

Clinical trial efficacy: What does it really tell you?

Joseph Spahn, MD *Denver, Colo*

The primary goal of most clinical trials is an evaluation of the efficacy of the drug being evaluated. Therefore, it is important to understand if study outcomes are a true reflection of a drug's "real-life" effectiveness. Clinical trials generally evaluate three types of outcomes: subjective, objective, and health-related. Clinical trials that use subjective measures as endpoints usually evaluate outcomes such as symptom scores, the need for rescue medication, and quality-of-life measures. The majority of clinical trials rely on objective measures to test the efficacy of asthma medications. These include lung function (peak expiratory flow rate, forced expiratory volume at one second [FEV₁]), level of bronchial hyperresponsiveness (methacholine challenge or exercise challenge), and markers of inflammation (exhaled nitric oxide, sputum eosinophils, bronchoalveolar lavage, and bronchial biopsy). FEV₁ remains the gold standard of efficacy measures; however, in pediatrics, where FEV₁ values are often within the normal range, re-adjustment of what constitutes various levels of asthma severity should occur. Occasionally, studies will include an assessment of health outcomes, either as a primary or secondary measure, during the course of the study. Commonly measured health outcomes include a reduction in need for rescue systemic glucocorticoids, a reduction in the need for emergent asthma care, a reduction in asthma hospitalizations, and a reduction in asthma deaths. Studies designed to assess impact of treatment regimens on morbidity and mortality remain a high priority, as do studies designed to predict response to current asthma therapies. (*J Allergy Clin Immunol* 2003;112:S102-6.)

Key words: Clinical trial outcomes, lung function, bronchial hyperresponsiveness, FEV₁

Clinical research is vitally important for providing a greater understanding of the history, pathogenesis, and most effective therapies for asthma. Although the insight gained from clinical research comes from a wide range of trial types, some forms of clinical research carry more weight than others. In general, randomized, controlled trials with a rich body of data are considered the gold standard of clinical research. Randomized, controlled, trials with a limited body of data are usually ranked second, with nonrandomized and observational studies, and panel consensus guidelines following.

The process of evaluating a clinical research study raises multiple questions. First, reviewers need to deter-

Abbreviations used

- BHR: Bronchial hyperresponsiveness
- BIS: Budesonide inhalation suspension
- CHSA: Parent-reported symptoms, healthcare utilization, and functional health status
- FEV₁: Forced expiratory volume at one second
- FEV₁%: Percent-predicted FEV₁
- FP: Fluticasone propionate
- PEFR: Peak expiratory flow rate

mine what question is being addressed, and more importantly, how it is being addressed. Determining primary and secondary outcomes is important, as is evaluating the patient population to be sure it is appropriate for the objective of the study. Finally, reviewers need to decide if the study is powered appropriately and if the duration of the study is long enough to answer the question.

The primary message most clinicians want to take away from a study is the efficacy of the drug being evaluated. Therefore, it is important to understand if study outcomes are a true reflection of a drug's "real-life" effectiveness. Clinical trials generally evaluate three types of outcomes: subjective, objective, and health-related. The following sections will review the measures included in each of these outcomes, as well as the correlation between these outcomes and real-world efficacy.

SUBJECTIVE MEASURES

Clinical trials that use subjective measures as endpoints usually evaluate outcomes such as symptom scores, the need for rescue medication, and quality-of-life measures. Subjective measures are the primary endpoint used in most studies of asthma in very young children because objective measures such as forced expiratory volume at one second (FEV₁) or peak expiratory flow rate (PEFR) are difficult to perform in this patient population.¹ However, assessment of asthma symptoms and quality-of-life measures present an additional set of problems: researchers need to rely on caregivers to report asthma symptoms and impact.¹ The use of validated instruments can help assure reliable capture of subjective data. The following are examples of two placebo-controlled trials of asthma controller therapy in very young children; one study evaluated nebulized budesonide, and the other evaluated montelukast. Of importance, both were found to be superior to placebo. However, the placebo effect largely reinforces the need to perform placebo-controlled studies to determine the real effect of the drug studied, especially when objective measures can not be obtained.

From the National Jewish Medical and Research Center, and University of Colorado Health Sciences Center, Denver, Colorado.

Reprint requests: Joseph Spahn, MD, Associate Professor of Pediatrics, National Jewish Medical and Research Center, University of Colorado Health Sciences Center, 1400 Jackson St, Denver, CO 80206.

© 2003 American Academy of Allergy, Asthma and Immunology
0091-6749/2003 \$30.00 + 0
doi:10.1067/mai.2003.1855

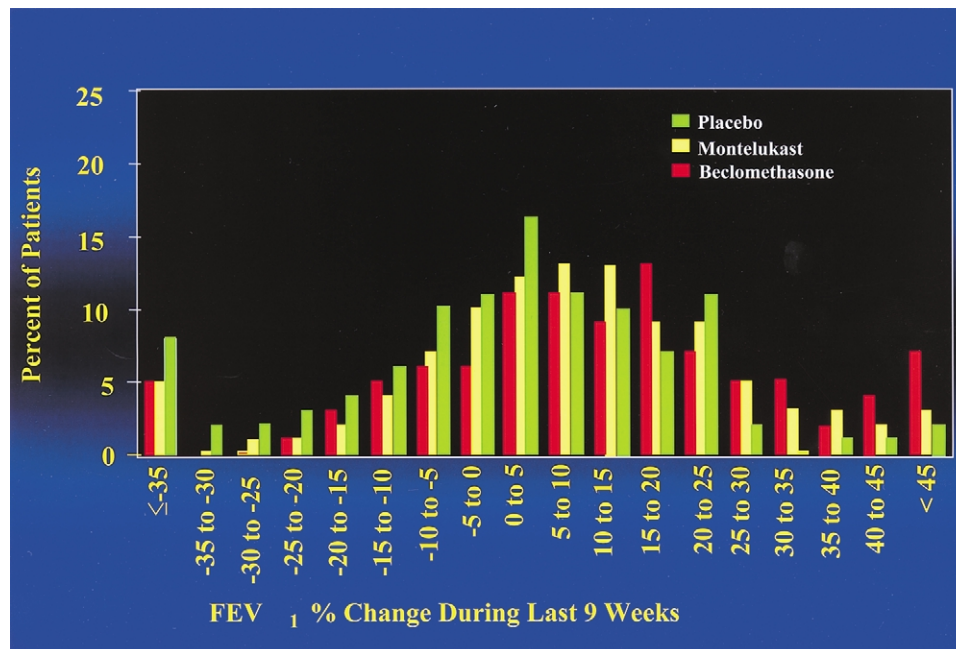


FIG 1. When data are presented as mean values, the variability in response is often lost. In a study by Zhang et al, the within-patient FEV₁ variability was 13.7%. Patients with a 10% increase in FEV₁ at 3 weeks had between a 50% decrease and a 75% increase in FEV₁ at week 12. (Adapted from Zhang et al, *Eur Resp J* 2002;20:1102-9.)

A double-blind, randomized, parallel-group study used asthma symptoms and use of rescue medication as endpoints to compare the efficacy and safety of four dosing regimens of budesonide inhalation suspension (BIS) or placebo in 480 asthmatic infants and children, ages 6 months to 8 years, with moderate persistent asthma.² Efficacy was assessed by twice-daily recording of asthma symptom scores at home and by use of rescue medication.² In addition, peak flow measurements were recorded twice daily and spirometry was recorded at clinic visits for those children able to perform these tests.²

Children enrolled in the study had a mean asthma symptom score of approximately 1.3 (scale, 0-3). All dosing regimens with BIS produced statistically significant improvement in daytime and night-time asthma symptom scores compared with placebo.² Separation between active treatment and placebo in daytime and night-time symptom scores were observed by week 2 of treatment for all BIS treatment regimens. In addition, children receiving BIS showed a significantly reduced number of days needing rescue medication compared with children receiving placebo.² A significant increase in peak flow measurement was observed in most active treatment groups compared with placebo in the subset of children able to do pulmonary function testing. All treatment groups showed numerical improvement in FEV₁.²

In another example, a double-blind, multicenter, multinational study at 93 centers worldwide evaluated the effects of montelukast on daytime and overnight asthma symptoms, daily use of β -agonist, days without

asthma, frequency of asthma attacks, number of patients discontinued because of asthma, need for rescue medication, physician and caregiver global evaluations of change, asthma-specific caregiver quality of life, and peripheral blood eosinophil counts.¹ During the study, 689 children aged 2 to 5 years were randomly assigned to 12 weeks of treatment with placebo (228 patients) or 4 mg of montelukast as a chewable tablet (461 patients).¹ Montelukast produced significant improvements compared with placebo in multiple parameters of asthma control including daytime asthma symptoms, overnight asthma symptoms, the percentage of days with asthma symptoms, the percentage of days without asthma, the need for β -agonist or oral corticosteroids, physician global evaluations, and peripheral blood eosinophils.¹ The clinical benefit of montelukast was evident within 1 day of starting therapy. Caregiver global evaluations, the percentage of patients having asthma attacks, and improvements in quality-of-life scores favored montelukast, but were not significantly different from placebo.¹

OBJECTIVE MEASURES

The majority of clinical trials rely on objective measures to test the efficacy of asthma medications. These include lung function (PEFR, FEV₁), level of bronchial hyperresponsiveness (methacholine challenge or exercise challenge), and markers of inflammation (exhaled nitric oxide, sputum eosinophils, bronchoalveolar lavage, and bronchial biopsy).

Lung function (PEFR, FEV₁)

The vast majority of studies require children to have an FEV₁ of 60% to 80% of predicted as a prerequisite for enrollment. However, most children in clinical practice with asthma have FEV₁ values within the normal range (80% to >100% predicted).³ Arguments for using children with relatively low lung function include that the Food and Drug Administration mandates using FEV₁ as a primary outcome measure, and if children with normal lung function are studied, one may not be able to demonstrate a significant improvement in FEV₁ compared with placebo, or to demonstrate a difference between two active treatments.

Despite this issue, FEV₁ has become the "gold-standard" measure for assessing response to controller therapies. The change in pre- β -agonist FEV₁ is standard, although one of the leading cohort studies of children with asthma, the Childhood Asthma Management Program (CAMP), used post- β -agonist FEV₁ as its primary outcome.⁴ Medications that improve baseline FEV₁ values, in general, are also associated with reductions in symptoms, need for short-acting β -agonists, reductions in bronchial hyper-responsiveness (BHR), and reductions in airway inflammation. Therefore most clinical trials use FEV₁ as a primary endpoint.

Twenty asthmatics with mild-to-moderate disease were included in a single-blind study evaluating the effectiveness of fluticasone propionate (FP).⁵ Baseline data were compared with those from 26 normal subjects. High-dose treatment with FP resulted in significant improvements in both lung function, as measured by FEV₁, and BHR.⁵ A reduction in airway inflammatory cells after treatment with FP suggested that improved lung function is associated with a reduction in the inflammatory process of asthma.⁵

Fuhlbrigge et al³ explored the relationship between the percent-predicted FEV₁ (FEV₁%) and subsequent asthma attacks in a longitudinal study of pediatric lung health. Data from a retrospective cohort of 13,842 children seen annually over a 15-year interval were analyzed for measurements of pulmonary function. In addition, a standard questionnaire was completed by a parent or guardian until the child reached ninth grade; then the survey was completed by the child.³ For each observation, the report of an attack during the past year was paired with FEV₁ recorded at the field survey one year earlier. Results showed that a progressive decrease in the proportion of persons reporting an attack was associated with increasing FEV₁% predicted.³ The authors concluded that the strong association between FEV₁% predicted and risk of asthma attack over the subsequent year supports an emphasis on objective measures of lung function in assessment of risk for adverse asthma outcomes.³

There are some issues with using FEV₁ as a primary variable. Clinical trials often present mean data. In general, the mean FEV₁ response improves approximately 10% to 15% with inhaled corticosteroids and 5% to 7.5% with leukotriene receptor antagonists. When data are presented as mean values, the variability of response is often

lost. Two recent studies have demonstrated a variability in response to treatment (as measured by FEV₁) and a unimodal distribution for all treatment groups (Fig 1).^{6,7} This suggests that patients cannot simply be divided into distinct responder and nonresponder treatment groups. Instead, patient response to treatment is somewhat unique to that particular patient.

Sharek et al⁸ sought to assess relationships between multiple parameters of asthma status in a cohort of inner-city children. They looked at parent-reported symptoms, healthcare utilization, functional health status (CHSA), diary card data, and spirometry at baseline, 32, and 52 weeks. Parent-reported symptoms, parent-reported health care utilization, and CHSA correlated with each other at baseline and throughout the study. They also found that FEV₁, PEFR, and diary card data did not correlate with other measures at baseline, and were inconsistently correlated to other parameters during the 52 weeks.⁸ They concluded that there is a clear dissociation between lung function and other asthma status measures, and that clinical trials should be designed to match interventions with outcomes that best measure effectiveness.⁸

A final problem with using FEV₁ as a primary endpoint is measuring efficacy in patients with normal lung function. Fuhlbrigge et al³ found that the majority of lung function measures were within the normal range. Furthermore, having a normal FEV₁ did not protect a patient from having subsequent asthma attacks. In the Childhood Asthma Management Program study, the investigators found that long-term budesonide therapy improved baseline FEV₁, but the change was clinically insignificant (3%) and it failed to improve post- β -agonist FEV₁.⁴ On the other hand, budesonide was effective in reducing BHR, the need for rescue β -agonist and symptoms, and was associated with significant reductions in hospitalizations (43%), urgent care visits (45%), and need for rescue prednisone (43%). This study clearly demonstrates the dissociation of lung function and other clinical measures, in that significant improvements in nearly every clinical parameter were noted during long-term inhaled glucocorticoid therapy in the absence of any meaningful change in lung function.

What these studies suggest is that there is a significant within-and-between-patient variability in FEV₁ values regardless of therapy. Also, the correlations between baseline FEV₁ values or changes in FEV₁ and changes in symptoms, quality-of-life measures, and patient global evaluations are weak. Therefore, to truly evaluate efficacy, FEV₁ alone as an endpoint is not sufficient.

Level of BHR

Bronchial hyperresponsiveness is associated with asthma severity, airway inflammation, and basement membrane thickness. Medications that decrease BHR are effective asthma medications. Therefore, BHR is occasionally used as an endpoint in asthma efficacy studies. However, it is possible to have BHR and not have asthma. Furthermore, patients in asthma remission will often continue to display symptoms of BHR. An important

restriction to measuring BHR in clinical studies is that methacholine challenges are time-consuming and require expertise to perform.

Markers of inflammation

Because asthma is an inflammatory disease, endpoints that measure levels of inflammation can be useful in clinical trials. Exhaled nitric oxide is easily measured and is noninvasive. In addition, levels of nitric oxide correlate with several clinical parameters and, therefore, may be a predictor for response to therapy. However, these correlations are modest at best. Therefore, using nitric oxide exhalation as the only endpoint is not a practical measure of efficacy.

Sputum can be obtained and evaluated for the presence of airway inflammatory cells and mediators. As with nitric oxide exhalation, this is a noninvasive, direct measurement of airway inflammation. Also, eosinophilia/eosinophil cationic protein is less affected by inhaled corticosteroid therapy and can be utilized in these studies. However, this evaluation is time-consuming (2-3 hours), and analysis is difficult to perform and interpret. Also, some patients cannot produce sufficient sputum. Finally, safety may be an issue, because the procedure to gather sputum can induce bronchospasm despite β -agonist treatment.

HEALTH OUTCOMES

Occasionally, studies will include an assessment of health outcomes, either as a primary or secondary measure, during the course of the study. Although these measures alone cannot be used to determine efficacy, they can further support real-world efficacy of a drug that has already undergone trials using objective outcomes. Commonly measured health outcomes include a reduction in need for rescue systemic glucocorticoids, a reduction in need for emergent asthma care, a reduction in asthma hospitalizations, and a reduction in asthma deaths.

Approximately 2 million emergency department visits and 500,000 hospitalizations are attributed to asthma each year. A major goal for asthma research is to reduce health care utilization in relation to acute asthma. To accurately measure the effect asthma medications have on emergency department visits and hospitalizations, long-term studies with large cohorts are required.

Donahue et al⁹ attempted to determine if anti-inflammatory treatment for asthma reduces the risk of asthma hospitalization in a retrospective cohort study. Of the 16,941 eligible persons, 742 (4.4%) were hospitalized for asthma. The overall relative risk (RR) of hospitalization among those who received inhaled corticosteroids was 0.5 (95% CI, 0.4-0.6) after adjustment for β -agonist dispensing.⁹ Additional adjustment for age, race, other asthma medications, and amount and type of ambulatory care for asthma did not substantially affect the inverse relationship between use of inhaled corticosteroids and hospitalization. Cromolyn was similarly associated with reduced risk, especially among children (RR, 0.8; 95%

CI, 0.7-0.9).⁹ In contrast, increasing β -agonist use was associated with increasing hospitalization risk, even after adjustment for other factors and medications.⁹ The steroid-associated protection was most marked among persons who received the largest amount of β -agonist. The authors concluded that inhaled corticosteroids and, to a lesser extent, cromolyn, confer significant protection against exacerbations of asthma leading to hospitalization.⁹ These results support the use of inhaled corticosteroids by persons who require more than occasional β -agonist use to control asthma symptoms.

Suissa et al¹⁰ assessed the efficacy of inhaled corticosteroids in preventing death from asthma. Using the Saskatchewan Health databases to form a cohort consisting of 30,569 subjects, they found that of the 562 reported deaths, 77 were classified as the result of asthma. After matching the 66 subjects who died of asthma for whom there were complete data with 2681 controls, they found that 53% of the case patients and 46% of the control patients had used inhaled corticosteroids in the previous year, most commonly low-dose beclomethasone.¹⁰ The rate of death from asthma during the first three months after discontinuation of inhaled corticosteroids was higher than the rate among patients who continued to use the drugs. On the basis of the results, the authors concluded that regular use of low-dose inhaled corticosteroids is associated with a decreased risk of death from asthma.¹⁰

CONCLUSIONS

Asthma is a complex disease. Just as multiple genes are involved in asthma pathogenesis, there is no single clinical parameter that provides a global assessment of disease severity or response to treatment. However, FEV₁ remains the gold standard by historic precedent and because it provides an objective, relatively easily performed measure of airflow obstruction.

In pediatrics, where FEV₁ values are often within normal values, re-adjustment of what constitutes various levels of asthma severity should occur. For example, the FEV₁/forced vital capacity ratio $\pm$ β -agonist-reversibility, independent of baseline FEV₁ could be utilized as entry criteria for clinical research studies.

Finally, studies designed to assess impact of treatment regimens on morbidity and mortality remain a high priority, as do studies designed to predict response to current asthma therapies.

I thank Maria Bavishi for her assistance in preparing and editing this manuscript.

REFERENCES

1. Knorr B, Franchi LM, Bisgaard H, Vermeulen JH, LeSouef P, Santanello N, et al. Montelukast, a leukotriene receptor antagonist, for the treatment of persistent asthma in children aged 2 to 5 years. *Pediatrics* 2001;108:E48.
2. Baker JW, Mellon M, Wald J, Welch M, Cruz-Rivera M, Walton-Bowen K. A multiple-dosing, placebo-controlled study of budesonide inhalation suspension given once or twice daily for treatment of persistent asthma in young children and infants. *Pediatrics* 1999;103:414-21.

3. Fuhlbrigge AL, Kitch BT, Paltiel AD, Kuntz KM, Neumann PJ, Dockery DW, et al. FEV(1) is associated with risk of asthma attacks in a pediatric population. *J Allergy Clin Immunol* 2001;107:61-7.
4. Zeiger RS, Dawson C, Weiss S. Relationships between duration of asthma and asthma severity among children in the Childhood Asthma Management Program (CAMP). *J Allergy Clin Immunol* 1999;103:376-87.
5. Booth H, Richmond I, Ward C, Gardiner PV, Harkawat R, Walters EH. Effect of high dose inhaled fluticasone propionate on airway inflammation in asthma. *Am J Respir Crit Care Med* 1995;152:45-52.
6. Zhang J, Yu C, Holgate ST, Reiss TF. Variability and lack of predictive ability of asthma end-points in clinical trials. *Eur Respir J* 2002;20:1102-9.
7. Malmstrom K, Rodriguez-Gomez G, Guerra J, Villaran C, Pineiro A, Wei LX, et al. Oral montelukast, inhaled beclomethasone, and placebo for chronic asthma. A randomized, controlled trial. Montelukast/Beclomethasone Study Group. *Ann Intern Med* 1999;130:487-95.
8. Sharek PJ, Mayer ML, Loewy L, Robinson TN, Shames RS, Umetsu DT, et al. Agreement among measures of asthma status: a prospective study of low-income children with moderate to severe asthma. *Pediatrics* 2002;110:797-804.
9. Donahue JG, Weiss ST, Livingston JM, Goetsch MA, Greineder DK, Platt R. Inhaled steroids and the risk of hospitalization for asthma. *JAMA* 1997;277:887-91.
10. Suissa S, Ernst P, Benayoun S, Baltzan M, Cai B. Low-dose inhaled corticosteroids and the prevention of death from asthma. *N Engl J Med* 2000;343:332-6.