

Evidence-Based Diagnostic Strategies for Evaluating Suspected Allergic Rhinitis

Karna Gendo, MD, and Eric B. Larson, MD, MPH

Allergic rhinitis is an increasingly common disease, with a prevalence of at least 10% to 25% in the United States. Diagnostic allergy tests, such as skin tests and in vitro tests, can assist clinicians in determining whether nasal symptoms are allergic in origin. In addition, safe and effective medications are available to treat allergic rhinitis.

The initial strategy should be to determine whether patients should undergo diagnostic testing or receive empirical treatment. This paper reviews the test characteristics of the history, skin tests, and in vitro tests in diagnosing allergic rhinitis from the perspective of decision thresholds. A combination of pertinent medical history features in a practice with a high baseline prevalence of allergic rhinitis justifies the common practice of empirical treatment since allergy medication has minimal toxicity and side effects.

The situation is more complex when the patient needs a diagnostic test, because reported sensitivities and specificities of skin tests and in vitro tests vary widely. As a result, it is difficult to calculate the post-test probability of allergic rhinitis with any confidence. The decision to initiate diagnostic testing must rely on clinical judgment to select patients who would benefit most from determining their allergic status while minimizing unnecessary testing and medications. Diagnosing allergy to a specific antigen allows patients to avoid the allergen and makes them candidates for allergen immunotherapy, which can decrease the need for medications.

Ann Intern Med. 2004;140:278-289.

www.annals.org

For author affiliations, see end of text.

Allergic rhinitis is an increasingly common disease, with a prevalence of at least 10% to 25% in the United States (1). Rhinitis is inflammation of the nasal membrane that causes periods of nasal discharge, sneezing, and congestion that persist for at least 1 hour per day (2). Rhinitis is considered allergic when allergen-specific IgE initiates the immunologic reaction that causes symptoms. The process of sensitization produces a specific IgE after T-cell release of interleukin-4 and interleukin-13, cell-to-cell signals, and B-cell switching to produce IgE rather than IgG (3, 4). Mast cells and basophils have receptors for this specific IgE. When allergens bind specific IgE on mast cell and basophil surfaces, these cells release histamine, leukotrienes, prostaglandins, cytokines, chemokines, tryptase, and other mediators that cause inflammation in the nasal membrane and symptoms.

Allergy diagnostic tests attempt to detect specific IgE that binds common allergens, such as dust mites, pollens, animal proteins, and mold spores. Clinicians must determine which patients should undergo diagnostic testing, which patients should receive empirical treatment, and which patients require only watchful waiting or nonspecific treatment. Empirical treatment is a common approach, especially since effective medications are available over the counter. If diagnostic testing is indicated, a clinician can either refer the patient to an allergy specialist for skin tests or order in vitro tests. The 2 types of in vitro tests are a panel that tests for the presence of specific IgE to any of the 6 most common allergens (Phadiatop test, Pharmacia-Upjohn, Uppsala, Sweden) and one or more specific IgE in vitro tests chosen specifically by the clinician because of a history of exposure to certain allergens.

We aim to compare these diagnostic testing and treatment strategies, to evaluate the quality of studies that examine diagnostic test performance of skin tests and in vitro

tests, and to formulate an evidence-based approach to the patient who presents to the general internist with nasal symptoms. We will use the principles of probability theory to show what we used to accomplish this goal.

METHODS

We searched MEDLINE from January 1995 to March 2003 (our search strategies are available by request) and reviewed references obtained from practice guidelines, review articles, and allergy specialists.

Articles on Diagnosing Allergic Rhinitis and Establishing a Gold Standard

We selected studies that performed a head-to-head comparison of skin tests and in vitro tests, with an emphasis on studies that performed nasal provocation tests in addition to skin tests and in vitro tests. We excluded studies that did not report findings for patients with and without allergic rhinitis. We excluded studies on children and non-English-language articles. We abstracted the total number of persons studied, sensitivities, specificities, and gold standard test used. The gold standard test is a reference test that should reliably measure the true state of the patients (that is, whether allergic rhinitis is present) (5).

Allergic rhinitis has been defined in the literature by 3 basic methods, which correspond to 3 gold standard tests. The first gold standard test relies on correlating the patient's symptoms and signs with clinical criteria established by expert opinion. This method is easy to use and applies to large numbers of patients but is prone to bias, especially if the clinician deciding whether a patient meets the gold standard criteria knows the results of the allergy test. Another source of error is missing data, which makes it difficult to be certain whether the patient meets diagnostic criteria. An advantage of this method is that clinical find-

ings usually influence treatment decisions; therefore, a person who meets the diagnostic criteria would be treated as if he or she had allergic rhinitis. Ideally, several physicians should independently review the data and assign the gold standard diagnosis. They should confer when they disagree or ask another physician to assign a diagnosis. For purposes of our review, the gold standard established by this method is referred to as the *clinical criteria gold standard*.

The second method uses a composite of the history and examination and one or more of the following: skin tests, in vitro tests, and nasal provocation tests. If necessary, additional in vitro tests or a second opinion by a different physician are used to reconcile discrepant results and arrive at a final diagnosis. This method is used most often in research. It has a serious shortcoming if the index test (for which sensitivity and specificity are being measured) is part of the composite gold standard because it will overestimate the correlation between the index test and the gold standard test and therefore overestimate sensitivity and specificity. A study that uses this method should specify the test that is used as part of the composite gold standard. Researchers should downgrade studies that use the index test as part of the gold standard test. This gold standard is referred to as the *composite gold standard*.

The third method uses a nasal challenge in which the nasal passages are exposed directly to an allergen in increasing concentrations. The patient may be prepared with nasal prewashes before the challenge. Symptoms or concentrations of mast cell mediators released into nasal lavage fluids are measured (6). This method most closely resembles the setting of disease. Certain challenge protocols have demonstrated that the threshold of sensitivity to nasal provocation statistically significantly correlates with symptom scores during the exposure to natural pollen (7). However, the methods and outcome measurements are not standardized, although recent guidelines have been proposed (8). In addition, higher doses of allergen than are found in nature are needed to elicit a response (9). This gold standard is referred to as the *challenge gold standard*.

The best gold standard, in our opinion, is the composite gold standard that does not include the index test. The other gold standards, including the composite gold standard that includes the index test and the nasal challenge gold standard, vary in quality depending on the quality of the study protocol. The clinical criteria gold standard seems least attractive unless there are strong measures to limit bias.

Some studies reported positive and negative likelihood ratios; for the remaining studies, we calculated likelihood ratios from the sensitivity and specificity. The likelihood ratio is the ratio of post-test odds after a test result to the pretest odds, and it indicates how much the odds change after a test. The positive likelihood ratio is the probability of a positive test result in patients with allergic rhinitis divided by the probability of a positive test result in patients without allergic rhinitis. The positive likelihood ratio

is sensitivity/(1 – specificity). The negative likelihood ratio is the probability of a negative test result in patients with allergic rhinitis divided by the probability of a negative test result in patients without allergic rhinitis. The negative likelihood ratio is (1 – sensitivity)/specificity (10). Likelihood ratios larger than 5.0 or less than 0.2 generate moderate to large shifts in disease probability, likelihood ratios of 2.0 to 5.0 and 0.5 to 0.2 generate small changes in probability, and likelihood ratios of 1.0 to 2.0 and 0.5 to 1.0 generate very small, often clinically insignificant changes in disease probability (11).

The pretest probability of allergic rhinitis depends on the overall prevalence of allergic rhinitis among patients with rhinitis and on the clinical setting and results of the history and examination. Several studies have measured the frequency of these clinical findings in patients with allergic rhinitis and in patients with nonallergic rhinitis. These studies use a gold standard test to establish a definite diagnosis of allergic rhinitis. The problems of gold standard tests in allergic rhinitis, which are noted in the preceding paragraphs, apply to these studies.

FREQUENCY OF CLINICAL FINDINGS AND TEST RESULTS IN PATIENTS WITH ALLERGIC RHINITIS

In the most recent population surveys, the prevalence of allergic rhinitis is more than 10% to 25% (1, 12). Precise epidemiologic data are difficult to obtain because of studies that use different definitions of allergic rhinitis. Peak current population prevalence is reported most often in young adults: 12% in people 16 to 24 years of age in a Michigan community and 23% in college freshman (12–14). Most of these data come from survey studies. The baseline prevalence of allergic rhinitis is probably higher among patients who report nasal symptoms to clinicians.

The prevalence of allergic rhinitis among patients with nasal symptoms was estimated in a study of 411 patients 16 to 65 years of age registered in a general practitioner's office in southwest London. Eligible participants identified from the practice register were asked to complete a postal screening questionnaire to identify those with rhinitis symptoms for at least 2 years. The questionnaires were mailed in March, April, and May, before the grass pollen season. A stratified sample of patients were interviewed at their homes in June, July, and August (15). The prevalence of seasonal allergic rhinitis due to tree and grass pollen in patients with seasonal symptoms was approximately 30% and 60%, respectively (15). The prevalence of allergic rhinitis due to dust mite and cat allergens in patients with perennial symptoms was approximately 30% (15).

Other studies of patients with nasal symptoms showed that patients who consult their generalist physicians for rhinitis have symptoms that interfere with activities of daily life or occupational life and they wait an average of 2.5 weeks after the onset of symptoms before consulting their physician (16). Symptoms typically occur between

Table 1. Accuracy of Medical History for the Diagnosis of Allergic Rhinitis*

Medical History Category	Question	Positive Likelihood Ratio	Negative Likelihood Ratio	Reference
Triggers	Are allergic nasal symptoms from pollen or animals?	6.69	0.15	14
	Do cats increase symptoms?	5.08	0.44	19
	Are allergic eye symptoms from pollen or animals?	4.43	0.09	14
	Are symptoms from animals?	4.22	0.34	20
	Do cats or dogs provoke or increase the nose problem?	4.21	0.81	21
	Do weeds increase symptoms?	6.53	0.65	19
	Do trees increase symptoms?	4.86	0.52	19
	Does grass increase symptoms?	3.95	0.40	19
	Are symptoms provoked by pollen?	2.52	0.49	22
	Do molds increase symptoms?	2.09	0.73	19
	Do house dust, house dust mites, or pollen provoke or increase the nose problem?	3.34	0.39	21
	Are symptoms provoked by dust?	1.76	0.49	19
		1.21	0.65	22
	Do you have symptoms when house cleaning?	2.33	0.73	20
	History	Have you had hay fever?	4.80	0.58
Have you been tested for allergy (puncture skin tests or in vitro tests) and were the results positive?		3.47	0.56	21
Have you been diagnosed with allergic rhinitis by a physician?		2.64	0.80	23
		1.94	0.13	14
	Has a doctor already diagnosed you with asthma, eczema, or allergic rhinitis?	2.08	0.76	21
Family history	Does any family member have asthma, eczema, or allergic rhinitis?	3.41	0.70	21
Months out of the year	Are symptoms worse in summer?	3.33	0.52	20
	Seasonal: In which of the past 12 months (or in which season) did this nose problem occur?	1.69	0.70	21
	Do you have seasonal exacerbation?	1.59	0.59	22
	Perennial: Did the nose problems occur in all of the past 12 months?	1.39	0.78	21
Eyes	Are nasal symptoms accompanied by itchy-watery eyes?	2.49	0.51	21
		1.33	0.50	22
	Were you told by your doctor that you have an allergic eye condition?	1.40	0.07	14
	Is your eye irritation apart from respiratory infections?	1.11	0.86	14
General questions	Do you think you are allergic?	3.12	0.33	21
	Have you ever had allergic nasal symptoms?	1.73	0.49	23
Nasal symptoms	Nasal symptoms in the past year, including sneezing, runny nose, or blocked nose: When you did not have a cold or flu?	1.33	0.65	14
		1.38	0.14	21
	Do you have diurnal variation of symptoms?	1.06	0.89	22

* All listed studies used a composite gold standard.

†Pharmacia-Upjohn, Uppsala, Sweden.

‡ Data from reference 24, with permission.

Table 1—Continued

Patients, n	Were Patients Seen Consecutively?	Was the Chief Symptom Listed for Entry into Study?	Did All Patients Receive Gold Standard Test?	Was the Gold Standard Test Interpreted Independent of Index Test?	Notes
290	No	Yes	Yes	No	
678	Yes	Yes	Yes	No	Computer-assisted
290	No	Yes	Yes	No	
100	Yes	Yes	Yes	No	
269	Yes	Yes	Yes	No	At least 1 trigger needed
678	Yes	Yes	Yes	No	Computer-assisted
678	Yes	Yes	Yes	No	Computer-assisted
678	Yes	Yes	Yes	No	Computer-assisted
466	Yes	Yes	Yes	No	
678	Yes	Yes	Yes	No	Computer-assisted
269	Yes	Yes	Yes	No	At least 1 trigger needed
678	Yes	Yes	Yes	No	Computer-assisted
466	Yes	Yes	Yes	No	
100	Yes	Yes	Yes	No	
466	Yes	Yes	Yes	No	
269	Yes	Yes	Yes	No	
400	No	Yes	Yes	Yes	The only study that did not use index test in gold standard test; gold standard test is the Phadiatop test†
290	No	Yes	Yes	No	
269	Yes	Yes	Yes	No	
269	Yes	Yes	Yes	No	
100	Yes	Yes	Yes	No	
269	Yes	Yes	Yes	No	Months listed should correlate with pollen calendar of the region
466					
269	Yes	Yes	Yes	No	
269	Yes	Yes	Yes	No	
466	Yes	Yes	Yes	No	
290	No	Yes	Yes	No	
290	No	Yes	Yes	No	
269	Yes	Yes	Yes	No	
400	No	Yes	Yes	Yes	
290	No	Yes	Yes	No	
269	Yes	Yes	Yes	No	At least 1 symptom; gold standard test based on definition of disease, according to European consensus guidelines‡
466	Yes	Yes	Yes	No	

Table 2. Accuracy of Skin Tests and in Vitro Tests for the Diagnosis of Allergic Rhinitis Reported in Studies That Do Not Include the Index Test in the Gold Standard

Test	Allergen	Reference	Positive Likelihood Ratio	Negative Likelihood Ratio	Sensitivity		Specificity	Gold Standard
					%			
Puncture skin test	Cat	27	4.93	0.08	93.6	81.0	Challenge	
	Cat	28	3.56	0.51	57	84	Clinical	
	Tree	29	16.17	0.03	97	94	Challenge	
	Grass	29	3.23	0.04	97	70	Challenge	
	Grass	28	6.82	0.28	75	89	Clinical	
	Mold	30	11.75	0.05	95.2	91.9	Challenge	
	Mite	31	4.06	0.03	97.4	76	Challenge	
Intradermal skin test	Mold (<i>Alternaria</i> sp.)	30	8.80	0.05	95.2	89.2	Challenge	
	Cat	27	0.89	1.24	60	32.4	Challenge	
	Cat	28	2.45	0.28	81	67	Clinical	
	Grass	32	1.05	0.98	33.3	68.4	Challenge	
In vitro test	Grass	28	4.94	0.25	79	84	Clinical	
	Cat	28	7.00	0.41	63	91	Composite	
	Cat	27	9.38	0.14	87.2	90.7	Challenge	
	Cat	28	3.39	0.48	61	82	Clinical	
	Tree	29	—	0.18	82	100	Challenge	
	Grass	28	80.0	0.20	80	99	Composite	
	Grass	29	3.13	0.09	94	70	Challenge	
	Grass	28	4.64	0.41	65	86	Clinical	
	Weed	28	20.33	0.40	61	97	Composite	
	Mold	28	14.00	0.46	56	96	Composite	
	Mold	30	15.0	0.2	81.0	94.6	Challenge	
	Mite	28	6.33	0.84	19	97	Composite	
	Mite	31	1.19	0.46	87.8	26.3	Challenge	
Phadiatop test*	Multiallergen	20	16.00	0.04	96	94	Composite	
	Multiallergen	33	3.95	0.28	77.1	80.5	Clinical	

* Pharmacia-Upjohn, Uppsala, Sweden.

March and May but can vary because of local pollination patterns (17, 18).

Medical History

Symptoms can help identify patients with allergic rhinitis (Table 1). Animal and pollen triggers are particularly helpful, as are the personal history and family history of allergy; the positive likelihood ratio ranged from 3 to 5 for each of these factors. No finding could definitely rule out allergic rhinitis since the negative likelihood ratios are not very low. The few studies that evaluated the physical examination for allergic rhinitis were not high quality, and we do not report their findings (25, 26).

All the studies evaluating the medical history, with one exception (23), use a composite gold standard that potentially includes the history item and probably will tend to overestimate the relationship between the history item and the result of the gold standard test. These estimates of sensitivity and specificity probably err on the high side for both, and therefore positive likelihood ratios are probably too high and negative likelihood ratios are probably too low.

The presence of atypical symptoms, such as pain, bleeding, fever, cough, unilateral symptoms, purulent discharge, headache, and dyspnea, indicates alternative or complicating conditions, such as sinusitis, nasal polyps, mechanical obstruction, tumors, granulomas, cerebrospinal rhinorrhea, nonallergic rhinitis, rhinitis medicamentosa, hypothyroidism, pregnancy, asthma, and medication effects (1).

Test Results

Table 2 and Appendix Table 1 (available at www.annals.org) describe performance characteristics of puncture skin tests, intradermal skin tests, and in vitro tests. Table 2 presents studies that exclude the index test from the gold standard test. Appendix Table 1 (available at www.annals.org) presents studies that potentially include the index test in the composite gold standard. As expected, these studies had considerably more favorable likelihood ratios.

For example, Wood and colleagues (27) evaluated puncture skin tests in a high-quality study that excluded the index test from the challenge gold standard. The reported positive and negative likelihood ratios of puncture

Table 2—Continued

Patients	Were Patients Seen Consecutively?	Was the Chief Symptom Listed for Entry into Study?	Did All Patients Receive Gold Standard Test?	Was the Gold Standard Test Interpreted Independent of Index Test?	Notes
<i>n</i>					
120	No	Yes	Yes	Yes	
181	Yes	Yes	Yes	Yes	
69	Yes	Yes	Yes	Yes	
69	Yes	Yes	Yes	Yes	
181	Yes	Yes	Yes	Yes	
58	No	Yes	Yes	Yes	
993	No	Yes	No	Yes	
58	No	Yes	Yes	Yes	
120	No	Yes	Yes	Yes	Data apply only to patients with negative puncture skin test results (flawed test design)
181	Yes	Yes	Yes	Yes	
28	Not stated	Yes	Yes	Yes	Studied a very specific sample: patients with history of seasonal allergic rhinitis and negative puncture skin test results
181	Yes	Yes	Yes	Yes	
181	Yes	Yes	Yes	Yes	Second-generation test
120	No	Yes	Yes	Yes	First- and second-generation tests
181	Yes	Yes	Yes	Yes	Second-generation test
69	Yes	Yes	Yes	Yes	First-generation test
181	Yes	Yes	Yes	Yes	Second-generation test
69	Yes	Yes	Yes	Yes	First-generation test
181	Yes	Yes	Yes	Yes	Second-generation test
181	Yes	Yes	Yes	Yes	Second-generation test
58	No	Yes	Yes	Yes	First-generation test
181	Yes	Yes	Yes	Yes	Second-generation test
60	No	Yes	No	Yes	First-generation test
100	Yes	Yes	Yes	Yes	Second-generation test
8329	No	Yes	Yes	Yes	Second-generation test

skin tests were 4.93 and 0.08, respectively. Crobach and colleagues (34) studied the same allergen and used a composite gold standard that included the index test (puncture skin test) in the gold standard. The reported positive and negative likelihood ratios of puncture skin tests were 126.0 and 0.25, respectively (34).

Escudero and colleagues (30) evaluated intradermal tests in a high-quality study that excluded the index test from the challenge gold standard. The reported positive and negative likelihood ratios were 8.80 and 0.05, respectively (30). However, 2 recent studies (27, 32) evaluating intradermal skin test performance by using a challenge gold standard reported positive and negative likelihood ratios near 1.0. In these 2 studies, intradermal skin tests were performed only on the participants with negative results on puncture skin tests. This practice is understandable since performing intradermal skin tests on patients with positive results on puncture skin tests can increase the risk for a systemic reaction to skin testing (38, 39). However, it is also a flaw in the study design.

Williams and colleagues (28) evaluated in vitro tests in

a high-quality study that excluded the index test from the composite gold standard. The reported positive likelihood ratio ranged from 6.33 to 80.0, and the negative likelihood ratio ranged from 0.20 to 0.84, depending on the allergen tested (28).

Eriksson (20) evaluated the multiallergen in vitro Phadiatop test in a high-quality study that excluded the index test from the composite gold standard. The reported positive and negative likelihood ratios were 16.00 and 0.04, respectively.

ALLERGY DIAGNOSTIC TESTS

Skin Tests

In puncture skin tests, a drop of allergen is placed on the volar surface of the forearm or on the back. About 15 to 25 allergens are tested, depending on the clinical situation. Allergen, a positive control histamine solution, and negative control saline solution are introduced into the epidermis by using a single-point, multipoint, or bifurcated needle. The use of control solutions detects inadvertent

Table 3. Causes of False-Positive and False-Negative Skin Test Results

Causes of false-positive skin test results
Test sites too close together (<2 cm)
Dermatographism
Irritant reaction from testing solution
Contamination of testing solution with another allergen
Injecting >0.05 mL of testing solution (intradermal tests)
Using high-concentration testing solution (intradermal tests)
Causes of false-negative skin test results
Use of antihistamines
Use of tricyclic antidepressants
Use of long-term oral steroid therapy
Use of topical steroids
Insufficient penetration of skin with needle
Low potency of testing extract
Age >50 y
Chronic renal insufficiency
Spinal cord injury
Peripheral nerve injury
Testing in the week after anaphylaxis
Subcutaneous injection (intradermal tests)

antihistamine ingestion and dermatographism (6). The allergen reacts with any available specific IgE on the mast cells, causing them to release histamine, which forms a wheal and surrounding erythema. The wheal diameter is measured in 15 to 20 minutes. Several semi-quantitative scales exist. The current consensus for a positive result is if a wheal caused by an allergen is at least 3 mm larger than the wheal caused by saline and is associated with surrounding erythema (40).

In intradermal skin tests, allergen (0.01 to 0.02 mL) is injected intracutaneously with a 26-gauge needle. The diameters of the wheal and flare are measured in 15 to 20 minutes. Intradermal skin tests are not necessary if the corresponding previous puncture skin test result is positive (6, 39). Skin testing is safe because the most common complication is a large local reaction; the risk for moderate to severe adverse reactions is 0.02% to 0.04% (41, 42).

False-positive skin test results can occur from dermatographism, placing the skin testing sites too close together, an excessively large test solution injection volume in intradermal testing, and irritation caused by the skin test solution. In addition, the presence of specific IgE in the patient does not indicate disease because a patient could possibly be asymptotically sensitized (6). False-negative test results can occur from poor skin testing technique and the recent use of antihistamines, tricyclic antidepressants, or topical steroids (Table 3) (38). Each patient has his or her own characteristic skin reactivity, which can affect results as well (6).

In Vitro Tests

Examples of the first-generation in vitro tests for specific IgE are the Phadebas RAST (Pharmacia-Upjohn) and the enzyme-linked allergosorbent test. Specific allergen is bound covalently onto cyanogen bromide-activated cellulose disks. Human serum is added to the disk, allowing any

specific IgE to bind to the allergen. This bound, specific IgE is detected with ^{125}I or enzyme-labeled antihuman IgE detection antibody. The resulting radioactive counts per minute from the ^{125}I or the residual optical density from the enzyme colorimetric reaction are proportional to the amount of specific IgE in the original serum. When the response is compared with that of standard reference serum, one can calculate the amount of specific IgE and report the level in arbitrary units or classes (6).

The second-generation in vitro tests improve on the first-generation in vitro tests and generate higher test sensitivities and specificities. ImmunoCAP (Quest Diagnostics, Teterboro, New Jersey) is a matrix capsule used in the Pharmacia CAP system (Pharmacia-Upjohn) that contains high-quality allergen bound to a hydrophilic carrier. This matrix enhances IgE-binding capacity and lowers nonspecific IgE binding. The Pharmacia CAP system consists of the particular immunoassay reagents, instrumentation, and information management software developed for measuring specific and total IgE. Automation using nonradioactive polyclonal and monoclonal anti-IgE detection antibody systems has made these in vitro tests safer and easier to perform. Calibration systems using the total serum IgE curve allow the estimation of the quantity of specific IgE in the serum (6). A value greater than 0.35 kU/L is defined as a positive result. In vitro test panels that contain dust mite, animal, mold, and regional pollen allergens are available.

Levels of specific IgE measured by different commercial assays are not equivalent because each commercial assay differs in the composition of allergen reagents, methods of measurement, and standardization procedures (6). A recent study of 6 diagnostic laboratories using 5 different types of in vitro tests in performing 7800 tests found that different assays were generally not concordant. The second-generation Pharmacia CAP system was the exception, demonstrating similar and precise results from 2 different laboratories (43). Clinicians should determine the exact type of in vitro test used by the diagnostic laboratory and request diagnostic performance characteristics (sensitivity, specificity, and type of gold standard) of the in vitro test used from the diagnostic laboratory.

The Phadiatop test is a multiallergen, second-generation in vitro test that contains a mixture of several allergens bound in the matrix to detect whether any of the most common specific IgE types are present in the serum. It reports either a positive or negative result for a statistically significant level of specific IgE but does not detect the presence of a particular specific IgE type (20). Therefore, skin testing or in vitro tests are performed afterward to identify the relevant allergen, if necessary.

False-negative Phadiatop test results can occur if the patient is sensitized to an allergen that is not present in the mixture of allergens; however, most allergic patients have multiple sensitizations. False-positive test results occurred in the Phadiatop paper-disk system (which was an analogous first-generation system) because some asymptomatic patients

had low levels of multiple specific IgE types that reacted in summation with one disk to create a positive result (37).

MAKING DECISIONS ABOUT TESTING AND EMPIRICAL TREATMENT

Threshold Model of Decision Making: To Test, Treat Empirically, or Wait

The threshold model of decision making helps to decide when tests will help and when either empirical treatment or watchful waiting is preferred. It uses a patient's pretest probability of disease, the accuracy and harms of the test, and the treatment threshold. The treatment threshold is a useful concept for linking diagnostic certainty and the harms and benefits of treatment to a treatment decision. It is the probability of disease above which to give treatment and below which to withhold treatment. Probability theory states that the treatment threshold is $\text{harms}/(\text{harms} + \text{benefits})$, where the benefits and harms are those of the treatment (44).

Testing is indicated only if the test result could alter the decision to treat. Since allergy testing is safe, the harms of testing are not a factor in the testing decision. Empirical treatment is indicated if the pretest probability is above the treatment threshold and a negative test result cannot lower the probability below the treatment threshold. Watchful waiting is indicated if the probability is low and a test cannot increase the probability enough to exceed the treatment threshold.

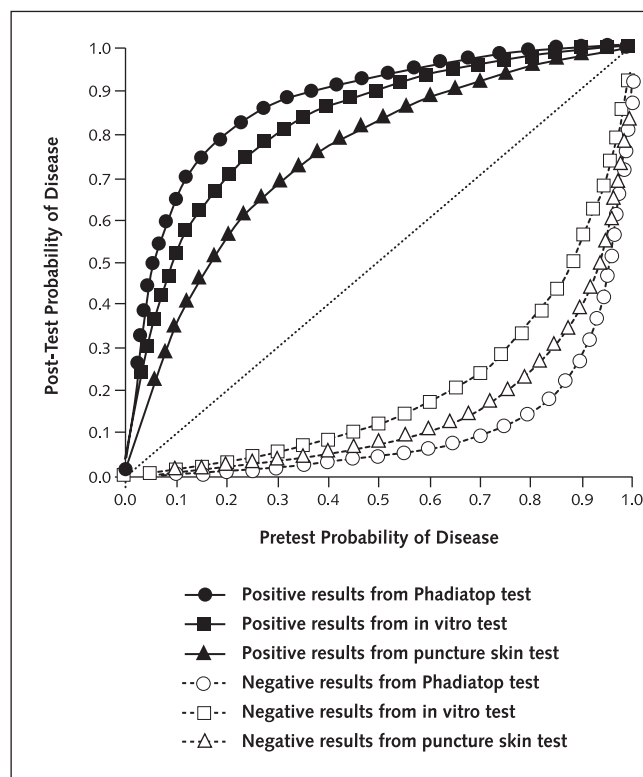
Synthesizing Recommendations

We have provided information to assist in deciding among empirical treatment, allergy testing, and watchful waiting. This information is admittedly imprecise and based on insufficiently good-quality evidence to support definitive recommendations. At the same time, clinicians must make practical decisions with the information available.

The prevalence of allergic rhinitis among patients presenting with seasonal nasal symptoms is probably more than 50%, but we cannot provide guidance about making a precise probability estimate without going beyond the evidence. The clinician must form his or her own subjective probability estimate on the basis of the information we have provided and clinical judgment about a particular patient. The results of good-quality studies of test performance vary widely; we can report that a positive test result increases the odds for allergic rhinitis by at least 3-fold, but the likelihood ratios after a negative test result are too variable to make any generalization. The treatment threshold probability for medication is probably low, on the basis of high benefit and low harm of treatment. However, a decision analysis to determine this treatment threshold or analyses for more complex allergy treatments have not been performed.

Because of these uncertainties, we are reluctant to make any firm recommendations about when to test. In-

Figure. The effect of prevalence on the predictive values of diagnostic tests using diagnostic test performance data.



To use this graph, estimate the pretest probability of allergic rhinitis and decide whether that probability justifies empirical treatment. Then, note the post-test probabilities corresponding to the pretest probability after positive and negative test results (*vertical axis*). If both probabilities justify the same action, testing will not change the decision about treatment and a choice (empirical treatment or watchful waiting) based on the pretest probability alone should be made. Data from Wood and colleagues (27) and Eriksson (20), with permission.

stead, we have graphed the post-test probability for 3 types of tests to help the practitioner interpret tests and perhaps when to perform testing in a specific patient (**Figure**). To make the calculations, we used high-quality studies of test performance; however, we caution against assuming that the post-test probabilities are typical of results in actual practice.

Effect of Testing on Treatment Decisions: Three Examples

Allergy tests are most useful in making a diagnosis in patients with a low pretest probability of allergic rhinitis because the positive likelihood ratios are consistently high (**Table 2**). These tests have the largest effect on disease probability when the pretest probability is low to intermediate (15% to 55%) (**Figure** and **Appendix Table 2** [available at www.annals.org]). Since the negative likelihood ratios are inconsistent but seldom below 0.3, we are not confident that negative test results lower the probability of allergic rhinitis past the treatment threshold in patients with intermediate to high pretest probabilities. Therefore, these tests will make the biggest difference in diagnostic

Table 4. Treatment Approach in Adolescents and Adults*

Management Based on Classification	Initial Treatment Options	Treatment Options on Follow-up Visit (2–4 wk)	Consider Immunotherapy?
Mild intermittent rhinitis	Oral antihistamine Nasal antihistamine Decongestant		No
Moderate to severe intermittent rhinitis	Oral antihistamine Nasal antihistamine Decongestant Nasal steroid Nasal cromolyn		Yes
Mild persistent rhinitis	Oral antihistamine Nasal antihistamine Decongestant Nasal steroid Nasal cromolyn	If improved, continue for 1 mo; if no response, add a medication	Yes
Moderate to severe persistent rhinitis	Nasal steroid	If improved, step down therapy and continue treatment for 1 mo; if no response, evaluate for infection, review diagnosis, assess adherence, and consider: Increasing nasal steroid dose Adding antihistamine for itching and sneezing Adding nasal ipratropium for rhinorrhea Adding decongestant or short-term oral steroid for congestion If congestion still persists, consider a surgical referral	Yes

* Adapted from Bousquet and colleagues (1).

thinking when the level of diagnostic suspicion is relatively low and a positive test result could influence a decision to treat.

Empirical Medication

In the case of medication, nonsedating antihistamines and nasal steroids are relatively free of side effects, significantly reduce symptoms, and are reasonably inexpensive (2, 45, 46). Therefore, benefits substantially exceed harms and the treatment threshold for medication should be relatively low (probably lower than 25%). The presence of triggers or a pertinent personal or family history of allergy (Table 1) can increase the pretest probability of allergic rhinitis, exceeding the treatment threshold by so much that a negative test result would not lower the probability enough to change the treatment decision. In this instance,

empirical treatment is indicated. Since the baseline prevalence of allergic rhinitis among patients with nasal symptoms is 30% to 60%, most patients with triggers or a personal or family history would qualify for empirical medication.

Evidence-based guidelines developed in collaboration with the World Health Organization can assist the clinician in selecting appropriate medications (Tables 4 and 5) (1). The first step is to classify the allergic rhinitis on the basis of the severity of illness and frequency of symptoms. Moderate to severe allergic rhinitis affects sleep and activities of daily living. Persistent allergic rhinitis lasts for more than 4 days per week and more than 4 weeks per year (Table 6).

Oral, nonsedating antihistamines or nasal antihistamine sprays, which are effective for sneezing, rhinor-

Table 5. Effects of Medication Treatment on Treatment*

Medication	Symptoms				
	Sneezing	Rhinorrhea	Nasal Congestion	Nasal Itch	Eye Symptoms
Antihistamine, oral	++	++	+	+++	++
Antihistamine, nasal	++	++	+	++	0
Antihistamine, ocular	0	0	0	0	+++
Nasal steroids	+++	+++	+++	++	++
Cromolyn, nasal	+	+	+	+	0
Cromolyn, ocular	0	0	0	0	++
Decongestants, oral	0	0	+	0	0
Decongestants, nasal†	0	0	++++	0	0
Anticholinergics, nasal	0	++	0	0	0
Antileukotrienes	0	+	++	0	++

* Adapted from Bousquet and colleagues (1) and van Cauwenberge and colleagues (47), with permission. Note: Consult package insert for cautions and contraindications. 0 = not effective for symptoms; + = mildly effective for symptoms; ++ = moderately effective for symptoms; +++ and ++++ = very effective for symptoms.

† Do not use > 3–5 d.

rhea, itch, and eye symptoms, are recommended for mild intermittent allergic rhinitis. Nasal steroids are effective for all symptoms, including congestion, and are useful for moderate to severe intermittent allergic rhinitis or mild persistent allergic rhinitis. Nasal steroids are first-line treatment for moderate to severe persistent allergic rhinitis. A 2- to 4-week trial suffices to decide whether treatment is working.

Allergen Avoidance and Specific Immunotherapy: When Testing Guides Treatment

For patients who do not respond to medication, treatment can incorporate allergen avoidance and specific immunotherapy. In this setting, the results of allergy testing inform the patient on what to avoid and inform the clinician on what allergens to include in a patient's specific immunotherapy formula. The purpose of allergy testing is no longer to diagnose but to inform about highly specific treatment.

Since patients often misidentify what they are allergic to, the results of allergy tests inform patients on what to avoid (19–22, 48). Dust mite avoidance measures include washing bed linens in hot water (130 °F); regular vacuuming; and encasing the pillows, mattress, and box springs with mite-proof casings. Removing upholstered furniture, wall-to-wall carpeting, chenille spreads, and stuffed toys from the bedroom may achieve more control (2). Other allergen-avoidance measures include removing pets from the bedroom (preferably removing them from the house altogether) and regularly cleaning areas of the home that promote mold growth, such as shower stalls, basements, and windowsills (49). If symptoms persist after these avoidance measures are used, specific immunotherapy may be considered.

Specific immunotherapy consists of subcutaneous injections of allergen for at least 3 to 5 years that gradually enable the patient to tolerate exposure to the allergen and potentially reduce the long-term use of medications. For severe intermittent (seasonal) allergic rhinitis that is resistant to medications, specific immunotherapy statistically significantly reduces symptoms during treatment and for up to 6 years after discontinuing immunotherapy (50–52).

Puncture Skin Tests or Second-Generation *In Vitro* Tests

The traditional role of *in vitro* tests has been for patients who cannot omit medications that interfere with skin tests; who have extensive skin disease or dermatographism; or who have relative contraindications to skin testing, such as pregnancy, unstable asthma, or use of β -blockers (39, 53). Our results show that puncture skin tests; intradermal skin tests; and well-performed second-generation *in vitro* tests, such as the Pharmacia CAP or UniCAP systems (Pharmacia-Upjohn), have similar diagnostic performance. Clinicians may prefer using skin tests to determine the components of a patient's immunotherapy for

Table 6. Classification of Allergic Rhinitis from the Allergic Rhinitis and Its Impact on Asthma and World Health Organization Guidelines*

Classification	Symptoms
Intermittent Mild	<4 d/wk or <4 wk Normal sleep <i>and</i> no impairment of daily activities, sport, or leisure; normal function at work or school; and no troublesome symptoms
Persistent Moderate to severe	4 d/wk and >4 wk One or more: Abnormal sleep Impaired daily activities, sport, or leisure Abnormal function at work or school Troublesome symptoms

* Data from Bousquet and colleagues (1), with permission.

mula since most clinical trials documenting efficacy of immunotherapy used skin tests to determine a patient's allergic status; fewer data on the efficacy of specific immunotherapy based on *in vitro* test results are available (54, 55). However, current practice parameters published by the American Academy of Allergy, Asthma, & Immunology and the American College of Allergy, Asthma & Immunology support the use of either skin tests or well-performed *in vitro* tests to determine the individual components of a patient's specific immunotherapy (55).

In patients with a very low pretest probability, clinicians may prefer to use the Phadiatop test to rule out allergy because it is a single test with a negative likelihood ratio of 0.04 (20). However, if a history of a trigger, pertinent personal history, or a positive family history is elicited, skin tests or specific *in vitro* tests are better initial choices because these test results inform patients about their specific allergies.

LIMITATIONS

One limitation in studies reporting sensitivities and specificities is the use of different gold standard tests. Studies using gold standard tests that use only history findings and physical examination generate sensitivities and specificities that are lower than those of studies that incorporate specific IgE testing in their gold standards. The gold standard using a composite of the history, examination, and specific IgE testing was the most commonly used, but the methods of combining results varied widely.

Another limitation is that most studies draw from a sample of patients referred to an allergy specialist. This referred sample probably has a higher pretest probability of allergic rhinitis than that of patients presenting to the primary provider with nasal symptoms. This affects the generalizability of the results.

A third limitation is that definitions of positive results of skin tests and *in vitro* tests vary in the literature. For example, many skin test studies compare the diameters of the wheals generated by allergens with the diameter of the

wheel generated by the histamine control solution to classify test results as positive.

These limitations generate variability in reported sensitivities and specificities. However, it may not be necessary to have absolutely precise estimates of allergy test diagnostic performance because allergic rhinitis is a nonfatal disease with safe and effective treatments. However, relatively unbiased test performance characteristics can help to identify the circumstances in which testing should affect management.

RECOMMENDATIONS

The prevalence of allergic rhinitis in patients presenting to their primary care provider with nasal symptoms is estimated to be 30% to 60%. The disease probability is clinically significantly increased if the patient has animal or pollen triggers, a history of allergy, or a family history of allergy. For most patients, the pretest probability of allergic rhinitis is well above the treatment threshold since allergy medication is relatively safe and effective.

Diagnostic allergy testing is indicated if the test result could alter the decision to treat. These tests have consistently high likelihood ratios when test results are positive and relatively high likelihood ratios when test results are negative. Therefore, testing will greatly affect diagnostic thinking when the suspicion is relatively low and will not particularly help when suspicion is high. We believe this reasoning justifies empirical medication as an initial step for patients with a convincing history of allergy.

When patients do not respond to medication and treatment must be increased to incorporate allergen avoidance and specific immunotherapy, it is important to know what the patient is allergic to. In this setting, allergy testing guides treatment.

The diagnostic performance of well-done, second-generation *in vitro* tests is similar to that of skin tests, provided that clinicians are aware of the exact type of *in vitro* test used by the laboratory. Clinicians may prefer using skin tests to determine what components to include in a patient's specific immunotherapy formula or may prefer using the Phadiatop test to evaluate patients with a very low pretest probability because it is a single test.

From the Northwest Asthma and Allergy Center and The Center for Health Studies, Group Health Cooperative, University of Washington, Seattle, Washington.

Acknowledgments: The authors thank Gail Shapiro, MD; John Sheffield, MD; and Michael Kennedy, MD, for their thoughtful reviews of the manuscript and helpful suggestions.

Potential Financial Conflicts of Interest: None disclosed.

Requests for Single Reprints: Eric B. Larson, MD, MPH, The Center for Health Studies, Group Health Cooperative, 1730 Minor Avenue, Suite 1600, Seattle, WA 98101-1448.

Current author addresses are available at www.annals.org.

References

- Bousquet J, van Cauwenberge P, Khaltaev N, Bousquet J. Allergic rhinitis and its impact on asthma. In collaboration with the World Health Organization. Executive summary of the workshop report. 7-10 December 1999, Geneva, Switzerland. *Allergy*. 2002;57:841-55. [PMID: 12169183]
- Ricketti AJ. Allergic rhinitis. In: Grammer LC, Greenberger PA, eds. *Patterson's Allergic Diseases*. 6th ed. Philadelphia: Lippincott Williams & Wilkins; 2002:159-82.
- Melvold RW. Review of immunology. In: Grammer LC, Greenberger PA, eds. *Patterson's Allergic Diseases*. 6th ed. Philadelphia: Lippincott Williams & Wilkins; 2002:1-24.
- Zeiss CR, Pruzansky JJ. Immunology of IgE-mediated and other hypersensitivity states. In: Grammer LC, Greenberger PA, eds. *Patterson's Allergic Diseases*. 6th ed. Philadelphia: Lippincott Williams & Wilkins; 2002:43-54.
- Gordis L. *Assessing the Validity and Reliability of Diagnostic and Screening Tests*. Epidemiology. 2nd ed. Philadelphia: WB Saunders; 2000.
- Hamilton RG, Adkinson NF Jr. 23. Clinical laboratory assessment of IgE-dependent hypersensitivity. *J Allergy Clin Immunol*. 2003;111:S687-701. [PMID: 12592314]
- Bousquet J, Maasch H, Martinot B, Hejaoui A, Wahl R, Michel FB. Double-blind, placebo-controlled immunotherapy with mixed grass-pollen allergoids. II. Comparison between parameters assessing the efficacy of immunotherapy. *J Allergy Clin Immunol*. 1988;82:439-46. [PMID: 3170992]
- Malm L, Gerth van Wijk R, Bachert C. Guidelines for nasal provocations with aspects on nasal patency, airflow, and airflow resistance. International Committee on Objective Assessment of the Nasal Airways, International Rhinologic Society. *Rhinology*. 2000;38:1-6. [PMID: 10780040]
- Bousquet J, Lebel B, Dhivert H, Bataille Y, Martinot B, Michel FB. Nasal challenge with pollen grains, skin-prick tests and specific IgE in patients with grass pollen allergy. *Clin Allergy*. 1987;17:529-36. [PMID: 3436032]
- Go AS. Refining probability: an introduction to the use of diagnostic tests. In: Friedland DJ, ed. *Evidence-Based Medicine: A Framework for Clinical Practice*. Stamford, CT: Appleton and Lange; 1998:11-34.
- Jaeschke R, Guyatt GH, Sackett DL. Users' guides to the medical literature. III. How to use an article about a diagnostic test. B. What are the results and will they help me in caring for my patients? The Evidence-Based Medicine Working Group. *JAMA*. 1994;271:703-7. [PMID: 8309035]
- Sly RM. Changing prevalence of allergic rhinitis and asthma. *Ann Allergy Asthma Immunol*. 1999;82:233-48; quiz 248-52. [PMID: 10094214]
- Greisner WA 3rd, Settignano RJ, Settignano GA. Natural history of hay fever: a 23-year follow-up of college students. *Allergy Asthma Proc*. 1998;19:271-5. [PMID: 9801740]
- Kilpelainen M, Terho EO, Helenius H, Koskenvuo M. Validation of a new questionnaire on asthma, allergic rhinitis, and conjunctivitis in young adults. *Allergy*. 2001;56:377-84. [PMID: 11350300]
- Sibbald B, Rink E. Epidemiology of seasonal and perennial rhinitis: clinical presentation and medical history. *Thorax*. 1991;46:895-901. [PMID: 1792637]
- Demoly P, Allaert FA, Lecasle M. ERASM, a pharmacoepidemiologic survey on management of intermittent allergic rhinitis in every day general medical practice in France. *Allergy*. 2002;57:546-54. [PMID: 12028122]
- Wuthrich B, Brignoli R, Canevascini M, Gerber M. Epidemiological survey in hay fever patients: symptom prevalence and severity and influence on patient management. *Schweiz Med Wochenschr*. 1998;128:139-43. [PMID: 9522418]
- Solomon WR. Airborne pollen prevalence in the United States. In: Grammer LC, Greenberger PA, eds. *Patterson's Allergic Diseases*. 6th ed. Philadelphia: Lippincott Williams & Wilkins; 2002:131-44.
- Lindberg S, Malm L. Comparison of allergic rhinitis and vasomotor rhinitis patients on the basis of a computer questionnaire. *Allergy*. 1993;48:602-7. [PMID: 8116859]
- Eriksson NE. Allergy screening with Phadiatop and CAP Phadiatop in combination with a questionnaire in adults with asthma and rhinitis. *Allergy*. 1990;45:285-92. [PMID: 2382793]
- Annesi-Maesano I, Didier A, Klossek M, Chanal I, Moreau D, Bousquet J.

- The score for allergic rhinitis (SFAR): a simple and valid assessment method in population studies. *Allergy*. 2002;57:107-14. [PMID: 11929412]
22. **Charpin D, Sibbald B, Weeke E, Wuthrich B.** Epidemiologic identification of allergic rhinitis. *Allergy*. 1996;51:293-8. [PMID: 8836332]
23. **Kauppi P, Laitinen IA, Laitinen H, Kere J, Laitinen T.** Verification of self-reported asthma and allergy in subjects and their family members volunteering for gene mapping studies. *Respir Med*. 1998;92:1281-8. [PMID: 9926141]
24. International Consensus Report on the diagnosis and management of rhinitis. International Rhinitis Management Working Group. *Allergy*. 1994;49:1-34. [PMID: 8080072]
25. **Ng ML, Warlow RS, Chrishanthan N, Ellis C, Walls R.** Preliminary criteria for the definition of allergic rhinitis: a systematic evaluation of clinical parameters in a disease cohort (I). *Clin Exp Allergy*. 2000;30:1314-31. [PMID: 10971479]
26. **Joko H, Hyodo M, Gyo K, Yumoto E.** Chromametric assessment of nasal mucosal color and its application in patients with nasal allergy. *Am J Rhinol*. 2002;16:11-6. [PMID: 11895189]
27. **Wood RA, Phipatanakul W, Hamilton RG, Eggleston PA.** A comparison of skin prick tests, intradermal skin tests, and RASTs in the diagnosis of cat allergy. *J Allergy Clin Immunol*. 1999;103:773-9. [PMID: 10329809]
28. **Williams PB, Dolen WK, Koepke JW, Selner JC.** Comparison of skin testing and three in vitro assays for specific IgE in the clinical evaluation of immediate hypersensitivity. *Ann Allergy*. 1992;68:35-45. [PMID: 1736718]
29. **Petersson G, Dreborg S, Ingestad R.** Clinical history, skin prick test and RAST in the diagnosis of birch and timothy pollinosis. *Allergy*. 1986;41:398-407. [PMID: 3789325]
30. **Escudero AI, Sanchez-Guerrero IM, Mora AM, Soriano V, Lopez JD, Garcia FJ, et al.** Cost-effectiveness of various methods of diagnosing hypersensitivity to *Alternaria*. *Allergol Immunopathol (Madr)*. 1993;21:153-7. [PMID: 8237720]
31. **Clarke PS.** The diagnosis of perennial rhinitis due to house dust mite (*Dermatophagoides pteromyssinus*) demonstrated by nasal provocation tests. *Ann Allergy*. 1987;59:25-8. [PMID: 3605794]
32. **Nelson HS, Oppenheimer J, Buchmeier A, Kordash TR, Freshwater LL.** An assessment of the role of intradermal skin testing in the diagnosis of clinically relevant allergy to timothy grass. *J Allergy Clin Immunol*. 1996;97:1193-201. [PMID: 8648012]
33. **Tschopp JM, Sistek D, Schindler C, Leuenberger P, Perruchoud AP, Wuthrich B, et al.** Current allergic asthma and rhinitis: diagnostic efficiency of three commonly used atopic markers (IgE, skin prick tests, and Phadiatop). Results from 8329 randomized adults from the SAPALDIA Study. Swiss Study on Air Pollution and Lung Diseases in Adults. *Allergy*. 1998;53:608-13. [PMID: 9689343]
34. **Crobach MJ, Hermans J, Kaptein AA, Ridderikhoff J, Petri H, Mulder JD.** The diagnosis of allergic rhinitis: how to combine the medical history with the results of radioallergosorbent tests and skin prick tests. *Scand J Prim Health Care*. 1998;16:30-6. [PMID: 9612876]
35. **Plebani M, Borghesan F, Faggian D.** Clinical efficiency of in vitro and in vivo tests for allergic diseases. *Ann Allergy Asthma Immunol*. 1995;74:23-8. [PMID: 7719878]
36. **Williams PB, Siegel C, Portnoy J.** Efficacy of a single diagnostic test for sensitization to common inhalant allergens. *Ann Allergy Asthma Immunol*. 2001;86:196-202. [PMID: 11258690]
37. **Matricardi PM, Nisini R, Pizzolo JG, D'Amelio R.** The use of Phadiatop in mass-screening programmes of inhalant allergies: advantages and limitations. *Clin Exp Allergy*. 1990;20:151-5. [PMID: 2357615]
38. **Demoly P, Michel E, Bousquet J.** In vivo methods for study of allergy skin tests, techniques, and interpretation. In: Middleton E, Ellis EF, eds. *Allergy Principles & Practice*. 5th ed. St. Louis: Mosby; 1998:430-9.
39. **Tripathi A, Booth BH.** Diagnosis of immediate hypersensitivity. In: Grammer LC, Greenberger PA, eds. *Patterson's Allergic Diseases*. 6th ed. Philadelphia: Lippincott Williams & Wilkins; 2002:145-59.
40. **Bernstein IL, Storms WW.** Practice parameters for allergy diagnostic testing. Joint Task Force on Practice Parameters for the Diagnosis and Treatment of Asthma. The American Academy of Allergy, Asthma and Immunology and the American College of Allergy, Asthma and Immunology. *Ann Allergy Asthma Immunol*. 1995;75:543-625. [PMID: 8521115]
41. **Ten RM, Klein JS, Frigas E.** Allergy skin testing. *Mayo Clin Proc*. 1995;70:783-4. [PMID: 7630219]
42. **Lin MS, Tanner E, Lynn J, Friday GA Jr.** Nonfatal systemic allergic reactions induced by skin testing and immunotherapy. *Ann Allergy*. 1993;71:557-62. [PMID: 8267250]
43. **Szeinbach SL, Barnes JH, Sullivan TJ, Williams PB.** Precision and accuracy of commercial laboratories' ability to classify positive and/or negative allergen-specific IgE results. *Ann Allergy Asthma Immunol*. 2001;86:373-81. [PMID: 11345278]
44. **Bent S, Friedland DJ.** Treatment and testing thresholds. In: Friedland DJ, ed. *Evidence-based Medicine: A Framework for Clinical Practice*. Stamford, CT: Appleton and Lange; 1998:59-83.
45. **Day J, Carrillo T.** Comparison of the efficacy of budesonide and fluticasone propionate aqueous nasal spray for once daily treatment of perennial allergic rhinitis. *J Allergy Clin Immunol*. 1998;102:902-8. [PMID: 9847429]
46. **Meltzer EO, Grant JA.** Impact of cetirizine on the burden of allergic rhinitis. *Ann Allergy Asthma Immunol*. 1999;83:455-63. [PMID: 10582731]
47. **van Cauwenberge P, Bachert C, Passalacqua G, Bousquet J, Canonica GW, Durham SR, et al.** Consensus statement on the treatment of allergic rhinitis. European Academy of Allergology and Clinical Immunology. *Allergy*. 2000;55:116-34. [PMID: 10726726]
48. **Li JT, Andrist D, Bamlet WR, Wolter TD.** Accuracy of patient prediction of allergy skin test results. *Ann Allergy Asthma Immunol*. 2000;85:382-4. [PMID: 11101180]
49. **Brooks GD, Bush RK.** Allergens and other factors important in atopic disease. In: Grammer LC, Greenberger PA, eds. *Patterson's Allergic Diseases*. 6th ed. Philadelphia: Lippincott Williams & Wilkins; 2002:81-122.
50. **Frew AJ.** Immunotherapy of allergic disease. *J Allergy Clin Immunol*. 2003;111:S712-9. [PMID: 12592316]
51. **Durham SR, Walker SM, Varga EM, Jacobson MR, O'Brien F, Noble W, et al.** Long-term clinical efficacy of grass-pollen immunotherapy. *N Engl J Med*. 1999;341:468-75. [PMID: 10441602]
52. **Varney VA, Gaga M, Frew AJ, Aber VR, Kay AB, Durham SR.** Usefulness of immunotherapy in patients with severe summer hay fever uncontrolled by antiallergic drugs. *BMJ*. 1991;302:265-9. [PMID: 1998791]
53. **Ownby DR.** Diagnostic tests in allergy. In: Lieberman P, Anderson JA, eds. *Allergic Diseases, Diagnosis and Treatment*. 2nd ed. Totowa, NJ: Humana; 2000:29-41.
54. **Nalebuff DJ, Fadal RG.** RAST-based immunotherapy. *Rhinology*. 1984;22:11-9. [PMID: 6729355]
55. **Allergen immunotherapy: a practice parameter.** American Academy of Allergy, Asthma and Immunology. American College of Allergy, Asthma and Immunology. *Ann Allergy Asthma Immunol*. 2003;90:1-40. [PMID: 12542071]

Appendix Table 1. Accuracy of Puncture Skin Tests and In Vitro Tests for the Diagnosis of Allergic Rhinitis Reported in Studies That Potentially Included the Index Test in the Gold Standard

Test	Allergen	Positive Likelihood Ratio	Negative Likelihood Ratio	Sensitivity	Specificity	Gold Standard	
				%			
Puncture skin test	Cat	126.0	0.25	75.6	99.4	Composite	
	Tree	64.1	0.17	83.3	98.7	Composite	
	Grass		52.9	0.05	95.3	98.2	Composite
			55.56	0.00	100	98.2	Composite
	Weed	47.0	0.58	42.3	99.1	Composite	
	Mold		250	0.25	75.0	99.7	Composite
			—	—	100	100	Composite
	Mite	48.6	0.03	97.1	98.0	Composite	
		26.83	0.06	93.9	96.5	Composite	
	In vitro test	Cat	292.7	0.12	87.8	99.7	Composite
Tree		305.0	0.09	91.5	99.7	Composite	
Grass			232.25	0.07	92.9	99.6	Composite
			—	—	100	100	Composite
Weed		66.67	0	100	98.5	Composite	
Mold			109.4	0.13	87.5	99.2	Composite
			—	—	100	100	Composite
Mite		—	0.07	93.1	100	Composite	
Phadiatop test*	Multiallergen	—	—	100	100	Composite	
		7.65	0.01	98.7	87.1	Composite	

* Pharmacia-Upjohn, Uppsala, Sweden.

Current Author Addresses: Dr. Gendo: Northwest Asthma and Allergy Center, 4540 Sand Point Way NE, Seattle, WA 98105-3941.

Dr. Larson: The Center for Health Studies, Group Health Cooperative, 1730 Minor Avenue, Suite 1600, Seattle, WA 98101-1448.

Appendix Table 1—Continued

Reference	Patients, n	Were Patients Seen Consecutively?	Was the Chief Symptom Listed for Entry into Study?	Did All Patients Receive Gold Standard Test?	Was the Gold Standard Test Interpreted Independent of Index Test?	Notes
34	365	Yes	Yes	Yes	No	
34	365	Yes	Yes	Yes	No	
34	365	Yes	Yes	Yes	No	
35	106	Not stated	Yes	Yes	No	
34	365	Yes	Yes	Yes	No	
34	365	Yes	Yes	Yes	No	
35	106	Not stated	Yes	Yes	No	
34	365	Yes	Yes	Yes	No	
35	106	Not stated	Yes	Yes	No	
34	365	Yes	Yes	Yes	No	Second-generation test
34	365	Yes	Yes	Yes	No	Second-generation test
34	365	Yes	Yes	Yes	No	Second-generation test
35	106	Not stated	Yes	Yes	No	Second-generation test
34	365	Yes	Yes	Yes	No	Second-generation test
34	365	Yes	Yes	Yes	No	Second-generation test
35	106	Not stated	Yes	Yes	No	Second-generation test
34	365	Yes	Yes	Yes	No	Second-generation test
35	106	Not stated	Yes	Yes	No	Second-generation test
34	365	Yes	Yes	Yes	No	Second-generation test
35	106	Not stated	Yes	Yes	No	Second-generation test
36	181	Yes	Yes	Yes	No	Second-generation test
37	600	No	No	Yes	No	Second-generation test

Appendix Table 2. Allergic Rhinitis Probability after Puncture Skin Tests and in Vitro Tests Using Diagnostic Test Performance Data*

Pretest Probability	Post-Test Probability											
	If Puncture Skin Test Result Is Positive	If Puncture Skin Test Result Is Negative	Probability Shift if Skin Test Result Is Positive	Probability Shift if Skin Test Result Is Negative	If in Vitro Test Result Is Positive	If in Vitro Test Result Is Negative	Probability Shift if in Vitro Test Result Is Positive	Probability Shift if in Vitro Test Result Is Negative	If Phadiatop Test Result Is Positive	If Phadiatop Test Result Is Negative	Probability Shift if Phadiatop Test Result Is Positive	Probability Shift if Phadiatop Test Result Is Negative
← % →												
5	20.6	0.4	15.6	4.6	33.0	0.7	28.0	4.3	45.7	0.2	40.7	4.8
10	35.4	0.9	25.4	9.1	51.0	1.5	41.0	8.5	64.0	0.5	54.0	9.5
15	46.5	1.4	31.5	13.6	62.3	2.4	47.3	12.6	73.8	0.7	58.8	14.3
20	55.2	1.9	35.2	18.1	70.1	3.4	50.1	16.6	80.0	1.1	60.0	18.9
25	62.2	2.6	37.2	22.4	75.8	4.5	50.8	20.5	84.2	1.4	59.2	23.6
30	67.9	3.3	37.9	26.7	80.1	5.7	50.1	24.3	87.3	1.8	57.3	28.2
35	72.6	4.1	37.6	30.9	83.5	7.1	48.5	27.9	89.6	2.2	54.6	32.8
40	76.7	5.0	36.7	35.0	86.2	8.6	46.2	31.4	91.4	2.8	51.4	37.2
45	80.1	6.1	35.1	38.9	88.5	10.4	43.5	34.6	92.9	3.4	47.9	41.6
50	83.1	7.3	33.1	42.7	90.4	12.4	40.4	37.6	94.1	4.1	44.1	45.9
55	85.8	8.8	30.8	46.2	92.0	14.7	37.0	40.3	95.1	4.9	40.1	50.1
60	88.1	10.6	28.1	49.4	93.4	17.5	33.4	42.5	96.0	6.0	36.0	54.0
65	90.1	12.8	25.1	52.2	94.6	20.8	29.6	44.2	96.7	7.3	31.7	57.7
70	92.0	15.6	22.0	54.4	95.6	24.8	25.6	45.2	97.4	9.0	27.4	61.0
75	93.7	19.2	18.7	55.8	96.6	29.7	21.6	45.3	98.0	11.3	23.0	63.7
80	95.2	24.0	15.2	56.0	97.4	36.1	17.4	43.9	98.5	14.5	18.5	65.5
85	96.5	30.9	11.5	54.1	98.2	44.4	13.2	40.6	98.9	19.4	13.9	65.6
90	97.8	41.6	7.8	48.4	98.8	55.9	8.8	34.1	99.3	27.7	9.3	62.3
95	98.9	60.0	3.9	35.0	99.4	72.8	4.4	22.2	99.7	44.7	4.7	50.3

* Data on puncture skin tests and in vitro tests from Wood and colleagues (27), and data on Phadiatop test (Pharmacia-Upjohn, Uppsala, Sweden) from Eriksson (20), with permission. These studies were selected because the index test was not used in the gold standard test.