Evidence-based recommendations or “Show me the patients selected and I will tell you the results”

Leif Bjørmer*

Department of Respiratory Medicine & Allergology, 221 85 Lund, Sweden

Summary  Modern treatment decision making in medicine is recommended to be evidence based. In order to have a high grade of evidence, the studies should have sufficient power, be randomized and double blinded. Moreover, the evaluation parameters should be solid and reproducible. While there is a lot focus on primary study design much less is focused on patient eligibility and to what extent the patients included in a clinical trial are representative for the patients treated in "real life". By knowing the profile and action of the drug it is thus often possible to design inclusion criteria’s that already from the start, with high probability, predict the results. If you wish to prove that the addition of long-acting beta-2 agonist is better that increasing the anti-inflammatory treatment you should select mild stable and highly reversible patients. On the other hand, if you wish to prove that increasing anti-inflammatory treatment is beneficial, you choose patients proven to be steroid responsive and slightly under treated. Applying common inclusion and exclusion criteria’s often render only a few percent of the patient population eligible. This is often forgot, when the results from these strictly selected patient populations are extrapolated into "evidence based” treatment recommendations directed towards a much larger and less selected patient population. Thus when evidence are graded, it is important also to consider to what extent the results are extendable to a much larger "real life” patient population. Modern asthma management must consider pathophysiological mechanisms that not necessarily are reflected by lung function parameters. Demands from medical authorities as the medical authorities in Europe (EMEA) asking for 15% reversibility of asthma study patients, helps to conserve the imbalance between clinical trials and real life.

© 2006 Published by Elsevier Ltd.

Introduction

Modern treatment decision-making in medicine is recommended to be evidence based. In a practical sense it means that we should rely more on randomised placebo-controlled trials, and less on
what is usually referred to as personal or clinical experiences. The medical evidences are graded according to the strength of evidence, with evidence level $A$ being a very strong recommendation based upon a large number of patients tested in a randomised placebo-controlled manner, with sufficient evaluation period and solid evaluation parameters. Level $D$ is more or less based upon preliminary observations and expert opinions.\(^1\)

It is of course necessary to act according to available evidence and the opposite would be a disaster. However, while we are focussing upon the results of a clinical trial, it is equally important to consider whether the patients included are representative of the patient in front of me and that I am about to treat. If we focus more on the results and less on the inclusion criteria you may come up with wrong treatment recommendations. This was excellently addressed in a recent letter by Bisgaard and Szefler\(^2\) in *The Lancet*. While the use of long-acting beta-2 agonists (LABAs) are well documented in adults, these results have in clinical practise been extrapolated to children, even though the evidence for effect in this group is ill documented. In Denmark, since 2003 more children have been treated with fixed combinations of fluticasone and salmeterol (FP/SM), than by FP alone (Fig. 1), potentially raising the risk of overtreatment of LABAs, especially in mild asthmatic patients.

I hypothesise that by knowing the inclusion criteria used for a clinical trial, it is also possible to predict the results with a high degree of probability. This can be exemplified by a number of clinical trials documented in the literature.

**Statement:** “The addition of a long-acting beta-2 agonist is better than increasing the anti-inflammatory treatment”

If you wish to document that a bronchodilator is better than increasing the anti-inflammatory treatment, i.e. increasing the inhaled corticosteroid (ICS) dose, you need to select patients according to criteria that are likely to favour bronchodilation. Several large multicentre studies have convincingly shown that the addition of an LABA is better than increasing (doubling) the dose of ICS. One recent example is the comparison of a fixed combination of FP and SM (FP/SM), compared to a higher dose of FP.\(^3\) In this large multicentre study including 3421 patients, it was shown that FP/SM was slightly superior to FP in achieving optimal clinical control; this was true for all stages of severity. However, the patients included in the study were already known to be excellent beta-2 agonist responders with a documented reversibility of over 25\% after a short-acting beta-2 agonist (SABA). In another study comparing the addition of SM or montelukast (ML) to a fixed dose of FP (FP/SM or FP/ML), FP/SM was found to be superior to FP/ML giving better symptom control and better improvement in lung function.\(^4\) Also in that study the response to SABA was excellent, with a reversibility of more than 20\%. Interestingly, FP/SM has also been documented to be superior to FP in patients with mild steroid-naïve asthma. This is in contrast to the combination of budesonide and formoterol (BUD/FM), which was unable to show the same superiority.\(^5\) While BUD/FM improved FEV\(_1\) slightly when compared to BUD, there was no difference in exacerbation rates or symptom control. The reason for this difference is probably not due to crucial differences in pharmacological effects between FP/SM and BUD/FM but rather due to differences in degree of pre-study beta-2 agonist responsiveness. In the BUD/FM study, the pre-study reversibility criterion was 12\% improvement in FEV\(_1\) after SABA or a PEF variability of $\geq 15\%$. The actual reversibility is not written in the publication but is assumed to be less than in the FP/SM trial, with a pre-requested SABA reversibility of $\geq 15\%$ (the actual mean reversibility was 27–29\%).

The conclusion from a clinical study is also highly dependent on the choice of primary evaluation parameters. This was clearly shown in a study.
comparing triamcinolone (TCA) versus SM as controller therapy in patients with mild-moderate asthma. While monotherapy with TCA and SM showed similar improvement of lung function and reduced the need for SABA compared to placebo, SM was clearly inferior in controlling underlying inflammation (i.e., induced sputum eosinophils and exhaled nitric oxide), and was also associated with more exacerbations and treatment failures. Similar discrepancies can be noted when comparing ML versus SM as add-on therapy to FP. While FP/SM was found to be superior in improving lung function and symptom control in patients with a pre-defined excellent SABA response, in another study FP/ML was shown to be superior to FP/SM in preventing exacerbation-induced bronchoconstriction after regular treatment.

Statement: “Increasing the anti-inflammatory treatment is better, or at least as good as, adding a long-acting beta-2 agonist”

Airway inflammation is the underlying cause of asthma symptoms. If you want to prove the hypothesis that increasing anti-inflammatory treatment is better or equally effective than adding an long acting beta-2 agonist (LABA), you have to (1) choose patients with highly likely insufficient inflammatory control, and (2) choose the correct evaluation parameters.

Choosing patients with insufficient inflammatory control

Data from the FACET (Formoterol And Corticosteroids Establishing Therapy) study were published in 1997. This was one of the first studies which convincingly showed that the addition of an LABA not only had a beneficial effect on lung function but also protected from asthma exacerbations. One major inclusion criterion was “insufficient clinical control on ICS treatment alone”. Moreover, during the run-in all patients received a high daily dose of BUD (1600 µg), and those who did reach sufficient clinical stability were excluded. Thus, 262 of 1114 patients (24%) entering the run-in were excluded, mainly due to insufficient steroid responsiveness. The study results showed that the addition of FM to each BUD step was beneficial. However, the data also showed that increasing the BUD dose from 200 to 800 µg/day was clearly more beneficial than adding FM, with a 41% decrease in severe exacerbation rate compared to an 18% decrease when FM was added.

The IMPACT (Investigation of Montelukast as Partner Agent for Complementary Therapy) study is another example showing that increasing anti-inflammatory treatment may be as beneficial as adding LABAs. This study included 1490 patients with documented insufficient clinical control on FP 200 µg daily during a run-in period of 4 weeks. The primary aim of the study was to see whether the addition of ML was as good as adding SM. The primary evaluation parameter in this study was, as in the FACET trial, the number of severe asthma exacerbations. The study concluded that the addition of ML was as good as adding SM with a similar decrease in the numbers of asthma exacerbations, nocturnal awakenings and improvement of quality of life (QoL).

Choosing the right evaluation parameter

Both the FACET and the IMPACT studies showed that the addition of LABA resulted in a better lung function measured as PEF and or FEV1. However, this improvement was not associated with either protection from exacerbations or improvement in QoL. Thus, it is clear that lung function improvement is not clearly related to other important measures of asthma control. Other important parameters related to asthma control are bronchial hyper-responsiveness as well as surrogate markers of inflammation. In the study
by Sont et al., it was convincingly shown that the use of bronchial hyper-responsiveness to metacholine was superior to symptoms and lung function as a guiding instrument for adjusting steroid treatment. Surrogate markers of inflammation have also been used as a guidance instrument with similar success. In the study by Green et al., ICS treatment was adjusted according to the degree of inflammation estimated from the induced sputum eosinophils and measured at regular intervals. By using sputum eosinophils as a guiding instrument, the number of asthma exacerbations decreased significantly in parallel to an achieved lower number of eosinophils in sputum. This could be achieved, without increasing the total ICS dose in the sputum group compared to the controls. In another study, exhaled nitric oxide was used as a guiding instrument in the treatment of patients with chronic asthma. ICS dose adjustment made in order to keep exhaled nitrous oxide (ENO) below 15 PPB resulted in fewer exacerbations despite a lower total ICS dose used in the ENO-controlled arm. If these surrogate markers of inflammation are being used as monitoring instruments, then increasing the dose of ICS seems to be superior to adding LABA in patients with insufficient control on low-to-medium doses of ICS.

How representative are clinical trial asthma patients compared to a general asthma population?

A clinical drug on its way to the market will need to pass through a number of "needle eyes" before the drug can be used on a larger patient population (Fig. 2). In early Phases I–II, very strict inclusion criteria are used in order to prove the concept, evaluate the therapeutic ratio and establish the correct dose for future trials. In Phase III, the inclusions are still very strict, and it is during this phase that most clinical documentation is collected. Phase III studies are also the sole basis for evidence-based decision making. However, the strength of evidence is linked to the degree of representativeness, i.e. it must be possible to extrapolate the results from concluded clinical trials into the real life situation. This was investigated in a recent study, 928 prospective patients with obstructive lung disease from three specialist clinics and 6 GP centres were classified according to clinical phenotype, use of medication, presence of co-morbidity, treatment and smoking history. The aim was to see how common a typical "trial asthma patient" was compared to "real life" asthma patient population. "Inclusion criteria" for suitable clinical "trial asthma patient" were defined as (a) Doctors diagnosed asthma without concomitant COPD, (b) FEV1 between 50% and 85% of predicted normal, (c) Proven reversibility of FEV1 of at least 12% during the last 12 months, (d) No smoking or if a previous smoker a smoking history of less than 10 packs per year, (d) No significant co-morbidity that theoretically could interfere with the treatment. Of the 928 patients entered, 870 patients were eligible but only 18 patients (2%) remained when the selection criteria were applied. Finally, if a patient is entered into a clinical trial evaluating the impact of additional therapy, then one additional inclusion criterion should be evidence of active uncontrolled disease. If this criterion is added, then only 11 patients (1.3%) would remain (Fig. 3).
Conclusions

From fairly straightforward clinical trial inclusion criteria, less than 2% of patients can be regarded as eligible for inclusion in a clinical trial. Many clinical trials have inclusion criteria that limit the eligible study population even more. Thus, it is questionable whether it is possible to extrapolate knowledge gained from the results of such a strictly selected patient study into the real life situation. A further complicating factor is the requirements of the medical authorities. Despite the poor association between lung function measurements and clinical asthma control, lung function parameters are regarded as important by the medical authorities, regulators and reimbursement boards. The medical authorities in Europe (EMEA) are presently demanding that asthma patients should have a FEV1 reversibility of at least 15% to be considered as real asthma patients in clinical trials. This is certain to preserve this imbalance between clinical trials and real life. There is a need for recognition of the contribution from large “outcomes research studies” which reflect real life more accurately. For the future, it is important that every (randomised controlled) clinical study is accompanied by a statement, where responsible investigators argue for how results are to be interpreted and extended to the larger real life population.

References