CORRESPONDENCE

Distribution of therapeutic response in asthma control

To the Editor:

The paper by BAUMGARTNER *et al.* [1] raises very important questions pertaining to the interpretation and usefulness of randomised clinical trials, the use of surrogates of little relevance to the patient and the comparison of results using group average values. However, these problems are not adequately addressed.

Whilst reference is given to the Global Initiative for Asthma guidelines to define asthma control, the focus of the study is on percentage of asthma control days defined on a subset of asthma control measures only. This definition carries two weaknesses. First, it ignores important parameters that define asthma control. Lung function, including level of bronchial obstruction and its variability, is an important parameter to account for, so that days with marked obstruction are not misclassified. Secondly, and more importantly, the use of percentage of days ignores the fact that only sustained daily control defines overall control. In this study, $\leqslant 30\%$ of patients had control for $\geqslant 90\%$ of days. In our definition of asthma control we used much more stringent criteria that incorporated all parameters and integrated these over 3 months to provide an overall assessment [2]. We showed that one of the main reasons for not achieving control was bronchial obstruction and that this approach was meaningful to patients, as good control correlates with an improved or even normalised quality of life [3].

The use of distribution analysis to compare therapies is an interesting approach as it allows an assessment of treatment efficacy on a patient basis. It is important to question why the mean percentage of asthma control days showed a clear and significant difference between active treatments but the distribution analysis did not. Examination of the distribution data suggests a simple shift to the right with beclomethasone compared with montelukast treatment. About 5–10% of the patients are shifted from one column to the next, therefore ending with little or no difference for most columns and a clear difference at the extremes. As expected, this difference looks similar to the mean treatment difference. Therefore, one can speculate that most patients improve more with beclomethasone than they would with montelukast treatment. The only design that allows the within-patient difference to be addressed is the crossover design, and its sophisticated version, N of 1 [4]. The use of crossover studies in asthma is a challenge, but we should also recognise that the approach proposed by BAUMGARTNER *et al.* [1] is not suitable for parallel group designs.

It is not surprising that an equivalence testing using an 80% overlap failed to show a true difference. Assay sensitivity is an important issue in equivalence trials [5]. Equivalence boundaries have not been unequivocally established. Experience from development of alternative propellant programmes, general publications [6] and regulatory recommendations [7] suggest a zone of equivalence of ± 0.5 or a third of the difference between active and placebo treatment. This placebo-controlled study provides an opportunity to ascertain assay sensitivity as well as validate the boundaries chosen, but these critical results are not presented.

Using traditional measures of asthma, the study by BAUMGARTNER *et al.* [1] showed significant differences in favour of beclomethasone in several variables and a trend in all others, suggesting a true finding. The proposed new analysis method suggests the two treatments are equivalent. It is an illustration that the use of inappropriate methodologies can lead to misleading conclusions.

G.L. Braunstein
194 Boston Manor Road, Brentford, Middlesex, UK.

References

1. Baumgartner RA, Martinez G, Edelman JM, *et al.* Distribution of therapeutic response in asthma control between oral montelukast and inhaled beclomethasone. *Eur Respir J* 2003; 21: 123–128.
2. Bateman ED, Bousquet J, Braunstein GL. Is overall asthma control being achieved? A hypothesis-generating study. *Eur Respir J* 2001; 17: 589–595.
3. Bateman ED, Frith LF, Braunstein GL. Achieving guideline-based asthma control: does the patient benefit? *Eur Respir J* 2002; 20: 588–595.
4. Guyatt G, Sackett D, Taylor DW, Chong J, Robert R, Pugsley S. Determining optimal therapy - Randomized trials in individual patients. *N Engl J Med* 1986; 314: 889–892.
5. International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use - ICH Harmonised Tripartite Guidelines E9: Choice of control group and related issues in clinical trials. www.ich.org. Date updated: July 2000; Date last accessed: June 28 2003.
6. Hwang IK, Morikawa T. Design issues in noninferiority/equivalence trials. *Drug Inf J* 1999; 33: 1205–1218.
7. CPMP concept paper on the development of a CPMP points to consider on biostatistical/methodological issues arising from recent CPMP discussions on licensing applications: choice of delta. www.emea.eu.int. Date updated: September 1999; Date last accessed: June 28 2003.

From the authors:

We are in agreement with G.L. Braunstein′s comment that it is important to discuss end-points that are of clinical relevance to patients and to describe not just average values but results accounting for the entire population. His letter also provides us with an opportunity to rechallenge some of the commonly held assumptions about the clinical interpretation of asthma data and re-emphasise the clinical advantages of our approach to describing the response of patients to asthma therapy. We particularly wish to focus on the methodological issues that he has raised.

The best definition available of asthma control is from the consensus guidelines; unfortunately, this description is neither population-based nor validated. Difficulties have been encountered in efforts to identify and accurately validate the important clinical measures of asthma control, and it remains

unclear which outcome or combination of outcomes will define and serve to measure control of the disease. For example, a number of recent publications have focused on novel approaches to asthma control [1–3]. In our study, the asthma control day was chosen as the primary end-point because this end-point incorporates patient-oriented measures, such as symptom scores, the need for rescue medication, and incidence of asthma exacerbations, measures that are directly and clinically relevant to patients.

We believe that the analyses of clinical trials that rely solely on comparisons of mean values frequently result in loss of important clinical information that is available in the distribution of response in the population. The overlap in response distributions can often provide a clearer determination of clinical comparability because the entire range of responses of individuals in the study population is taken into account. The plot and overlap statistics give the reader a more comprehensive interpretation of the study data than means and p-values, conveying to physicians a better understanding of the range of expectations for patient outcomes.

In addition, comparisons based on means are questionable when using data that show a clear deviation from normality. The overlap statistic is based on the well-known nonparametric Mann-Whitney U-test statistic. This approach is strictly rank-based and therefore does not rely on the assumption of normally distributed data. Using these criteria, there was a clinically important degree of overlap (>85%) in the response of patients treated with either active therapy in the asthma control day end-point despite the mean differences. In addition, both active therapies produced an overlap response (>95%) for forced expiratory volume in one second with no important difference in mean values between both treatments.

The question of equivalence bands with response distributions requires a different consideration from that noted by G.L. Braunstein. The approach here has to be similar to that described for pharmacokinetic bioequivalence. For example, our paper prespecified a criterion for the lower boundary of a 95% confidence interval estimate of overlap being >80%. Values of 80% were suggested by ROM and HWANG [4] and correspond roughly to differences between means of approximately half a SD, which is generally considered small-to-moderate. This is consistent with the bioequivalence literature and its use in other areas [5].

We agree with G.L. Braunstein that crossover designs have the theoretical advantage of providing within-patient comparisons. However, our extensive experience with these studies leads us to the conclusion that the day-to-day biological variability of asthma (at least in moderate-to-severe asthmatic patients) is so large that it confounds the ability to make clear within-patient conclusions (for example, a patient "responds" to one therapy but not to another).

Finally, we call for an accelerated effort in additional methodological work in the field of asthma, including the development of population-based and validated definitions of severity, control and response (all three measuring different dimensions of disease in patients and of treatment) as well as clearly defined and validated definitions of asthma control outcomes and surrogate clinical end-points. It is only through the development of these methodologies and definitions that we will be able to make the best judgments for asthma patients.

T.F. Reiss, J.M. Edelman, J.F. Heyse
Merck & Co. Inc., Rahway, NJ, USA.

References

1. Sont JK, Willems LN, Bel EH, *et al.* Clinical control and histopathologic outcome of asthma when using airway hyperresponsiveness as an additional guide to long-term treatment. *Am J Respir Crit Care Med* 1999; 159: 1043–1051.
2. Reddel H, Ware S, Marks G, Salome C, Jenkins C, Woolcock A. Differences between asthma exacerbations and poor asthma control. *Lancet* 1999; 353: 364–369.
3. Green RH, Brightling CE, McKenna S, *et al.* Asthma exacerbations and sputum eosinophil counts: a randomised controlled trial. *Lancet* 2002; 360: 1715–1721.
4. Rom DM, Hwang E. Testing for individual and population equivalence based on the proportion of similar responses. *Stat Med* 1996; 15: 1489–1505.
5. United States Food and Drug Administration, Centre for Drug Evaluation and Research. Guidance for Industry: Statistical approaches to establishing bioequivalence. Rockville, United States Food and Drug Administration, 2001.