Correlation of airway obstruction and patient-reported endpoints in clinical studies

S. Shingo*, J. Zhang*, T.F. Reiss#

Correlation of airway obstruction and patient-reported endpoints in clinical studies. S. Shingo, J. Zhang, T.F. Reiss. ©ERS Journals Ltd 2001.
ABSTRACT: To establish the correlation among asthma efficacy parameters over a long period, data from over 1500 patients in two one-year asthma clinical trials with montelukast, a Cys-LT1 antagonist, were analysed. Airway obstruction measurements, forced expiratory volume in one second (FEV1) and peak expiratory flow (PEF), were measured at clinic visits. Patients recorded daytime symptom score, "as-needed" β-agonist use, and PEF on a daily basis.

Relationships among these parameters at baseline and during the one-year treatment period were established by correlation analyses. Multiple correlations between the airway obstruction (FEV1 and PEF) and patient-reported measurements were evaluated by canonical correlation analysis.

Pairwise correlations of the efficacy parameters over a one-year time period were stable. Canonical correlation between the airway obstruction and patient-reported asthma efficacy endpoints was low, indicating that each category of endpoints measures a distinctively different aspect of the disease.

It appears that at least one endpoint from each category should be used in asthma clinical studies.
Eur Respir J 2001; 17: 220–224.

*Dept of Clinical Biostatistics, Merck Research Laboratories, Rahway, NJ 07065, USA and #Dept of Pulmonary and Immunology, Merck Research Laboratories, Rahway, NJ 07065, USA.

Correspondence: S. Shingo, Merck Research Laboratories, PO Box 2000, RY33–404, Rahway, NJ 07065, USA
Fax: 1 7325946075

Asthma is a variable and multifaceted disease. Many efficacy endpoints are measured and reported in asthma clinical trials. Cross-sectional correlations among the most commonly used asthma-related endpoints such as forced expired volume in one second (FEV1), peak expiratory flow (PEF), "as-needed" β-agonist use, and daytime and night-time symptoms have been explored, mostly in a small group of patients over short periods of time, and mild to moderate correlations have been observed [1–3].

This paper investigates the relationship (over a one-year period) among the endpoints of FEV1, PEF, "as-needed" β-agonist use, and daytime symptoms, and the relationship between the pulmonary function measurements (FEV1, PEF) and patient-reported efficacy measurement endpoints (daytime symptom, "as-needed" β-agonist use). The relationships were investigated based on a large data set collected over a one-year period, from 1576 patients who entered two large clinical trials, which were conducted as multi-centre, double-blind, placebo-controlled studies to determine the safety, tolerability, and efficacy of one-year continuous treatment of montelukast, a Cys-LT1

antagonist [4, 5]. The relationships among these endpoints will provide more information about this multifaceted disease and establish the multi-dimensionality of these endpoints.

Methods

Study design

Study 1 was a multi-centre, double-blind, parallel-group, placebo-controlled study comparing montelukast (MNT), inhaled beclomethasone (ICS) and placebo (table 1). The study consisted of a 2-week single-blind placebo run-in period (I), a 12-week, double-blind, active treatment period (II) and a 3-week, double-blind washout period (III). Patients taking concomitant oral theophylline (limited to approximately 25% of patients) were allowed to continue this medication at a constant dose. At the end of Period II, a prespecified subset of patients, determined by initial allocation from the MNT and ICS groups, along with all the placebo group patients,

Table 1. – Study procedure

Study period (weeks)	Treatment			
	Study 1		Study 2	Visit spacing
Screening (1)	Open		Open	1 week
Baseline placebo run-in (2)	Single-blind	Placebo	Single-blind	Placebo	2 weeks
Treatment (15)	Double-blind	Placebo MNT ICS	Double-blind	Placebo MNT	3 weeks
Extension (37)	Double-blind	MNT ICS	Open	MNT ICS	4 – 8 weeks

MNT: montelukast; ICS: inhaled beclomethasone.

entered a 3-week placebo washout period to assess the clinical effects of withdrawal; these patients were discontinued from the study at the end of the washout period. The remaining patients in the MNT and ICS groups were qualified to continue on the same treatment for the 37-week extension period (IV) [4].

Study 2 was similar in design to Study 1 with the exception that there was no ICS treatment group in Period II (20% of patients were allowed concomitant use of inhaled steroid). At the end of Period II, a prespecified subset of patients from the MNT group, determined by initial allocation, entered a 3-week placebo washout period to assess the clinical effects of withdrawal. In contrast to Study 1, however, all patients were eligible to participate in the extension period, even those who received placebo. Qualifying, consenting MNT and placebo patients received either MNT or ICS during the extension. Patients using ICS concomitantly in the primary study period and randomized to the ICS group in the extension maintained their concomitant ICS doses (table 1) [5].

Patients

For both studies, healthy, nonsmoking males and females, aged $\geqslant 15$ yrs, with $\geqslant 1$ yr of intermittent or persistent asthma symptoms, were enrolled. To be eligible for randomization, patients had to have an FEV_1 between 50% and 80% of the predicted value (after withholding β-agonist for at least 6 h) and an absolute increase in FEV_1 of $\geqslant 15\%$ or greater, 20 – 30 min after inhalation of β-agonist on two occasions during Screening and Period I. In addition, patients were required to have a minimum daily average daytime asthma symptom score of 1.14 (0 – 6 point scale) and an average of at least one puff of β-agonist per day during the 2 week run-in period. Each patient received a mini-Wright Peak Flow Meter (Clement Clark, Columbus, Ohio, USA), and diary cards were distributed at each clinic visit.

Study procedure

Spirometry was performed at each clinic visit between 6 and 9 am, approximately 10 – 12 h after the previous dose of study medication (after Visit 2). The airway obstruction measurements (FEV_1 and PEF) were collected with a standard spirometer (Puritan-Bennett PB 100/PB110, Kansas City, KS, USA) and transmitted *via* modem to a central spirometry quality control centre, where a review of the data was performed to ensure uniform adherence to American Thoracic Society standards of acceptability and reproducibility [6, 7]. The largest FEV_1 and PEF values from a set of at least three manoeuvres, were selected as the visit value. Airway reversibility was evaluated at each visit during the run-in period (I) and at predefined visits during the active treatment period. A validated diary card containing daytime asthma symptom scales was used by patients to record diary card data on a daily basis. The scales had been shown to have acceptable evaluative measurement properties [8, 9]. The four daytime asthma symptom scales (assessing the frequency, severity and bothersome on a 0 – 6 scale) were combined into a mean daily score. Patients recorded daytime symptoms and "as-needed" β-agonist use on the diary card at bedtime, and night-time awakenings and "as-needed" β-agonist use on the diary card upon arising. PEF was measured twice daily by the patient: in the morning upon arising and in the evening at bedtime (before taking the study medication). The largest of three measurements at each time point was recorded on the diary card.

Statistical methods

Data from all patients with a baseline value and at least one treatment period measurement were included in the analysis. For patient-reported endpoints (on the diary card), mean values for daily daytime symptom score, β-agonist use, and PEF were calculated from the daily entries between two visits, and defined as a visit response. The mean of measurements, taken during Period I (placebo run-in), was defined as the baseline value.

For correlation coefficient calculation, visit-specific observed values were used. For all calculations described above, data from both studies and all treatment groups were pooled; the results based on each treatment group were similar.

Pairwise correlation. Pairwise correlations among all endpoints (FEV_1, daytime symptom score, "as-needed" β-agonist use, or PEFs (both measured-at-clinic and patient-reported)) were established by

Table 2. – Patient baseline characteristics (all treatment groups pooled)

Characteristic	Study 1	Study 2
Sample size n	895	681
Age yrs (range)	38 (15–85)	33 (15–79)
Sex: male (%)	354 (40%)	305 (45%)
female (%)	541 (60%)	376 (55%)
FEV_1 L	2.2 ± 0.6	2.5 ± 0.7
FEV_1 % predicted	65.2 ± 10.4	66.8 ± 10.8
Measured-at-clinic PEF ($L \cdot min^{-1}$)	352.9 ± 123.2	388.1 ± 117.4
Patient-reported morning PEF ($L \cdot min^{-1}$)	334.9 ± 97.4	385.3 ± 88.8
Daytime symptom score (0–6)	2.4 ± 0.9	2.5 ± 0.8
"As-needed" β-agonist use $puffs \cdot day^{-1}$	5.5 ± 3.8	5.4 ± 3.2

Results expressed as mean $\pm$ SD unless otherwise stated. FEV_1: forced expired volume in one second; PEF: peak expiratory flow.

calculating Pearson's correlation coefficients on the visit-specific observed values, at baseline and at each visit in the treatment and extension periods. Spearman's correlation coefficients, which measures the relationship of the ranks of the measurements, were also computed and were similar to the Pearson's correlation coefficients.

Canonical correlation. Canonical correlations were also calculated to establish the relationship between the two sets of endpoints, *i.e.* airway obstruction (FEV_1 and PEF) and patient-reported (daytime symptoms and "as-needed β-agonist use") endpoints. Canonical correlation, analogous to multiple regression, captures the relationship between two sets of variables.

Results

The baseline characteristics of the 1576 patients from the two studies, including demographic information and summary statistics of efficacy parameters, are displayed in table 2. Baseline summary statistics of FEV_1, FEV_1 % predicted, "as-needed" β-agonist use, daytime symptom score and PEFs (both measured-at-clinic and patient-reported), were comparable between the two studies.

Pairwise correlation coefficients

Pearson's correlations between each pair of efficacy endpoints are displayed in table 3. The correlation between FEV_1 and either patient-reported or measured-at-clinic PEF was strong (0.74 and 0.85, respectively), compared to that between FEV_1 and daytime symptom score, or FEV_1 and "as-needed" β-agonist use, which were mild (-0.13 and -0.22, respectively). The correlations between patient-reported PEF and measured-at-clinic PEF was strong (0.77). These correlations were generally stable over time (figures 1, 2). At the beginning of the study, the number of patients was 1576, then, at the end of the double-blind treatment period (Week 15), it was reduced to 1411 due to patients' withdrawal from the study. Since some of the patients did not enter the extension treatment (due to study design and/or patient's selection), the number of patients decreased to 426 toward the end of the extension period (Week 52).

Canonical correlations

The canonical correlation, calculated for each visit-specific value, between the two sets of endpoints, (FEV_1, measured-at-clinic PEF) and (daytime symptom, "as-needed" β-agonist use) was low (range 0.202–0.269). This relationship was confirmed by the canonical correlation between patient-reported PEF and the rest of the diary endpoints (daytime symptom score, "as-needed" β-agonist use), which was approximately 0.213–0.274, indicating that pulmonary function and patient-reported endpoints (daytime symptom, "as-needed" β-agonist use) measure two different aspects of asthma. The canonical correlations between

Table 3. – Pearson's correlation coefficients between efficacy endpoints

Efficacy endpoints	Mean*	Median (range)*
FEV_1 and measured-at-clinic PEF	0.85	0.86 (0.75–0.88)
FEV_1 and patient-reported PEF	0.74	0.75 (0.70–0.76)
FEV_1 and daytime symptom	-0.13	-0.13 (-0.17 – -0.08)
FEV_1 and β-agonist use	-0.22	-0.22 (-0.26 – -0.18)
Measured-at-clinic PEF and patient-reported PEF	0.77	0.78 (0.69–0.80)
Daytime symptom and measured-at-clinic PEF	-0.15	-0.16 (-0.19 – -0.09)
Daytime symptom and patient-reported PEF	-0.19	-0.19 (-0.22 – -0.15)
Daytime symptom and β-agonist use	0.55	0.58 (0.42–0.62)
β-agonist use and measured-at-clinic PEF	-0.20	-0.21 (-0.30 – -0.09)
β-agonist use and patient-reported PEF	-0.23	-0.22 (-0.27 – -0.17)

*: Across visit-specific correlations over time. FEV_1: forced expired volume in one second; PEF: peak expired flow. All correlations were significantly different from zero ($p < 0.005$).
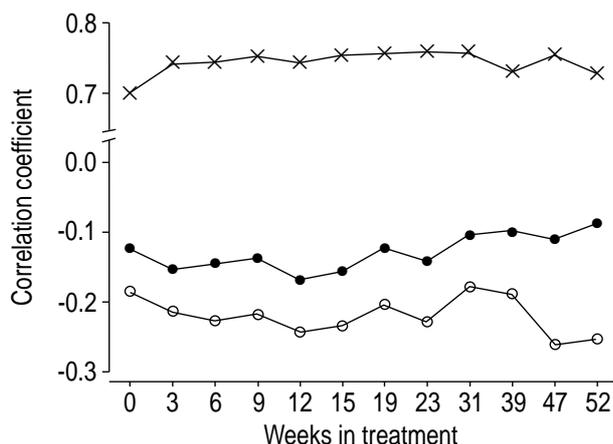
Fig. 1. – Correlation coefficients between forced expired volume in one second (FEV1) and patient-reported endpoints. ●: correlation between FEV1 and daytime symptom score; ○: correlation between FEV1 and "as-needed" β-agonist use; ×: correlation between FEV1 and patient-reported peak expiratory flow (PEF).

(FEV1, daytime symptom score) and "as-needed" β-agonist use, or between (FEV1, "as-needed" β-agonist use) and daytime symptom score were moderate (range: 0.53–0.66) (table 4).

Discussion

This report provides the largest cohort for correlation analyses among clinical trial endpoints. Previously, APTER and coworkers [1, 2] reported low correlations in a total of 34 subjects, among commonly used asthma efficacy parameters (daytime symptom score, PEF, and "as-needed" β-agonist use). JUNIPER et al. [10] and SANTANELLO et al. [8] reported low correlations between pulmonary function tests, quality of life and symptom scores in short-term observational studies. A result from a recent study [11] suggested that asthmatic patients' moods may be strongly affected by pulmonary function; thus patients' fluctuating moods may influence perception of asthma symptoms. Furthermore, ZHANG et al. [9] incorporated many of these measures in defining an asthma exacerbation from information in short term studies suggesting that the correlation among these endpoints was low. In the present report, based on data collected from 1576 patients in the two studies over a one-year period, it was observed that correlations among the endpoints were stable over time, and that daytime symptom score and "as-needed" β-agonist use had low correlations with both FEV1 and PEF. A strong correlation was
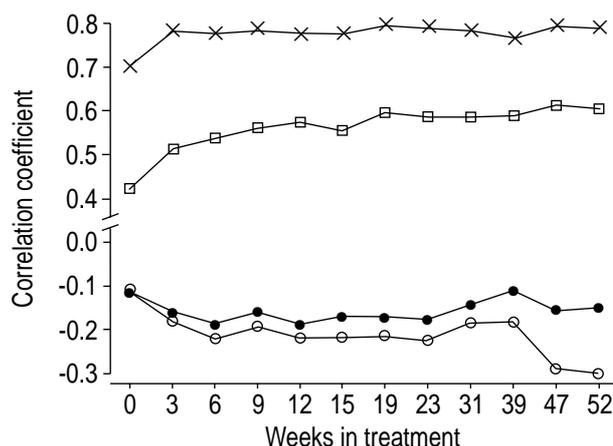

Fig. 2. – Correlation coefficients between measured-at-clinic peak expiratory flow (PEF) and patient-reported endpoints. ●: correlation between measured-at-clinic PEF and daytime symptom score ○: correlation between measured-at-clinic PEF and "as-needed" β-agonist use ×: correlation between measured-at-clinic PEF and patient-reported PEF □: correlation between daytime symptom score and "as-needed" β-agonist use.

observed between FEV1 and PEF, and between measured-at-clinic and patient-reported PEF. These findings suggest that FEV1 or PEF (measures of airway obstruction), and daytime symptom score and "as-needed" β-agonist use (measures of patient perception), assess different aspects of asthma. This was also supported by the canonical correlation analysis. The canonical coefficient between the two sets of endpoints, i.e., airway obstruction (measured at the clinic), patient-reported endpoints of daytime symptom score and "as-needed" β-agonist use, demonstrated low correlation. These low correlations among classes of surrogate asthma endpoints, support the concept that at least one endpoint from each category should be used in asthma clinical trials, e.g. FEV1 (or PEF) as a measure of airway obstruction, and daytime symptom or "as-needed" β-agonist use, as a measure of patient-perceived disease burden, because they measure different dimensions of asthma. Because the correlation between daytime symptom score and "as-needed" β-agonist use was only moderate, it could be argued that these endpoints are sufficiently different to justify their independent use as asthma endpoints in clinical trials.

These findings are consistent with and support the Global Initiative for Asthma Consensus Criteria [12] for asthma severity and control through the use of multiple parameters, including measures of airway function, patient symptoms and rescue medication use; this approach highlights the need to measure multiple parameters, including symptoms and exacerbation, in

Table 4. – Canonical correlation coefficients

	Mean
(FEV1, clinic-measured PEF) and (daytime symptom, "as-needed" β-agonist use)	0.202–0.269
(Patient-reported PEF) and (daytime symptom, "as-needed" β-agonist use)	0.213–0.274
(FEV1 and daytime symptom) and ("As-needed" β-agonist use)	0.556–0.660
(FEV1 "as-needed" β-agonist use) and (Daytime symptom)	0.531–0.663

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

addition to measurements of airway function, when assessing individual patients.

In conclusion, this analysis has, in a large patient cohort, demonstrated that surrogate asthma endpoints have a range of correlations, demonstrating the utility of their measurement in both clinical trials and in monitoring individual patient progress on therapy.

References

1. Apter AJ, ZuWallack RL, Clive J. Common measures of asthma severity lack association for describing its clinical course. *J Allergy Clin Immunol* 1994; 94: 732–737.
2. Apter AJ, Afflek G, Reisine ST, *et al.* Perception of airway obstruction in asthma: sequential daily analysis of symptoms, peak expiratory flow rate, and mood. *J Allergy Clin Immunol* 1997; 99: 605–612.
3. Shingo S, Zhang J, Reiss TF. Relationship between peripheral eosinophils and asthma efficacy end points: FEV1, PEFR and daytime symptom scores. Proceedings of The American Lung Association/American Thoracic Society International Conference; 1997 May 17–21, San Francisco, CA, USA. *Am J Respir Crit Care Med* 1997; 155 (4): A889.
4. Malmstrom K, Rodriguez-Gomez G, Guerra J, *et al.* Comparison of montelukast and inhaled beclomethasone in the treatment of chronic asthma: a randomized, placebo-controlled 12-week study. *Ann Intern Med* 1999; 130 (6): 487–495.
5. Reiss TF, Chervinsky P, Dockhorn RJ, Shingo S, Seidenberg BS, Edwards TB. Montelukast, a once-daily leukotriene receptor antagonist, in the treatment of chronic asthma, a multicenter, randomized, double-blind trial. *Arch Intern Med* 1998; 158 1213–1220.
6. American Thoracic Society. Standardization of spirometry: 1994 update. *Am J Respir Crit Care Med* 1995; 152 (3): 1107–1136.
7. Malmstrom K, Botto A, Zhang J, Reiss TF. Centralized spirometry quality control lowers the variability in multicenter asthma clinical trials. *Eur Respir J* 1996; 9: 273s, P0813.
8. Santanello NC, Barber BL, Reiss TF, Friedman BS, Juniper EF, Zhang J. Measurement characteristics of two asthma symptom diary scale for use in clinical trials. *Eur Respir J* 1997; 10: 646–651.
9. Zhang J, Song C, Reiss TF. Development and validation of a sensitive measure of worsening asthma. *Am J Resp Crit Care Med* 1997; 155: A892.
10. Juniper EF, Guyatt GH, Epstein RS, Ferrie PJ, Jaeschkl R, Miller TK. Evaluation of impairment of health-related quality of life in asthma: development of a questionnaire for use in clinical trials. *Thorax* 1992; 47 (2): 76–83.
11. Hyland ME. The mood-peak flow relationship in adult asthmatics: a pilot study in individual differences and direction of causality. *Br J Med Psychol* 1990; 63: 379–384.
12. National Institute of Health. Global Initiative for Asthma: Global strategy for asthma management and prevention. Bethesda, Maryland, USA. 1995; publication No. 95; p0813.