
Immunotherapy for asthma: unfavorable studies

Jay M. Portnoy, MD

Introduction: Many studies have demonstrated efficacy of immunotherapy (IT) for asthma. Two meta-analyses reached the same conclusion. Because some studies failed to demonstrate efficacy, it is instructive to review reasons why that may have been the case.

Methods: Two meta-analyses were used to identify studies that did and did not demonstrate efficacy. These were reviewed to identify the sources of variability.

Results: Most studies are able to demonstrate a treatment effect of IT although they have wide confidence intervals. The studies were heterogeneous with respect to selection of subjects, underlying populations, the main treatment protocol, concomitant treatments, the duration of treatment and followup, and the effect of the protocol on concurrent treatment. There were also problems with the blinding. In most cases, it was difficult to determine how the doses of extract given compared between studies.

Conclusions: Studies that fail to demonstrate a treatment effect of IT may do so for many reasons. Given the heterogeneity of IT clinical trials as a group, it is not surprising that variable results occur. The use of multiple studies in a meta-analysis with its underlying assumption of homogeneity is problematic.

Ann Allergy Asthma Immunol 2001;87(Suppl):28–32.

INTRODUCTION

Immunotherapy (IT) has been used for many years as a treatment for rhinitis and asthma triggered by allergic exposures. Many studies looking at the effect of IT for both of those diseases in children and adults have been published in the medical literature. Although most have demonstrated a positive effect of the treatment (termed “favorable” studies in this article), others either failed to show such an effect or found a detrimental effect of IT (termed “unfavorable” studies). Because of these apparently contradictory results, several meta-analyses were performed in an attempt to reach a conclusion about the effectiveness of IT. Two of those meta-analyses^{1,2} concluded that the studies used in their evaluation, when statistically combined, clearly demonstrate the efficacy of IT for treatment of asthma.

Given the conclusions of these meta-analyses, it is instructive to re-

view the sources of variability in the studies and reasons why some of the studies may have failed to evince a treatment effect. Clinical trials involving the same topic may reach contradictory conclusions for a number of reasons. Horwitz³ identified nine methodologic sources for this variation. Those sources included six items concerned with the design of the trials and three concerned with their interpretation. Because there are conflicting studies of IT for treatment of asthma, the sources of variation identified by Horwitz will be used as an outline to determine whether the conclusions of the meta-analysis are supported by the studies themselves. The meta-analyses cited above were used to identify studies for this review as listed in Table 1.

SOURCES OF VARIATION IN CLINICAL TRIALS

The customary reasons given for variations in the results of clinical trial include inadequate statistical power and inadequate study design that leads to bias. Studies that show a significant treatment effect obviously have sufficient numbers of subjects to demonstrate the effect, so lack of power is an

issue only for studies that fail to show an effect. In the last 20 years, an increasing number of studies provide an indication of the number of subjects that would be required to demonstrate both a statistically and clinically significant effect with a stated degree of certainty. Studies that fail to detect a treatment effect when one is actually present are said to have committed a type II error (a type I error would occur if a treatment is found to be effective when in fact it is not).⁴ The use of meta-analyses represents an attempt to circumvent the deficiencies of studies that have insufficient subjects by adding their numbers together, thus increasing the power of the analysis. Such analyses are performed with an assumption that the studies are sufficiently comparable for their results to be combined. Although this may be a reasonable assumption for some topics, clinical trials of IT tend to be so heterogeneous that variations may be a significant problem. Consideration only of clinical trials that include randomization, blinding, and use of placebo control groups in their protocols has reduced the issue of bias in these types of study.

Horwitz identified six design issues that cause variability in clinical trials. These issues are: 1) differences in eligibility criteria; 2) baseline differences in the available populations; 3) variability in treatment protocols for the principal therapy; 4) variable concomitant therapies; 5) protocol requirements of the randomized clinical trial; and 6) management of intermediate outcomes, such as when multiple measured outcomes give conflicting results. Issues involving interpretation of the results include: regulatory variables in interpretation of trial results, frailty of double-blinding, and trial variability and occurrence of unexpected outcomes. These issues will be considered in an attempt to identify

Children's Mercy Hospitals and Clinics, Kansas City, MO.

Received for publication November 10, 2000.

Accepted for publication November 15, 2000.

Table 1. Studies Evaluated for this Review

Study	Reference	Year	Reviewed by		N	Antigen	Effective	
			Abramson	Ross			Abramson	Ross
Aas	28	1971	no	yes	80	mite		yes
Adkinson	6	1997	no	yes	108	ragweed		no
Bousquet	20	1985	yes	yes	30	mite	yes	yes
Bruce	29	1977	yes	No	33	ragweed	no	yes
BTA	8	1979	yes	No	56	mite	no	
Creticos	5	1996	no	yes	50	ragweed		equiv
Dreborg	30	1986	yes	yes	30	<i>Cladosporium</i>	yes	
D'Souza	23	1973	yes	yes	91	mite	yes	no
Frankland	31	1954	yes	No	57			
Haugaard	14	1992	no	yes	24	cat or dog		yes
Johnstone	32	1968	no	yes	130	ragweed		yes
Malling	26	1986	no	yes	22	<i>Cladosporium</i>		yes
Maunsell	18	1971	yes	No	34	mite	yes	
Murray	22	1985	no	yes	11	mite		no
Newton	19	1978	yes	yes	14	mite	no	no
Ohman	9	1984	yes	yes	17	cat	equiv	yes
Olsen	33	1997	no	yes	31	mite		yes
Ortolani	7	1984	no	yes	15	grass		yes
Pauli	21	1984	yes	yes	18	mite	no	no
Pichler	34	1997	no	yes	30	mite		yes
Reid	25	1986	no	yes		grass		yes
Smith	16	1971	yes	yes	22	mite	yes	yes
Sundin	12	1986	yes	yes	39	cat or dog	equiv	yes
Taylor	10	1978	yes	yes	10	cat	yes	no
Valovirta	13	1984	yes	yes	27	dog	no	yes
Van Bever	15	1992	no	yes	18	mite		yes
Van Metre	11	1988	yes	No	22	cat	yes	
Warner	17	1978	yes	yes	51	mite	yes	yes

Note that the Abramson and Ross meta-analyses considered 6 studies differently with respect to whether or not they showed a treatment effect of immunotherapy.

sources of variation as an explanation for conflicting results in the IT clinical trials.

Subjects who are enrolled in a clinical trial must pass a number of eligibility decisions that are designed to ensure that an appropriate group of patients enter the study. Variations in these criteria can introduce differences in the types of subjects being studied. For example, two clinical trials^{5,6} that failed to demonstrate an effect of IT incorporated strict entry criteria into their protocols for reasons that were unique to those studies. In the Adkinson study,⁶ the investigators screened 350 children to obtain 121 who were ultimately enrolled. The protocol required subjects to undergo a 1-year lead-in time during which noncompliance, either with medications or environmental control, led to expulsion

from the study. This resulted in a highly selected group of subjects with unique characteristics. In the Creticos study,⁵ the investigators screened 1,000 candidates to obtain 90 subjects. Of that number, 77 entered the study, 64 completed the first year, and 54 completed the full 2 years of the study. This was done to identify asthmatic adults who were exclusively sensitive to ragweed and who had symptoms of asthma during ragweed season. These studies used very strict selection criteria to eliminate confounding factors for specific reasons. In doing so, they are likely to have studied subjects that differ from those used in clinical trials that did show an IT effect. In contrast, Ortolani,⁷ in a study that was able to demonstrate efficacy of grass extract, excluded a similarly large number of subjects by selecting 15 of 200 who

were sensitive only to grass pollen. Therefore, it seems unlikely that strict subject selection, per se, determines whether IT is effective, unless the criteria alters the subject population in a manner that results in a group of patients for whom IT is not effective, such as those with perfect medication and environmental control compliance.

The Adkinson study⁶ had another unique feature in that each patient's medication was reduced before and during the course of the treatment according to a strictly defined protocol. This created a source of variability in the concomitant medical therapy. Although other clinical trials permit a reduction in medication use during the course of therapy when indicated, they commonly measure those reductions as outcomes, rather than using them as independent variables. The duration of

the Adkinson protocol and the consistency of the medication reduction are not matched by any of the other studies. Because the study was designed this way specifically to determine whether IT provides benefit beyond medical treatment and environmental control, this factor alone may have reduced the likelihood that the study would show an effect. Another example of variability introduced by concomitant therapy was noted in the British Thoracic Association study,⁸ in which, for reasons that are unknown, patients who were in the placebo group did better than the active group if they were also on corticosteroids.

Baseline differences in patient populations that are available to the investigators may introduce variability that is not accounted for by the protocol design. Such differences are likely to occur between studies performed in Europe and those performed in the United States where different populations reside and exposures to allergens are not the same. For example, 5 of 6 trials using animal extracts were performed in the United States, where exposure to pets may differ from exposure in Europe because of cultural differences.⁹⁻¹⁴ Eight studies^{8,15-21} of dust mite were done in Europe, where mite exposure is very high, whereas two studies^{22,23} were done in the United States with lower, more seasonal exposure. This may have contributed to the fact that 6 of 8 European studies showed an effect of mite IT, whereas one study done in the United States was negative and one was equivocal. Since none of the studies measured personal dust mite exposure, it is possible that IT is more effective in locations where exposure is greater.

Differences in the primary treatment protocol are potentially the greatest source of variability in clinical trials of IT. Protocol differences occur for a number of factors including: antigen used, type of extract, duration of treatment, and the final maintenance dose. The two meta-analyses considered studies that used the following antigens: 3 cat, 2 cat or dog, 1 dog, 2 *Cladosporium*, 2 grass, 13 mite, and 4

ragweed (Table 1). These antigens came in the following forms: aqueous whole extract, lyophilized, alum precipitant, partially purified, and tyrosine absorbed. Maintenance doses were expressed in noon units, protein nitrogen units, weight per volume, RAST units, micrograms of major allergen, FDA units, unspecified units, and bioequivalent allergy units. Though in some cases it is possible to convert one type of unit into another so that the studies can be compared, the wide variety of units used in the clinical trials illustrates the huge heterogeneity that is intrinsic to this type of investigation. Because the dose of extract given has been shown to be related to its efficacy,²⁴ it is likely that this is a significant source of variability. Even so, there does not seem to be a systematic under-dosing in studies that failed to show an effect, as illustrated by the fact that the Taylor and Van Metre studies^{10,11} provided the same maintenance dose of cat, and yet, the former showed an equivocal effect whereas the latter showed a positive effect.

The majority of studies demonstrate a strong placebo effect. This is most likely related to an impact of the conduct of the clinical trial on the primary outcomes. In particular, subjects receiving IT are seen and evaluated on a weekly basis, making it likely that more effective treatment is given for their underlying disease.

Concomitant therapies are another source of variability. The Reid study²⁵ gave IT with antigens other than the extract to which subjects were sensitive in both active and placebo groups to simulate adverse reactions in both groups and improve the quality of the blinding. By doing so, they may have inadvertently introduced another source of variability by altering concomitant treatment, leading to an effect that may not have been attributable only to the active treatment. Although the study demonstrated a treatment effect of IT with grass extract, it is possible that concomitant treatment with other relevant antigens increased the effectiveness of the primary treatment.

Intermediate outcomes are a source of variability when multiple outcomes are measured, some of which show a treatment effect and others that do not. This was the case for the studies of Ohman,⁹ Sundin,¹² and Valovirta.¹³ The Ohman study found an effect for asthma symptoms but not for bronchial hyperresponsiveness, whereas the Sundin study found an effect for bronchial hyperresponsiveness but not for asthma symptoms. Both were considered to have shown an effect in the Ross meta-analysis. The Valovirta study, which failed to show an effect both for symptoms and bronchial hyperresponsiveness, was considered to be favorable in the Ross meta-analysis because it showed an effect when conjunctival challenges were considered as the outcome.

The Ortolani study⁷ had displayed interesting features in that one-sixth of the patients in the active group dropped out because of adverse effects, as opposed to the control group, revealing a fault in the blinding of the study. Difficulties in the blinding also were likely to have been a factor in the studies by Sundin¹² and by Malling,²⁶ in which increased systemic reactions occurred predominantly in the active group, making it likely that the investigators could tell to which group those patients had been assigned. Faulty blinding because of adverse effects of active treatment was specifically described as not being present both in studies that either did show an effect, as seen in the Valovirta and Reid studies,^{13,25} and in studies that did not show an effect, as in the Taylor study.¹⁰ To avoid this faulty blinding problem, the Reid study gave IT with antigens other than the extract to which subjects were sensitive in both active and placebo groups to simulate adverse reactions in both groups. Because faulty blinding was present both in studies showing and in studies not showing an effect, it is not likely that the overall results of the meta-analyses were caused by this source of variability.

Variability between clinical trials also may be introduced by the duration of time that treatment is given. The study durations in the clinical trials of

this review ranged from 7 weeks²⁷ to 2 years,⁵ although the majority lasted for 1 year. There was no consistent tendency for studies of short duration to be favorable or unfavorable. In particular, the shortest duration study was able to demonstrate a significant effect after only 7 weeks, whereas the longest study was considered to be equivocal. These studies had substantially different protocols in that the Bousquet study²⁷ used rush IT to achieve a maintenance dose in 1½ days, whereas the Creticos study⁵ used a weekly build-up schedule that took much longer. Interestingly, that study found a treatment effect during the first year but not during the second, a finding that clinical trials lasting <1 year would not detect.

Perhaps the greatest source of variation is the use of differing outcomes to measure the efficacy of the treatment. Outcomes used in these studies included: peak expiratory flow rates, forced expiratory volume in 1 second, symptoms, medications, bronchial hyperresponsiveness of both antigen and histamine, and conjunctival challenges. The Abramson meta-analysis attempted to evaluate three of these outcomes independently and identified conflicting outcomes for the studies by Ohman⁹ and Sundin.¹² The Valovirta study¹³ was considered to have been unfavorable, although conjunctival challenge showed a treatment effect, because that outcome was not included in the analysis. The Ross meta-analysis attempted to combine its studies, although their outcomes were highly variable, by categorizing them into broad groups such as symptoms, pulmonary function, and challenge studies.

DISCUSSION

So what are we to conclude from studies that arrive at differing conclusions regarding the efficacy of IT? The meta-analyses described attempted to address this issue; however, the validity of that approach is questionable, given the extreme heterogeneity of the IT clinical trials. Because there are wide variations in protocols, it is likely that those studies were testing the ef-

ficacy of completely different treatments that are all bundled under the term IT. For example, short-term rush IT clinical trials may be testing a different treatment than long-term trials of IT are. Trials in which medications are adjusted systematically during the protocol may be testing the effect of medication adjustments as a primary treatment. Because antigen exposure has been measured infrequently, it is not clear what effect that has on the ability to demonstrate a treatment effect of IT, because that form of treatment may work differently depending on the amount of natural antigen exposure. It is not even clear that pollen counts are satisfactory surrogates for antigen exposure, especially because patients spend different amounts of time outdoors.

The only way to resolve these questions, therefore, would be for IT protocols to be more standardized so that the results can be compared. The questions being asked by a particular clinical trial need to be stated more clearly and the protocols designed specifically to answer those questions. Extracts need to be standardized, schedules and maintenance doses need to be consistent, and the same outcomes need to be measured in future studies of IT. Given the highly variable nature of the various protocols, the large confidence intervals, and the multiple sources of variation discussed above, it is remarkable that so many studies were able to show a treatment effect. Given all of those problems, it should not be surprising that a few studies were unable to demonstrate a treatment effect.

CONCLUSION

It is difficult to compare results from the various studies of IT because of the large number of variations in the clinical trial protocols for this topic. Ultimately, additional rigorous studies using standardized protocols of IT are needed before a conclusion can be reached regarding the role of IT in asthma.

ACKNOWLEDGMENTS

I thank Steve Simon, PhD, for his valuable statistical advice.

REFERENCES

1. Abramson M, Puy R, Weiner J. Is allergen immunotherapy effective in asthma? A meta-analysis of randomized controlled trials. *Am J Respir Crit Care Med* 1995;151:969–974.
2. Ross R, Nelson H, Finegold I. Effectiveness of specific immunotherapy in the treatment of asthma: a meta-analysis of prospective, randomized, double-blind, placebo-controlled studies. *Clin Therapeutics* 2000;22:329–341.
3. Horwitz R. Complexity and contradiction in clinical trial research. *Am J Medicine* 1987;82:498–510.
4. Colton T. *Statistics in Medicine*. New York: Little, Brown and Company, 1974:120–124.
5. Creticos P, Reed C, Norman P, et al. Ragweed immunotherapy in adult asthma. *N Engl J Med* 1996;334:501–506.
6. Adkinson N, Eggleston P, Eney D, et al. A controlled trial of immunotherapy for asthma in allergic children. *N Engl J Med* 1997;336:324–331.
7. Ortolani C, Pastorello E, Moss R, et al. Grass pollen immunotherapy: a single year double-blind, placebo-controlled study in patients with grass pollen-induced asthma and rhinitis. *J Allergy Clin Immunol* 1984;73:283–290.
8. British Thoracic Association. A trial of house dust mite extract in bronchial asthma. Mite Allergy Subcommittee of the Research Committee of the British Thoracic Association. *Br J Dis Chest* 1979;73:260–270.
9. Ohman JL Jr, Findlay SR, Leitermann KM. Immunotherapy in cat-induced asthma. Double-blind trial with evaluation of in vivo and in vitro responses. *J Allergy Clin Immunol* 1984;74:230–239.
10. Taylor WW, Ohman JL Jr, Lowell FC. Immunotherapy in cat-induced asthma. Double-blind trial with evaluation of bronchial responses to cat allergen and histamine. *J Allergy Clin Immunol* 1978;61:283–287.
11. Van Metre TE Jr, Marsh DG, Adkinson NF Jr, et al. Immunotherapy for cat asthma. *J Allergy Clin Immunol* 1988;82:1055–1068.
12. Sundin B, Lilja G, Graff-Lonnevig V,

- et al. Immunotherapy with partially purified and standardized animal dander extracts. I. Clinical results from a double-blind study on patients with animal dander asthma. *J Allergy Clin Immunol* 1986;77:478–487.
13. Valovirta E, Koivikko A, Vanto T, et al. Immunotherapy in allergy to dog: a double-blind clinical study. *Ann Allergy* 1984;53:85–88.
 14. Haugaard L, Dahl R. Immunotherapy in patients allergic to cat and dog dander. I. Clinical results. *Allergy* 1992;47:249–254.
 15. Van Bever H, Stevens W. Effect of hyposensitization upon the immediate and late asthmatic reaction and upon histamine reactivity in patients allergic to house dust mite (*Dermatophagoides pteronyssinus*). *Eur Respir J* 1992;5:318–322.
 16. Smith AP. Hyposensitization with *Dermatophagoides pteronyssinus* antigen: trial in asthma induced by house dust. *Br Med J* 1971;4:204–206.
 17. Warner JO, Price JF, Soothill JF, Hey EN. Controlled trial of hyposensitisation to *Dermatophagoides pteronyssinus* in children with asthma. *Lancet* 1978;2:912–915.
 18. Maunsell K, Wraith D, Hughes A. Hyposensitisation in mite asthma. *Lancet* 1971;1:967–968.
 19. Newton DA, Maberley DJ, Wilson R. House dust mite hyposensitization. *Br J Dis Chest* 1978;72:21–28.
 20. Bousquet J, Calvayrac P, Guerin B, et al. Immunotherapy with a standardized *Dermatophagoides pteronyssinus* extract. I. In vivo and in vitro parameters after a short course of treatment. *J Allergy Clin Immunol* 1985;76:734–744.
 21. Pauli G, Bessot J, Bigot H, et al. Clinical and immunologic evaluation of tyrosine-adsorbed *Dermatophagoides pteronyssinus* extract: a double-blind placebo-controlled trial. *J Allergy Clin Immunol* 1984;74:524–535.
 22. Murray A, Ferguson A, Morrison B. Non-allergic bronchial hyperreactivity in asthmatic children decreases with age and increases with mite immunotherapy. *Ann Allergy* 1985;54:151–154.
 23. D'Souza MF, Pepys J, Wells ID, et al. Hyposensitization with *Dermatophagoides pteronyssinus* in house dust allergy: a controlled study of clinical and immunological effects. *Clin Allergy* 1973;3:177–193.
 24. Van Metre TE Jr, Adkinson NF Jr, Lichtenstein L, et al. A controlled study of the effectiveness of the Rinkel method of immunotherapy for ragweed pollen hay fever. *J Allergy Clin Immunol* 1980;65:288–297.
 25. Reid M, Moss R, Hsu Y, et al. Seasonal asthma in northern California: allergic causes and efficacy of immunotherapy. *J Allergy Clin Immunol* 1986;78:590–600.
 26. Malling HJ, Dreborg S, Weeke B. Diagnosis and immunotherapy of mould allergy. V. Clinical efficacy and side effects of immunotherapy with *Cladosporium herbarum*. *Allergy* 1986;41:507–519.
 27. Bousquet J, Hejjaoui A, Skassa-Brociek W, et al. Double-blind, placebo-controlled immunotherapy with mixed grass-pollen allergoids. I. Rush immunotherapy with allergoids and standardized orchard grass-pollen extract. *J Allergy Clin Immunol* 1987;80:591–598.
 28. Aas K. Hyposensitization in house dust allergy asthma. A double-blind controlled study with evaluation of the effect on bronchial sensitivity to house dust. *Acta Paediatr Scand* 1971;60:264–268.
 29. Bruce C, Norman P, Rosenthal R, Lichtenstein L. The role of ragweed pollen in autumnal asthma. *J Allergy Clin Immunol* 1977;59:449–459.
 30. Dreborg S, Agrell B, Foucard T, et al. A double-blind, multi-center immunotherapy trial in children using a purified and standardized *Cladosporium herbarum* preparation. I. Clinical results. *Allergy* 1986;41:131–140.
 31. Frankland A, Augustin R. Prophylaxis of summer hayfever and asthma. *Lancet* 1954;1:1055–1057.
 32. Johnstone D, Dutton A. The value of hyposensitization therapy for bronchial asthma in children: a 14-year study. *Pediatrics* 1968;42:793–802.
 33. Olsen OT, Larsen KR, Jacobsen L, Svendsen UG. A 1-year, placebo-controlled, double-blind house-dust-mite immunotherapy study in asthmatic adults. *Allergy* 1997;52:853–859.
 34. Pichler CE, Marquardsen A, Sparholt S, et al. Specific immunotherapy with *Dermatophagoides pteronyssinus* and *D. farinae* results in decreased bronchial hyperreactivity. *Allergy* 1997;52:274–283.

Requests for reprints should be addressed to:
 Jay Portnoy, MD
 Section of Allergy, Asthma, & Immunology
 Children's Mercy Hospitals and Clinics
 2401 Gillham Road
 Kansas City, MO 64108
 E-mail: jportnoy@cmh.edu