease. Similarly, the present study related changes in the diary score, FEV₁, averaged morning and evening PEF, and β -agonist use to the patients' global ratings of change in their asthma. For each category of the global, the average (and quartiles) change from baseline were calculated in each asthma measure (diary symptom score, FEV₁, PEF and β -agonist use). Unlike JAESCHKE *et al.* [3] and others [4, 5], this study found differences in the magnitude of change for the better and worse global categories, and therefore did not combine categories into one measure of change. Rather, the minimal patient perceivable improvement (MPPI) and the minimal patient perceivable deterioration (MPPD) were examined. The MPPI and MPPD describe the smallest average change in a score that corresponds to patients perceiving improvement or deterioration.

The impact of treatment group, sex and age on the MPPI was analysed by comparing the MPPI for the placebo group *versus* the active treatment groups, for males *versus* females, and for patients older than the median age (33 yrs) *versus* those younger than the median age. Similar categories could not be used for the MPPD due to the small numbers reporting a deterioration in the asthma measurements.

Although the focus of this study is on patient perceivable change, it was also of interest to examine the association between the clinician and patient perceptions of change. Patient and clinician ratings were cross-tabulated on the global assessment of change in asthma to assess the level of agreement in their ratings. That is, the patient global change scores were compared to the clinician global change scores and the sample measure weighted kappa [7] calculated to describe the closeness of their agreement (using quadratic weights).

The study was designed *a priori* with a sample size of 50 patients per group to have 80% power to detect ($\alpha=0.05$, two-tailed test) a mean difference between groups in FEV₁ of 11% (per cent change from baseline). The resulting over recruitment to ~70 per group gave the study 80% power to detect a 9.3% difference in FEV₁. The study was not designed specifically for the analysis of minimal perceivable improvement, as the number of patients who would fall into each category could not be known prior to study completion. However, with 86 subjects in the minimally improved group, the standard error for FEV₁ change from baseline would be 0.064.

Results

A total of 281 patients were enrolled into the study. Eight patients discontinued prior to completing the study; therefore, complete data on 273 (97%) patients is available for this report. The baseline characteristics for all patients entered, are reported in table 1.

Tables 2–6 present the average change in asthma measures for each category of patient global rating. The global category "A little better" provides the average MPPI and the category "A little worse" provides the average MPPD. There were very few patients in the

Table 1. – Baseline characteristics by treatment group

	Placebo (n=69)	Active treatment (n=212)	All (n=281)
Age yrs			
Mean	37.23	34.19	34.94
Range	(18–63)	(18–63)	(18–63)
Sex n (%)			
Female	35 (51)	94 (44)	129 (46)
Male	34 (49)	118 (56)	152 (54)
Race n (%)			
Caucasian	61 (88)	183 (86)	244 (87)
Other	8 (12)	29 (14)	37 (13)
FEV ₁ L·s ⁻¹			
Mean±SD	2.27±0.72	2.31±0.64	2.30±0.66
Range	(1.04–4.28)	(0.93–4.04)	(0.93–4.28)
Daily β-agonist puffs·day ⁻¹			
Mean±SD	5.47±2.54	5.41±2.70	5.43±2.66
Range	(0.71–12.23)	(0.77–14.57)	(0.71–14.57)
PEF L·min ⁻¹			
Mean±SD	396.0±93.7	401.3±78.6	400.0±82.5
Range	(212.9–577.9)	183.9–612.9	(183.9–612.9)

FEV₁: forced expiratory volume in one second; PEF: peak expiratory flow.

"Moderately worse" and "Very much worse" categories. Therefore, little confidence is placed on the average changes in the asthma measures associated with these two categories.

Table 2 reports the average change in the asthma symptom score for each category of patient global rating. The MPPI was -0.31 and the MPPD was 0.19 for all treatment groups combined. The average MPPI differed by treatment groups (table 3). For the active treatment groups combined the MPPI was -0.37, whereas the MPPI for the placebo group was -0.10. The MPPI for males (-0.29) and females (-0.33) was similar. Using the median age (33 yrs) to divide the patients into two groups (older *versus* younger) the average MPPI for the older group was smaller (MPPI=-0.19) than the younger group (MPPI=-0.41).

In general, for the measures of airflow (FEV₁ and PEF), the average changes were in the expected direction (tables 4 and 5); that is, as patients reported improvement or deterioration in their asthma the average change in airflow measures increased or decreased, respectively. The MPPI for FEV₁ was 0.23 L (10.38% change from baseline) for all treatment groups combined. The active treat-

Table 2. – Average (quartiles) changes in symptom score by patient global change category

Patient-determined global category	n	Average (quartiles) change from baseline in symptom score*
Very much worse	1	0.17
Moderately worse	2	0.94
A little worse	10	0.19 (-0.23–0.45)
Unchanged	52	0.07 (-0.20–0.38)
A little better (MPPI)	86	-0.31 (-0.71–0.10)
Moderately better	74	-0.55 (-0.90–0.01)
Very much better	48	-0.97 (-1.54–0.13)

MPPI: minimal patient perceivable improvement. *: symptom scale score range 0–6, with 0=no symptoms, and 6=maximum symptoms.

Table 3. – Average minimal patient perceivable improvement (MPPI) values in symptom score, forced expiratory volume in one second (FEV₁), peak expiratory flow (PEF), and β-agonist use by treatment, sex and age groups

Subgroups	MPPI n	Average MPPI from baseline for asthma measures			
		Diary scale score	FEV ₁ L	PEF L·min ⁻¹	β-agonist use puffs·day ⁻¹
Active treatment	67	-0.37	0.25	22.1	-1.03
Placebo treatment	19	-0.10	0.18	7.2	-0.04
Male	42	-0.29	0.25	19.0	-0.67
Female	44	-0.33	0.22	18.6	-0.94
Older than median age*	39	-0.19	0.17	19.4	-0.54
Younger than median age	47	-0.41	0.28	18.3	-1.03

*: median age=33 yrs for all patients enrolled.

ment groups combined showed an MPPI for FEV₁ of 0.25 L compared to 0.18 L for placebo group. Males (0.25 L) and females (0.22 L) showed a similar MPPI for FEV₁, whereas, consistent with what was found for the asthma symptom score, older patients had a lower MPPI for FEV₁ (0.17 L) than younger patients (0.28 L) (table 3). Furthermore, the values for change in FEV₁ were consistent across patients in the MPPI group with different baseline FEV₁ values (the correlation between the change in FEV₁ and baseline FEV₁ values in the MPPI group was nearly zero).

The MPPI for PEF was 18.8 L·min⁻¹ (or 5.39%) for all treatment groups combined. Again, the placebo group had a lower MPPI (7.2 L·min⁻¹) than the active treatment group (22.1 L·min⁻¹), as shown in table 3. The MPPI for males and females were similar (19.0 L·min⁻¹ *versus* 18.6 L·min⁻¹, respectively). There were no apparent differences in average MPPI values between older (19.4 L·min⁻¹) and younger subjects (18.3 L·min⁻¹).

As expected, daily β-agonist inhaler use increased as patients reported a deterioration in their asthma and decreased as patients reported an improvement in their asthma (table 6). The MPPI for β-agonist inhaler use was -0.81 puffs·day⁻¹ for all treatment groups combined, -0.04 puffs·day⁻¹ for placebo and -1.03 puffs·day⁻¹ for active treatment groups. There were differences observed for the MPPI use of β-agonist inhaler between males and females (-0.67 and -0.84 puffs·day⁻¹, respectively) and between

Table 4. – Average (quartiles) changes in forced expiratory volume in one second (FEV₁) by patient global change category

Patient determined global category	n	Average (quartiles) change from baseline in FEV ₁ L
Very much worse	1	-0.62
Moderately worse	2	-0.01
A little worse	10	0.08 (-0.10–0.26)
Unchanged	52	0.11 (-0.12–0.34)
A little better (MPPI)	86	0.23 (0.01–0.48)
Moderately better	74	0.25 (-0.14–0.55)
Very much better	48	0.38 (0.10–0.63)

MPPI: minimal patient perceivable improvement.

Table 5. – Average (quartiles) changes in peak expiratory flow (PEF) by patient global change category

Patient determined global category	n	Average (quartiles) change from baseline in PEF L·min ⁻¹
Very much worse	1	-6.7
Moderately worse	2	-15.6
A little worse	10	-14.3 (-31.8–0.8)
Unchanged	52	3.4 (-9.8–19.8)
A little better (MPPI)	86	18.8 (0.7–36.9)
Moderately better	74	22.1 (0.4–43.3)
Very much better	48	35.2 (12.9–55.2)

MPPI: minimal patient perceivable improvement.

older and younger patients (-0.54 and -1.03 puffs·day⁻¹, respectively) (table 3).

The weighted kappa statistic describing the agreement between patient and investigator global responses was 0.66 (95% confidence interval 0.57–0.73). The MPPI values for each of the four asthma measures were relatively similar to the minimum investigator-perceived improvement. When the patient and investigator disagreed, the patient more often than not rated themselves as more improved than the investigator perceived them to be.

Discussion

Previous research determined at what point asthmatics begin to perceive deterioration in their lung function [8–10]. The analysis presented in this study determined the level at which a population of asthmatics in a clinical trial perceived change, improvement or deterioration, in lung function and other asthma measures after treatment. The clinical meaning was determined by relating changes in four measures of asthma (FEV₁, PEF, β -agonist use and an asthma symptoms score) to patients' perception of change in their asthma. This approach has been used previously in quality-of-life research to describe the amount of change in a quality-of-life measure that is clinically meaningful to a group of patients within a study [3–5, 11, 12].

The present study found that the distribution of average MPPI and MPPD scores for changes in symptom score, FEV₁, PEF and inhaled β -agonist use were not symmetric. That is, the average MPPI change value differed from the average MPPD value for each asthma measure in distance from the "unchanged" category. Therefore, the MPPI and MPPD were not combined into one measure of change. This finding differs from that reported by others [3–5] who

Table 6. – Average (quartiles) changes in β -agonist inhaler use by patient global change category

Patient determined global category	n	Average (quartiles) change from baseline in β -agonist use puffs·day ⁻¹
Very much worse	1	3.09
Moderately worse	2	4.11
A little worse	10	0.44 (-1.21–2.33)
Unchanged	51	0.13 (-0.67–0.56)
A little better (MPPI)	86	-0.81 (-1.67–0.28)
Moderately better	74	-1.47 (-2.07–-0.50)
Very much better	48	-2.10 (-3.10–-0.88)

MPPI: minimal patient perceivable improvement.

looked at minimal changes from baseline within an observational study design where the change was symmetrical (improvement and deterioration had approximately equal magnitude).

The present study found that there were differences in the clinically important improvements for all four asthma measures between treatment and age groups. Patients in the placebo group who rated themselves as minimally improved tended to have experienced less improvement in the four measures than the active treatment group. Likewise, older patients, on average, needed less improvement in asthma measures to report improvement than younger patients. For the four asthma measures, only puffs of β -agonist differed slightly in minimal improvement using the sex groups (males who reported minimal improvement experienced on average a lesser decrease in puffs than females who reported minimal improvement). However due to sample size limitations the authors were unable to test for statistical differences and, therefore, differences between groups may be due to chance alone. Differences in MPPI values for different levels of severity were not examined, since the MPPI value for change in FEV₁ did not systematically vary from baseline FEV₁ (the line describing the relationship between baseline FEV₁ and change in FEV₁ in the MPPI group was close to zero).

Previous studies examined minimally important changes from baseline for quality-of-life measures in prospective observational studies [3–5]. Within these studies only the change from baseline in an asthma cohort who reported change on a global question was calculated. This change from baseline was termed the minimal important difference (MID). The MID has been used as a way to determine whether the difference observed between treatment groups is clinically important [3–5, 12, 13]. However, the methodology underlying the calculation of the MID is based on change from baseline in an observational cohort of asthmatics who report change on a global basis, and, as such, cannot examine whether there are differences in values for what are perceived as important between treatment groups. Using blinded clinical trial data for the present analyses, it was found that the change from baseline for the placebo group who rated themselves as minimally improved was less than that seen in the active treatment group who rated themselves as minimally improved for each asthma measure. These findings are consistent with those expected based on an analysis in a study by NORMAN *et al.* [14] comparing "retrospective" change (change observed in a cohort who report change on a global basis) with "treatment" induced change (change observed as a result of an applied treatment).

It might easily be accepted that there would be differences in the MPPI between older and younger patients as was found in this study. Different age groups (older or younger patients) might be more sensitive to smaller changes in asthma measures to define improvement or deterioration. CONNOLLY *et al.* [8] showed that elderly asthmatic patients were less aware of bronchoconstriction than younger patients. However, in the present study, it had been assumed that no large differences would be found between the treatment groups in the amount of change in a measure perceived as minimal improvement, since the data were collected in a blinded, placebo-controlled, randomized clinical trial. A possible explanation for the differences found between treatment groups may be that

within a blinded, placebo-controlled, randomized clinical trial, all patients generally expect to improve. Therefore, the blinded placebo group might be more sensitive to smaller improvements in clinical asthma measures and rate themselves as improved more easily, even without evidence of greater change in asthma measures. Additionally, there may be a threshold range of effect, such that although many patients fell into the range, the active treatment group responses are distributed more towards the top of the threshold range and the placebo group more towards the bottom, as evidenced by the difference in average changes.

The patient and investigator minimal improvement was comparable in magnitude for each asthma measure. This finding was expected since the weighted kappa of 0.66 indicated relatively strong agreement between patients' and clinician's global ratings.

There are limitations to the use of this method in providing clinical meaning to changes in clinical measures. Although the average changes in the asthma measures showed an orderly progression in the expected direction with the categories of change in the global ratings, considerable variability existed among individuals. One explanation for this is that individual patients perceive symptoms differently and hence also perceive the magnitude of change in their asthma differently when answering the global change [8–10]. Another reason is that global ratings do not represent a "gold standard" for assessment of change in these measures, but rather provide direction and magnitude of average effects. This is consistent with the authors' previous research where it was demonstrated that the value for minimal improvement differed based on the type of global question asked [15].

With these limitations in mind, it then becomes clear that the minimal patient perceivable change for a measure defined from clinical trial data should not be used by a clinician to define clinical improvement or deterioration in the care of an individual patient. Rather, the results of this study might provide a basis for interpreting the clinical meaning of average changes in asthma measures reported within clinical trials.

In summary, statistically significant differences from placebo may not be sufficient to demonstrate clinical improvement or deterioration, if the clinical meaning of the difference is not understood [12]. The method of relating change in clinical measures to patients' overall rating of change in their disease may be useful to provide an estimate of what level of change from baseline or difference between treatment groups in a measure is clinically relevant to patients. However, this analysis of a blinded, placebo-controlled, randomized clinical trial demonstrated that what patients perceive as minimally clinically relevant may differ by treatment group and age. Therefore, while a level of minimal change in a measure that is defined as clinically relevant by this method might be useful, caution is advised against setting a universal benchmark of what is an important change for a measure based on this method for many reasons [14], including possible differences between treatment groups and other

demographic groups. Based on the present findings, further research appears warranted to determine whether minimal improved values for measures vary for other treatments and other types of patient populations.

Acknowledgements. The authors would like to thank Z. Zhao for assistance in the statistical analyses.

References

1. Busse WW, Reed CE. Asthma: definition and pathogenesis. *In: Middleton E, Reed CE, Ellis EF, Adkinson HF, Yunginger JW, eds. Allergy: Principles and Practice. 3rd Edn. St. Louis, CV Mosby 1988; p. 969.*
2. Santanello NC, Barber BL, Friedman B, Reiss TF, Zhang J. Measurement characteristics of two asthma symptom diaries validated in two separate clinical trials. *Eur Respir J* 1997; 10: 646–651.
3. Jaeschke R, Singer J, Guyatt GH. Measurement of health status. *Control Clin Trials* 1989; 10: 407–415.
4. Juniper EF, Guyatt GH, Willan A, Griffith LE. Determining a minimal important change in a disease-specific quality of life questionnaire. *J Clin Epidemiol* 1994; 47: 81–87.
5. Redelmeier DA, Guyatt GH, Goldstein RS. Assessing the minimal important difference in symptoms: a comparison of two techniques. *J Clin Epidemiol* 1996; 49: 1215–1219.
6. Noonan MJ, Chervinsky P, Brandon M, *et al.* Montelukast, a potent leukotriene receptor antagonist, causes dose-related improvements in chronic asthma. *Eur Respir J* 1998; 11: 1232–1239.
7. Spitzer RL, Cohen J, Fleiss JL, Endicott J. Quantification of agreement in psychiatric diagnosis. *Arch Gen Psychiatry* 1967; 17: 83–87.
8. Connolly MJ, Crowley JJ, Charan NB, Nielson CP, Vestal RE. Reduced subjective awareness of bronchoconstriction provoked by methacholine in elderly asthmatic and normal subjects as measured on a simple awareness scale. *Thorax* 1992; 47: 410–413.
9. Rubinfeld AR, Pain MC. Conscious perception of bronchospasm as a protective phenomenon in asthma. *Chest* 1977; 72: 154–158.
10. Rubinfeld AR, Pain MC. Perception of asthma. *Lancet* 1976; i: 882–884.
11. Lydick E, Epstein RS. Interpretation of quality of life changes. *Qual Life Res* 1993; 2: 221–226.
12. Juniper EF. Quality of life questionnaires: does statistically significant=clinically important? *J Allergy Clin Immunol* 1998; 102: 16–17.
13. Guyatt GH, Juniper EF, Walter SD, Griffith LE, Goldstein RS. Interpreting treatment effects in randomized trials. *BMJ* 1998; 316: 690–693.
14. Norman GR, Stratford P, Regehr G. Methodological problems in the retrospective computation of responsiveness to change: the lesson of Cronbach. *J Clin Epidemiol* 1997; 50: 869–879.
15. Barber BL, Santanello NC, Epstein RS. Impact of the global on patient perceivable change in an asthma specific QOL questionnaire. *Qual Life Res* 1996; 5: 117–122.