

Editorial

Methodology and quality of immunotherapy trials

Hans-Jørgen Malling

Allergy Clinic,
National University Hospital,
Copenhagen, Denmark

Hans-Jørgen Malling
Allergy Clinic 4222
Rigshospitalet Blegdamsvej 9
DK-2100
Copenhagen
Denmark

Accepted for publication 6 January 2004

The tradition of basing the clinical efficacy of immunotherapy on controlled studies is only 55 years old (1). The risk of not comparing the advantage of active treatment with placebo has been illustrated for *Hymenoptera* venom allergy showing that whole-body extract was as effective as placebo (2). In the evidence-based thinking of modern medicine, it is crucial that treatments proposed to patients fulfill at least the criterium that the clinical efficacy could not be questioned, i.e. there is a solid scientific documentation of the benefit of intervention (3).

During the last 10 years the international recognition of efficacy of sublingual immunotherapy is reflected in position papers and meta-analysis. In the 1993 EAACI position paper (4) only two studies supported efficacy. In 1998 (5) the number was doubled, and in the 2001 ARIA document (6) increased to 13 studies. The latest Cochrane analysis (7) identifies 22 studies fulfilling criteria to conclude on efficacy, identical to another review (8). Another review published in 2002 identified 23 studies of which 26% were categorized as unequivocally effective, 35% possible effective (significant improvement in either symptom or medication scores), and 39% without statistically documented efficacy (9). In the latter review eight studies investigating the efficacy of house dust mite allergy documented clinical efficacy in half of the studies.

The study in this issue of *Allergy* (10) further adds to the documentation of efficacy in house dust mite allergy. The immunotherapy study included 32 patients divided in two study groups, and during the 2-year treatment period a steady decrease in rhinitis score was observed. At termination of the study the symptoms in actively treated were reduced to about one-third of placebo-treated. The medication score did not differ significantly between the two groups probably caused by a general low intake of

drugs. The study is interesting as it only considered immunotherapy in those patients with residual symptoms after application of mite-reducing intervention (anti-mite mattress cover). Only about 25% of 120 screened patients using mattress cover fulfilled the inclusion criteria of residual symptoms of >40% of the possible maximal symptom score indicating that, in contrast to the general opinion (11) rigorous house dust mite avoidance and exposure control will control most of the clinical symptoms in the majority of house dust mite allergic patients. Like the problem in many immunotherapy studies the level of allergen exposure varies during the treatment period and the initial beneficial effect of intervention seems partly to wean-off during the treatment period reinforcing that the reduction in clinical symptoms is indeed caused by active immunotherapy.

Some years ago an editorial on sublingual immunotherapy (12) raised reservations on the quality of published studies. Studies included insufficient number of patients, the drop out rate was unacceptably high, and the randomization process was hampered by no baseline assessment. The conclusion of this editorial was that 'future studies should be large, properly randomized, and controlled. Baseline symptom levels should be prospectively assessed to ensure precise balancing of the groups for all relevant major end points'.

In order to bring the documentation for clinical efficacy to modern standard, to ensure that patients are offered treatment based on scientific evidence, and to minimize the risk of misuse of the limited financial resources for scientific evaluations on inconclusive studies, it may be time for scientific journals to reconsider their policy for accepting scientific studies. Several journals (*Lancet*, *New England Journal of Medicine*, etc.) only accept studies conducted according to the Consort guidelines (13).

These include important criteria in relation to the evaluation of the quality (and consequently the reliability) of published studies, but even more essential to the planning of conclusive studies.

Importance of pretreatment monitoring

Most immunotherapy studies do not include a pretreatment monitoring period. The advantages of this relate both to an optimal allocation to treatment groups and to the documentation of clinical efficacy. Direct comparison of the results obtained in the placebo and the actively treated group are based on comparing comparable groups with respect to parameters important for the primary outcome. Allocation to intervention groups based on subjective data (patients recall of disease severity in the preceding season) implies a high risk of biasing the groups, and consequently to introduce both statistical type I (not being able to detect a factual effect) and type II errors (demonstrating an effect that actually does not exist). The advantage of including a pretreatment monitoring period is that the disease severity can be precisely determined based on objective parameters and the patients randomized appropriately without bias. Furthermore, estimating changes from pretreatment a further compensation for variation in disease severity scoring might be obtained. By using the patient as his own control in relation to the clinical outcome the possibility to detect both improvements and deteriorations in disease severity evolve. A crucial factor for the approach to evaluate changes in disease severity is the constancy of allergen exposure during succeeding evaluation periods.

Randomization

Participants in clinical studies should be assigned to comparison groups on the basis of a random process characterized by unpredictability. The methods used to generate the random allocation sequence should minimize the risk of bias in group assignment. Restricted randomization is often used in trials including a smaller group of patients to achieve balance between groups in size and characteristics. Randomization could be in blocks (14), or by stratification according to predefined clinical criteria like age, gender, disease (15). Allocation to treatment groups using minimization has the advantage of making small groups closely similar in terms of several characteristics (16).

Blinding of intervention and assessment

The blinding of the trial is important to avoid bias (17). The use of placebo combined with blinding is important

to evaluate the 'true' treatment effect associated with the intervention. The blinding should also include the data analyst and the success of the blinding should be evaluated by asking participants.

Participants flow

The flow of participants through each stage of the study is critical for the evaluation of efficacy. Information for each group on the number of participants randomized, receiving treatment, completing the study protocol, and analysed for outcome is important. Many studies do not give information on whether some participants did not receive treatment as allocated, were lost to follow up, or excluded from the analysis. Ideally, the primary outcome (clinical efficacy) should be based on all patients randomized (intention-to-treat analysis). In order to evaluate the 'true' treatment effect, it is crucial to include participants leaving the study because of insufficient clinical efficacy in the primary outcome analysis. Likewise, patients leaving the study because of adverse effects should be included in the safety analysis.

Minimal clinically relevant efficacy

In the discussion of the clinical efficacy it is important to remember that allergic patients suffer from allergic symptoms and the need for intake of anti-allergic drugs. Consequently efficacy should be based on a reduction in these parameters. By including a high number of participants statistically significant, but clinically irrelevant differences might be observed. The magnitude of efficacy should be clinically relevant, i.e. the reduction in symptom scores and drug consumption should significantly reduce the morbidity of the disease. An attempt to define the minimal clinically relevant reduction in disease severity has been published (18).

Statistical methods

When planning a study the number of patients to be included should be large enough to have high probability (power) to detect statistical significance. Power calculations depend on the expected magnitude of efficacy and the variability of included patients and must be applied in the planning phase of the trial. The size of the clinical effect to be evaluated is inversely related to the sample size necessary to detect it (large samples are necessary to detect small differences) (19). Inclusion of insufficient numbers of patients increases the risk of statistical type I errors, but also in the case of differences in disease severity between the evaluated groups at inclusion to result in statistical type II errors. The choice between

parametric and nonparametric statistics could be discussed. In biologic science most small samples are not normally distributed. This may be circumvented partly by transformation of data, but showing statistical significant difference using nonparametric tests usually reinforces the conclusion. In order to improve the readout of clinical studies it is recommended to include confidence intervals giving the range of the differences or the magnitude of efficacy between the two rates, i.e. the range of uncertainty expected to include the true value.

In summary it is recommended that future immunotherapy studies should be planned according to the following essentials:

- The study should be a placebo-controlled, double-blind, randomized trial.
- The sample size should be large enough to detect statistically significant and clinical relevant changes in disease severity.
- Patients should be selected according to predefined clinical criteria.
- The primary and secondary outcome measures should be clearly defined.
- Allergen extracts and maintenance doses should be defined.
- The duration of treatment should be sufficient to allow clinical improvement to occur.

References

1. BRUUN E. Control examination of the specificity of specific desensitization in asthma. *Acta Allergol* 1949;**2**:122.
2. HUNT KJ, VALENTINE MD, SOBOTKA AK, BENTON AW, AMODIO FJ, LICHTENSTEIN LM. A controlled trial of immunotherapy in insect hypersensitivity. *New Engl J Med* 1978;**299**:157–161.
3. GUYATT GH, SACKETT DL, SINCLAIR JC, HAYWARD R, COOK DJ, COOK RJ. Users' guides to the medical literature. IX. A method for grading health care recommendations. Evidence-Based Medicine Working Group. *JAMA* 1995;**274**: 1800–1804.
4. MALLING H-J, WEEKE B. EAACI immunotherapy position papers. *Allergy* 1993;**48**(Suppl. 14):9–35.
5. BOUSQUET J, LOCKEY RF, MALLING H-J (eds). WHO Position Paper. Allergen Immunotherapy: therapeutic vaccines for allergic diseases. *Allergy* 1998;**53**(Suppl. 44):1–42.
6. BOUSQUET J, VAN CAUVENBERGE P, KHALTAEV N. Allergic rhinitis and its impact on asthma. *J Allergy Clin Immunol* 2001;**108**:147–334.
7. WILSON DR, Torres Lima M, Durham SE. Sublingual immunotherapy for allergic rhinitis. *The Cochrane Database Syst Rev* 2003;**2**:CD002893.
8. CANONICA GW, PASSALACQUA G. Noninjective routes for immunotherapy. *J Allergy Clin Immunol* 2003;**111**: 437–448.
9. MALLING H-J. Is sublingual immunotherapy clinically effective? *Curr Opin Allergy Clin Immunol* 2002;**2**:523–531.
10. TONNEL AB, SCHERPEREEL A, DOUAY B, MELLIN B, LEPRINCE D, GOLDSTEIN N et al. Allergic rhinitis due to house dust mites: evaluation of the efficacy of specific sublingual immunotherapy. *Allergy* 2004;**59**:491–497.
11. SHEIKH A, HURWITZ B. House dust mite avoidance measures for perennial allergic rhinitis. *Cochrane Database Syst Rev* 2001;**4**:CD001563.
12. FREW AJ, WHITE PJ, SMITH HE. Sublingual immunotherapy. *J Allergy Clin Immunol* 1999;**104**:267–270.
13. MOHER D, SCHULZ KF, ALTMAN DG. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials. *Ann Intern Med* 2001;**134**:657–662.
14. ALTMAN DG, BLAND JM. How to randomise. *BMJ* 1999;**319**:703–704.
15. ENAS GG, ENAS NH, SPRADLIN CT, WILSON MG, WILTSE CG. Baseline comparability in clinical trials. *Drug Inf J* 1990;**24**:541–548.
16. TREASURE T, MACRAE KD. Minimisation: the platinum standard for trials? Randomisation doesn't guarantee similarity of groups; minimisation does. *BMJ* 1998;**317**:362–363.
17. DEVEREAUX PJ, MANNS BJ, GHALI WA, QUAN H, LACCHETTI C, GUYATT GH. In the dark: physician interpretations and expert definitions of blinding in randomized controlled trials. *JAMA* 2001;**285**:2000–2003.
18. MALLING H-J. Immunotherapy as an effective tool in allergy treatment. *Allergy* 1998;**53**:461–472.
19. CAMPBELL MJ, JULIOUS SA, ALTMAN DG. Estimating sample sizes for binary, ordered categorical, and continuous outcomes in two group comparisons. *BMJ* 1995;**311**:1145–1148.