

Defining the responder in asthma therapy

David A. Stempel, MD,^a and Anne L. Fuhlbrigge, MD, MS^b *Seattle, Wash, and Boston, Mass*

Key words: Asthma, responder, inhaled corticosteroids, leukotriene receptor antagonists, statistical significance, effect size

In last month's Journal, Szeftler et al¹ presented an article from the National Heart, Lung, and Blood Institute Childhood Asthma Research and Education (CARE) Network on the characterization of response to 2 common therapies for childhood asthma. The article establishes a definition of response as the percent change in prebronchodilator FEV₁ $\geq$ 7.5%. The importance of the article is that it selected a working definition of *response* and then attempted to assess whether predictors of response could be determined. The article by Szeftler et al¹ takes a systematic look at one definition of response and serves as an example for future discussion of how one characterizes a responder.

A limitation in the interpretation of the asthma literature is the inconsistency in the definition of *response*. The heterogeneity and variability of asthma make it difficult to achieve consensus for this term. This is compounded by the lack of a standardized means to assess asthma control, a measure of response. Should response to an intervention be defined primarily by objective parameters, subjective symptom reports, or a composite measure? Is the observed response to an intervention clinically as well as statistically significant, and how has the choice of primary outcome and sample size influenced these results? In addition, which parameters should be more appropriately thought of as markers of disease severity than measures of response? If these issues are difficult for the specialist and researcher to resolve, how can we provide guidance to our primary care colleagues when they read the asthma literature? The definition of *responder* is increasingly important as the number of assessment tools multiplies beyond traditional clinical trials to include observational investigations, studies of patients with specific phenotypes

Abbreviations used

ACRN: Asthma Care Research Network
 AHR: Airway hyperresponsiveness
 CAMP: Childhood Asthma Management Program
 CARE: Childhood Asthma Research and Education
 eNO: Exhaled nitric oxide
 PEF: Peak expiratory flow
 START: Inhaled steroid treatment as regular therapy in early asthma study

and genotypes, and an increase in the number of surrogate markers of disease activity. This commentary is not a critique of the literature but rather is intended to encourage reflection on this issue.

Medication selection for the treatment of asthma should be based on the clinical evidence that demonstrates the most favorable outcomes to a specific therapeutic. Recommendations for treatment have been based on guidelines developed by expert panels, evidence-based systematic reviews of the literature, and frequently the combination of these 2 processes. Ideally, one would like to customize therapy to the specific patient. However, when clinicians review the literature to form their treatment choices, what is not standardized is the definition of *response*. Clinical studies commonly select a primary outcome, and then the investigation is appropriately powered to study the null hypothesis; is the observed effect on the primary outcome related to the intervention, or is it possibly a random occurrence? However, there are often secondary outcomes within a clinical trial that may be overpowered or underpowered based on the number needed to examine the primary outcome.

The Childhood Asthma Management Program (CAMP) illustrates the implications of the choice of a primary outcome measure when assessing the response to therapy.² CAMP was a randomized clinical trial examining the effect of inhaled corticosteroids or nedocromil compared with placebo on pulmonary growth and development. Postbronchodilator "FEV₁ was chosen as the primary outcome measure because it was widely accepted as the most clinically useful and predictive measure of lung function."² In the final analysis, the investigators found no significant change in postbronchodilator FEV₁ in response to therapy, and the null hypothesis was accepted. However, other clinical outcomes were examined as a part

From ^athe Infomed Northwest and University of Washington; and ^bthe Channing Laboratory, Brigham and Women's Hospital, Harvard Medical School, Boston.

Disclosure of potential conflict of interest: David A. Stempel serves as a consultant and on the speakers' bureau for GlaxoSmithKline. Anne L. Fuhlbrigge serves as a consultant for GlaxoSmithKline and has received research grant support from Boehringer Ingelheim.

Received for publication November 22, 2004; revised December 12, 2004; accepted for publication December 13, 2004.

Reprint requests: David A. Stempel, MD. E-mail: econmed@msn.com. J Allergy Clin Immunol 2005;115:466-9.

0091-6749/\$30.00

© 2005 American Academy of Allergy, Asthma and Immunology
 doi:10.1016/j.jaci.2004.12.1113

of the trial. Budesonide compared with placebo for the treatment of asthma significantly reduced hospitalizations, urgent care visits, oral corticosteroid use, symptom scores, episode-free days, and need for rescue albuterol. In addition, bronchial hyperreactivity improved. Based on these results, are the patients treated with budesonide considered responders or nonresponders? For the primary outcome, this is a negative study. Yet the 40% to 45% reduction of exacerbations with budesonide is both statistically significant and clinically relevant. The absence of improvement in postbronchodilator FEV₁ by itself does not imply that the patients, either individually or collectively, are nonresponders.

The study of the early intervention with budesonide in mild persistent asthma (START)³ demonstrated similar benefits with reduced exacerbations. This “study was designed to detect a 35% reduction in the hazard of having a severe asthma-related event with a 95% power...”³ The investigators reported that budesonide reduced the risk of the first severe asthma-related event by 44% for all patients (hazard ratio, 0.56; CI, 0.45–0.71). Thus, they were able to reject the null hypothesis for the primary endpoint; budesonide had a significant effect on exacerbation rates. However, because these severe events are uncommon, the study required a large sample size to achieve adequate power. The investigators also examined secondary outcomes, including FEV₁, and noted a significant difference in the rate of decline in FEV₁ in the budesonide-treated group ($P = .0005$). However, the absolute value of this difference was small, 0.88%. The observation that budesonide slowed the natural decline in lung function over time illustrates how, for this parameter, the study may have been overpowered. Although statistically significant, the observed difference in rate of decline of FEV₁ may have limited clinical value, because it is within the range of the reproducibility for spirometry.

On initial review, the CAMP² trial and the START³ study have conflicting results. CAMP observed no significant change in post bronchodilator FEV₁ associated with budesonide, whereas the START trial reported a significant change in the decline in FEV₁ over time. There were 7241 subjects enrolled in the START trial, contrasted with the 1041 subjects in CAMP. This may explain why the similar effect size in lung function produced statistical significance in one study but not the other. In fact, both the lung function results and the exacerbation data of CAMP and START are comparable, and the 2 studies had similar baseline lung functions. However, because of the choice of primary outcome, START required a larger study population and was able to achieve statistical significance for a secondary outcome, FEV₁.

In the current article from the CARE network,¹ the mean response to therapy was 6.8% and 1.9% change in FEV₁ with fluticasone and montelukast, respectively. For both treatment arms, the mean change in lung function was below the threshold chosen by the authors in their definition for response. Therefore, by definition, the majority of patients in the trial were not considered responders. But is FEV₁ by itself an appropriate working

definition of response? As a secondary outcome, the article reported that 12 patients had exacerbations requiring systemic corticosteroids: 10 subjects assigned to montelukast and 2 on fluticasone. This current article, although with shorter duration and fewer subjects, has findings consistent with CAMP and START demonstrating limited lung function improvement but significant reductions in exacerbations. Reduction in exacerbations, although infrequent in some patients, is a goal of guideline-directed asthma care and should be integrated into the definition of response. The relative weight of each outcome parameter needs further characterization.

An interesting observation in many placebo-controlled trials is that outcomes of interest are normally distributed within the treatment arms of a trial. This is illustrated in the publication of Zhang et al.⁴ The majority of placebo patients showed little change in lung function or symptom control, but a modest subgroup had noteworthy improvement and a smaller group clinical deterioration over the course of the trial. A similar distribution curve for response was observed in subjects treated with the active medications, with the curve moving in the direction of improvement. This distribution in the observed response to placebo may be the consequence in part of the natural course of the disease. For example, patients with asthma entering a study may be recovering from a viral illness. Dramatic improvements in FEV₁ of 20% to 40% may be noted as the illness resolves, even in those patients treated with placebo. In contrast, patients may be infected with a common virus during the study and have equivalent deterioration of this outcome parameter. The statistical analysis in accepting or rejecting the null hypothesis accommodates for these events that are not related to the study intervention.

How response is defined can alter the interpretation of the results of clinical trials. The Asthma Care Research Network (ACRN) has published a study on the variability of response to inhaled corticosteroids.⁵ This study addresses several issues related to the definition of response in asthma treatment. In this randomized, open trial, patients were placed on inhaled corticosteroids. The primary outcome was FEV₁, and secondary outcomes included methacholine PC₂₀, exercise-induced fall in FEV₁, exhaled nitric oxide (eNO), and sputum eosinophils. A good FEV₁ response to the study medication was defined as greater than 15%, a marginal response 5% to 15%, and poor response less than 5%. In contrast, the START study reported that a 2.24% improvement in prebronchodilator FEV₁ was significant. If one utilizes the ACRN definition of FEV₁ response in the START trial, although statistically significant, the changes reported would be considered poor.

The ACRN study also focused on factors, based on the patient phenotype or baseline clinical status of the subject, that predict response to a specific therapy. Thirteen of the 21 patients on inhaled corticosteroids had a good to marginal FEV₁ response.⁵ A good response was correlated with higher median eNO, greater median reversibility, and lower baseline FEV₁/forced vital capacity ratio. Similarly,

the report from the CARE Network noted that the responders had at baseline lower prebronchodilator FEV₁%, lower prebronchodilator FEV₁/forced vital capacity, lower methacholine PC₂₀, higher eosinophilic cationic protein, and higher IgE.¹ However, these parameters identified in both studies are also measures of disease activity, which implies that disease activity may predict the effect size of the response. If the disease is well controlled, limited response in lung function will be noted; if the disease activity is increased, the response to treatment will be greater.

In addition, the ACRN investigators⁵ demonstrated that 21 of the 26 subjects (81%) had a good to excellent response to inhaled corticosteroids when assessed by methacholine PC₂₀. Again, the pattern of response was related to markers of disease activity: higher levels of sputum eosinophils, lower baseline PC₂₀, and higher eNO. Should subjects who did not improve their FEV₁ but had improvement in PC₂₀ be considered responders or not? Differences in the parameter selected, the effect size of the outcome measured, or the asthma phenotype can modify which patients are considered responders.

The study by Reddel et al⁶ illustrates another central determinant in the definition of response: differences in maximal response may vary as a function of time. If a composite measure of response is used, response can be a function of how the composite measure is defined. Reddel et al⁶ noted initial improvements in AM peak expiratory flow (PEF) in the first days of treatment with budesonide, with maximal improvement in laboratory FEV₁ and quality of life assessments seen after 8 weeks of therapy. However, improvement in FEV₁ and PEF continued through weeks 12 to 14 of the study, and PD₂₀ was noted to improve up to conclusion of the study at week 72. The authors⁶ proposed a composite measure of response that included symptoms, nighttime awakening, β_2 -agonist use, PEF variation, and airway hyperreactivity. Whether subjects achieved this composite measure of response was a function of the length of treatment as well as whether one or multiple parameters were included in the definition of *responder*.

Another recent publication defined a responder with the use of a composite score.⁷ Seven parameters (daytime symptoms, rescue β_2 -agonist use, AM PEF, nighttime awakening, exacerbations, emergency visits, and adverse effects) were equally weighted to determine the level of control. They observed that a composite definition of a responder correlates with asthma control. The use of composite scores is potentially an important method of defining response. However, consensus still needs to be developed on which parameters should be included. Furthermore, many questions remain: should response be defined as improvement to any, some, or all of the parameters? Does response need to meet a minimal threshold, be a substantial change, or be a complete normalization? Do all outcomes merit the same weight? Should uncommon events like exacerbations have an equal or greater weight than symptoms? Should weighted symptoms and exacerbations have greater value than

objective measures such as lung function or eNO? How should the length of therapy be factored in? A limitation of composite scores is that the effect size of each parameter needs to be defined.

Two groups of investigators have proposed that biomarkers of inflammation may be important parameters to follow in the management of asthma. Sont et al⁸ demonstrated that airway hyperresponsiveness (AHR) could be used as a guide to alter the dosage of inhaled corticosteroid. In their study, patients treated with a management protocol based on AHR noted a significant decrease in the incidence of mild exacerbations (1.8-fold) and improved prebronchodilator FEV₁ compared with a group of patients on a guideline-based reference strategy. To achieve this level of improvement, the AHR group required nearly twice as much inhaled corticosteroid. Of interest, despite a reduced thickness of the subepithelial reticular layer among subjects treated with the AHR protocol, the authors failed to demonstrate a difference in the inflammatory cell numbers in the bronchial biopsies between the 2 treatment groups.

Green et al⁹ used a similar study design to compare subjects whose inhaled corticosteroid dose was adjusted based on sputum eosinophil counts to subjects whose therapy was adjusted according to a protocol based on the National Asthma Education and Prevention Program guidelines. Despite similar doses of inhaled corticosteroids in the 2 groups, subjects whose therapy was adjusted by their sputum eosinophils had fewer severe exacerbations, less use of oral corticosteroids, and fewer hospitalizations. However, there was no difference in symptom scores, total asthma quality of life scores, lung function, and use of rescue albuterol. These 2 studies^{8,9} demonstrate potential role of specific inflammatory markers in asthma disease management but also illustrate the limitation of these parameters in predicting all outcome measures.

The objective that the CARE Network¹ established in their recent publication¹ of identifying indicators of response is formidable and builds a platform for future research. Consensus still needs to be established on the specific parameters to include in the definition of *responder*. The present literature suggests that as baseline parameters approach normal, the effect size of the response diminishes. Failing to respond to one parameter such as lung function does not preclude improvement in other areas such as reduced exacerbations or increased asthma-free days. Osborne et al¹⁰ from Oregon Health Sciences have proposed a methodology for weighting outcome measures in predicting acute care in patients with asthma. This methodology may allow development of a composite response score. Finally, one needs to contemplate how to integrate the effect of a ubiquitous trigger altering a good response into a poor response. This effect of the variable nature of the disease and the influence of confounding variables on response reinforce why studies report mean data. Mean data eliminate the observational bias when the effect is not a direct result of the intervention. Predicting response to a specific therapy

is important, but first, consensus must be reached on how to define response to therapy.

The national guidelines¹¹ classify asthma based on combinations of physiologic measures (pulmonary function testing), symptoms/functional limitation, medication use, and morbidity (exacerbations). It acknowledges that patients may have 1, some, or all of these indicators for treatment. Although response to therapy can be defined as the elimination or amelioration of any of these parameters associated with disease activity, the clinical importance as well as the statistical significance of the findings need to be considered. In addition, emphasis needs to be placed on the consistency with which we choose and report the outcomes of our investigations. Attention now should focus on the development of composite measures that adjust for baseline parameters and attempt to weigh the contribution of each variable. Reporting of these measures should allow comparison across studies for easier integration of information into practice as it becomes available.

REFERENCES

1. Szeffler SJ, Phillips BR, Martinez FD, Chinchilli VM, Lemanske RF Jr, Strunk RC, et al. Characterization of within-subject responses to fluticasone and montelukast in childhood asthma. *J Allergy Clin Immunol* 2005;115:233-42.
2. The Childhood Asthma Management Program Research Group. The long-term effects of budesonide or nedocromil in children with asthma. *N Engl J Med* 2000;343:1054-63.
3. Pauwels RA, Pederson S, Busse WW, Tan WC, Chen YZ, Ohlsson SV, et al. Early intervention with budesonide in mild persistent asthma: a randomized double-blind trial. *Lancet* 2003;361:1071-6.
4. Zhang J, Yu C, Holgate ST, Reiss TF. Variability and lack of predictive ability of asthma and end-points in clinical trials. *Eur Respir J* 2002;20:1102-9.
5. Szeffler SJ, Martin RJ, King TS, Boushey HA, Cherniack RM, Chinchilli VM, et al. Significant variability in response to inhaled corticosteroids for persistent asthma. *J Allergy Clin Immunol* 2002;109:410-8.
6. Reddel HK, Jenkins CR, Marks GB, Ware SI, Xuan W, Salome CM, et al. Optimal asthma control, starting with high doses of inhaled budesonide. *Eur Respir J* 2000;16:226-35.
7. Bateman ED, Boushey HA, Bousquet J, Busse WW, Clark TJH, Pauwels RA, et al. Can guideline-defined asthma control be achieved? the gaining optimal asthma control study. *Am J Respir Crit Care Med* 2004;170:836-44.
8. Sont JK, Willems LNA, Bel EH, van Krieken JHJM, Vandenbroucke JP, Sterk PJ, et al. Clinical control and histopathologic outcomes of asthma when using airway hyperresponsiveness as an additional guide to long-term treatment. *Am J Respir Crit Care Med* 1999;159:1043-51.
9. Green RH, Brightling CE, McKenna S, Hargadon B, Parker D, Bradding P, et al. Asthma exacerbations and sputum eosinophils counts: a randomized controlled trial. *Lancet* 2002;360:1715-21.
10. Osborne ML, David MR, O'Hollaren MT, Buist AS, Vollmer WN. Predictors of acute care in adult asthmatics: a prospective study [abstract]. *Am J Respir Crit Care Med* 2004;169:A529.
11. Guidelines for the Diagnosis and Management of Asthma. NIH publication #97-4051. Bethesda (MD): US Department of Health and Human Services, National Institutes of Health; 1997.

Receive tables of contents by e-mail

To receive the tables of contents by e-mail, sign up through our Web site at

<http://www.mosby.com/jaci>

Choose *E-mail Notification*

Simply type your e-mail address in the box and click the *Subscribe* button

Alternatively, you may send an e-mail message to majordomo@mosby.com.

Leave the subject line blank and type the following as the body of your message:

subscribe jaci_toc

You will receive an e-mail message confirming that you have been added to the mailing list.

Note that TOC e-mails will be sent out when a new issue is posted to the Web site.