

Diagnostic Tests for Patients with Suspected Allergic Disease

Utility and Limitations

Paul P. VanArsdel Jr., MD; and Eric B. Larson, MD, MPH

Study Objective: To evaluate the clinical efficacy of diagnostic tests used for persons with suspected allergic disease.

Design: Information synthesis based on historical review of developments in the understanding of the pathophysiology of allergic diseases and on selected recent literature on efficacy of specific diagnostic tests.

Main Results: Skin testing is most effective when based on clues from the patient's history. The sensitivity and specificity of skin testing methods are compared; skin prick testing alone is often sufficient to identify or exclude immunoglobulin E (IgE)-mediated hypersensitivity, including food allergy. Except for penicillin and certain macromolecules, skin testing is not useful for evaluating drug allergy. Skin test titration may be useful for determining the starting dose for immunotherapy; otherwise it is rarely necessary. The patch skin test helps identify the cause of allergic contact dermatitis. Bronchial provocation testing is useful in special cases. Oral provocation testing may be used to identify allergy or other intolerance to suspected foods, food additives, and certain drugs. Provocation testing is time-consuming and requires special precautions. In-vitro methods for identifying allergen-specific IgE are especially useful when skin testing is unreliable, equivocal, or cannot be done. In-vitro tests should be used as adjuncts to the clinical interview and examination.

Conclusions: Tests that are effective for identifying allergenic substances usually can be determined from a careful patient interview. Clinicians should be aware of nonspecific test results and allergy tests of unproven effectiveness.

[Note that sections in this review are numbered so that they can be identified with cross-references as supporting evidence in the article, "Allergy Testing," published in the Position Paper section of this issue; see pages 317-320—The Editor.]

The diagnostic approach to allergic diseases involves identifying disease entities using clinical and pathophysiologic principles, and identifying certain substances that may cause or aggravate symptoms of such diseases. A basic concept of allergy and clinical immunology is that a subset of the population can become sensitive to certain substances that can produce anaphylactic, respiratory, cutaneous, or gastrointestinal symptoms in these persons, whereas other persons will not be affected (1). Most of the substances are common and are distributed widely in the form of pollens, mold spores, animal-derived dusts, and foods. The tendency to become sensitive is heritable, and most afflicted persons have first-order relatives with similar symptoms such as asthma, hay fever, or dermatitis. The term *atopic* is often applied to persons with these afflictions. The offending substances are true antigens immunologically, but are generally called *allergens*.

Allergic reactions are not limited to atopic patients. Nonatopic patients have a similar degree of risk for developing hypersensitivity to venomous insect stings, therapeutic or diagnostic agents, chemicals in the workplace, and cell-mediated reactions—allergic eczematous contact dermatitis in particular. Some adverse reactions, especially to food stuffs, are difficult to categorize and cause problems for physicians who evaluate the validity of diagnostic and therapeutic claims.

This review examines the utility and limitation of various diagnostic tests used to evaluate patients with suspected allergy and identifies problems unique to allergy testing. Allergy tests are used to define the subset of persons who, because of sensitivity to allergens, have symptoms caused by one or more allergen, and to identify the offending allergen or allergens. Conversely, negative results of allergy tests can differentiate persons whose symptoms are not caused by allergens from persons with allergen-mediated disease. The efficacy of allergy tests depends on the sensitivity, specificity, reproducibility, and safety of each test. Convenience and cost are also important considerations in evaluating comparative efficacy of different tests. This paper reviews the efficacy of in-vivo allergy tests (skin tests and provocation tests) and some of the in-vitro serologic and cellular tests used to evaluate persons with possible allergic disease.

1. Skin Tests

1.1 Immediate Type

Skin tests are used to detect immediate (antibody-mediated) or delayed (cell-mediated) sensitivity. The

Annals of Internal Medicine. 1989;110:304-312.

From University of Washington School of Medicine, Seattle, Washington. For current author addresses, see end of text.

commonest use of skin testing is to identify allergens that cause or provoke symptoms associated with immediate hypersensitivity. Such testing was the basis for the specialty of allergy as it developed 40 to 60 years ago. It was first used over a century ago to confirm the association between hay fever and pollen sensitivity. Early in this century, a small number of investigators began to report on the use of skin tests to identify allergens such as animal dander and various foods. Scientific credibility was added by Prausnitz and Küstner (2), who in 1921 reported the transfer of specific skin sensitivity by injecting serum from an allergic subject into the skin of a nonallergic subject. Within a few years this technique, the "P-K test," began to be used to supplement or replace direct skin testing of selected patients. In addition, it became a valuable research tool for identifying biochemical properties of the heat labile antibody called "skin-sensitizing antibody" or "reagin" for several decades before it was identified as Immunoglobulin E (IgE) about 20 years ago. The recognition of the IgE class of antibodies led to the development of in-vitro methods that supplanted P-K testing in clinical evaluations that required identification of specific serum antibodies.

At first, skin reactions were used to confirm fairly obvious relationships, such as ragweed hay fever and horse asthma. The number of allergens used in clinical practice steadily increased to include extracts derived from many pollens and molds, various dusts and insects, domestic and fur-bearing animal pelts, dozens of foodstuffs, and numerous bacterial vaccines and antigens. By the middle of this century, a patient having an allergy evaluation might be tested to as many as 300 substances purported to be allergenic.

Reliability

For a skin test to be reliable and clinically significant for a particular patient, the patient's symptoms or signs should have a reasonable possibility of being mediated by IgE antibodies to substances in the patient's environment; the allergens used must be safe and reliable, based on properly controlled clinical testing of subjects with appropriate allergic problems; and the patient must not use antihistaminic drugs for 2 to 5 days (depending on the drug used) before the test. Furthermore, the allergenic extracts must be prepared and administered at concentrations low enough to avoid nonspecific irritating reactions, but high enough to have appropriate sensitivity to identify IgE antibodies. They must be introduced into the skin by a reproducible technique that is associated with minimal trauma. The size of the skin reactions must be interpreted with the recognition that there is a threshold for clinical significance. Skin tests should be placed no less than 1 inch apart, and inflamed skin must be avoided. Finally, positive (histamine) and negative controls must accompany each set of tests.

As with other immunologic tests, standardization is difficult. In the simplest terms, each antigen should be standardized by measuring its reaction with an antibody standard. Yet antibodies may need to be standardized using antigen standards. Obviously standard-

ization was, in a sense, a bootstrap operation until highly purified antigens and monoclonal antibodies were developed.

Standardization of allergens based on their in-vivo and in-vitro reactivity with antibody has improved steadily in recent years through the efforts of the International Union of Immunological Societies of the World Health Organization and the Office of Biologics Research and Review of the U.S. Food and Drug Administration, among others. The Food and Drug Administration is responsible for determining safety and efficacy of allergenic extracts. Because of the efforts of the Food and Drug Administration and its expert advisory panels, almost all bacterial products (3) and more than 300 allergenic extracts (4) have been found to be ineffective, unsafe, or misbranded, and are no longer marketed. Many other extracts were determined to be safe and effective for testing, and remain in use. Few physicians are aware of the number of substances that were manufactured and used with no evidence of efficacy. Typical examples include cigar smoke, fiberglass, green tea, chalk, and newsprint.

Specificity and Safety

Some allergenic extracts of established efficacy are highly potent and must be diluted and administered properly for specificity and safety to be maintained. Testing of specificity by determining the prevalence of positive reactions in a normal (nonatopic) population is necessary to develop allergenic extracts for clinical use, and such tests have been reported from time to time for many decades. In 1961 Lindblad and Farr (5) reported their observations on intradermal tests of clinically nonallergic subjects using serial dilutions of several common allergenic extracts. They made two important observations. First, the high concentrations sometimes used in clinical practice produced positive reactions in 16% to 50% of subjects. Some were non-specific irritative reactions that could be attributed in part to the glycerol used as a stabilizing agent. The second observation was that intermediate concentrations produced reactions in 3% to 8% of subjects. These reactions were of particular interest because the sera of more than half of these subjects had skin-sensitizing antibody as determined by P-K testing, which provided good evidence that the subjects were immunologically but not clinically sensitive. This report should have alerted physicians who used intradermal testing that they were likely to assign significance to some clinically insignificant allergens if they used inappropriately high concentrations, or attached too much significance to small reactions (5 mm diameter or less). However, it is not known whether many physicians paid attention to the report.

Another important study was done by Hagy and Settignano (6) with college students. Using 15 common allergens, they scratch-tested 1836 (99.2%) of the freshman students at Brown University and Pembroke College in 1962 and 1963. As expected, most of the 25% who had a history of rhinitis or asthma or both had one or more positive reaction to skin tests. Seventeen percent of the students with a negative history

had one or more positive skin test reaction. A follow-up study of these students was conducted 3 years later to test the significance of these reactions. A substantial number of asymptomatic students with positive skin test reactions as freshmen had developed symptoms during the ensuing 3 years. Over 18% of students with positive pollen reactions developed symptoms of hay fever, compared with 1.7% of students with negative pollen reactions as freshmen (6). The conclusion was similar to that reached by Lindblad and Farr (5). For pollens, at least, skin tests represented true immunologic reactions. The presence of immunologic reactivity was a necessary, but not sufficient requirement for the development of clinical pollen allergy.

More recently, Barbee and coworkers (7) have used skin prick testing to five common allergens to determine the prevalence of sensitivity in a random sample of residents of Tucson, Arizona. They found that half their subjects had one or more positive reaction, and the prevalence in young adults was over 60%. They concluded that the high prevalence was related to subjects' perennial exposure to high concentrations of several airborne allergens in that region. Unfortunately, they have not yet correlated their skin test data with symptoms of atopic disease.

Placebo-controlled trials on stinging insect venoms have provided further information on skin-test validation. It was crucial to prove that subjects were anaphylactically sensitive by subjecting them to an actual sting. To determine whether venom desensitization (immunotherapy) was effective in protecting patients from future stings, it was necessary to establish the degree of sensitivity by intradermal skin-test titration using tenfold dilutions. Titration accomplished two things. First, it established a safe starting dose for immunotherapy for each patient. Second, it identified an upper limit to the testing concentration beyond which specificity was lost. Many allergic patients will react to venom concentrations of 0.01 or 0.1 $\mu\text{g/mL}$ or less. However, a significant number of nonallergic persons will react when the concentration is increased only a little more, to 3 $\mu\text{g/mL}$ (8).

Skin Testing Methods

The intradermal test is done with an allergenic extract diluted in an alkaline-buffered sterile saline solution that should contain 0.03% human serum albumin as a stabilizer. A half mL syringe with a #27 to #30 needle is used to produce a visible bleb (0.01 to 0.02 mL). Larger volumes increase the possibility of nonspecific reactions. A positive reaction consists of a wheal at least 5 mm in diameter, surrounded by significant erythema; lack of erythema should cast doubt on the significance of small reactions. To avoid antigen contamination, a different syringe and needle must be used for each antigen, and because of the possibility of microbial contamination, for each patient (9).

The intradermal test can be too sensitive when used routinely by inexperienced personnel. This sensitivity is a problem for two reasons. First, there is a small, but definite risk for inducing an adverse reaction in a patient who is highly sensitive. Second, nonallergic

persons may be falsely labeled as allergic based on nonspecific reactions. For this reason, epidermal methods of testing are used. Each method introduces antigen through a superficial penetration of the epidermis, such as a scratch, abrasion, puncture, or prick. Because of nonspecific reactions and lack of reproducibility, scratch and abrasion methods have been or should be abandoned (10).

The skin prick (or puncture) test, done in various ways, has proved specific and reasonably sensitive for most clinical applications. It was one of the first testing techniques to undergo rigorous objective evaluation. In 1950, Squire (11) reported on the quantitative analysis of skin prick tests, which were done by pricking the skin with a sharp needle passed through a drop of test solution placed on the skin; several concentrations of the solution were tested. The resulting wheal and flare reactions were traced, enlarged photographically onto graph paper, and measured by counting squares. Variations of this technique are used today. Using several concentrations of histamine, and comparing the reactions with reactions produced by intradermal injections of 0.001 mL each from a micrometer syringe, Squire determined that the skin prick test delivered 1/300 of this amount, or about 3×10^{-6} mL, with a coefficient of variation (using similar skin locations) of 16%. He then applied the method to the semiquantitative analysis of various horse antigens. Decades later, other studies (10) found similar differences in sensitivity between skin prick and intradermal testing. This difference is compensated for by using concentrated allergens (1/10 or 1/20 weight/volume) for skin prick testing. These amounts are usually stabilized in 50% glycerol and have two advantages over the more dilute solutions used for intradermal testing: They retain potency for a longer time, and they do not have to be kept in sterile vials. Furthermore, in contrast to intradermal testing (5), glycerol does not affect skin prick tests.

The commonest skin prick test is to introduce the tip of a stylet or a 27-gauge needle (or smaller) at a 15- to 20-deg angle through a drop of test allergen into the epidermis, then lift up, tenting the skin upward until the needle point pops loose. As with intradermal testing, a positive reaction consists of a wheal and erythema that reaches maximum size in 10 to 20 minutes. Because surface control sites usually show no reaction, in contrast to intradermal control sites, it is tempting to assign significance to any reaction. However, reactions less than 5 mm are of questionable significance (12).

If a hollow needle is used, it may carry some allergen from one test site to another, so it is advisable to use a new needle for each allergen. Lancets or solid needles offer a modest cost savings, because one needle can be used for testing several allergens on the same patient, if adequately wiped between tests. More details on such skin testing techniques were reviewed a few years ago by Nelson (13).

The test can also be done by introducing a needle perpendicularly through the test solution into the skin with a gentle stabbing motion, as was done by Squire

(11). When the two methods were compared in one investigation (14), there was little difference in sensitivity or in the coefficients of variation. Various types of needles and lancets have been developed and marketed based on claims of improved efficiency, lower cost, less discomfort, and improved reliability. However, there is no proof that any of these devices has an advantage over the properly controlled use of standard needles or lancets by experienced personnel.

Efficacy

Skin testing is clinically effective when limited to allergens that are suspected from a patient's history. The physician should review the patient's medical problem carefully, including the variations in symptoms according to environmental and seasonal changes. With appropriate selection of test allergens, a physician should rarely need to test a patient to more than 50 allergens (10) and, if food allergy is not suspected, fewer than 30 usually are sufficient.

Some experts in allergy practice use only skin prick testing; others use only intradermal tests. Each method alone is reasonably reliable if standardized extracts and appropriate controls are used and if the interpretation of the tests by skilled personnel is consistent with clinical findings. One common compromise is to use the simpler, relatively atraumatic skin prick test first. Intradermal testing can then be used for patients with a suggestive history and negative or equivocal results of skin prick tests to suspected allergens.

For identification or exclusion of IgE-mediated food allergy, we now have good evidence that skin prick testing is sufficient. In children, where the problem is much more prevalent than in adults, Bock and associates (15) found an excellent correlation between positive skin prick reactions to foods such as peanuts and eggs, and double-blind oral symptom-provoking challenges with corresponding foods. No patient with positive intradermal reactions only had positive oral challenges. If results are negative, a properly performed skin test, using a properly prepared and standardized allergen, excludes IgE-mediated hypersensitivity to that allergen with a high degree of certainty.

Skin testing for drug allergy is unreliable except for macromolecular agents (such as enzymes and xenogenic sera) and the penicillins. Penicillin allergy can be confirmed or excluded reliably by skin prick tests followed (if results are negative) by intradermal tests to the major determinant (penicilloyl polylysine) and a minor determinant mixture (16). Unfortunately, the minor determinant mixture has not yet been approved for marketing by the Food and Drug Administration, and testing with the major determinant alone is not sufficient to exclude allergy. One compromise, not altogether satisfactory, is to test with benzylpenicillin (one component of the minor determinant mixture). If results of this test also are negative, one should then administer a small test dose of the drug as a further precaution.

Skin Test Titration

The usual purpose of skin testing is to find out if a substance or substances are contributing to a person's

symptoms. One or, at most, two skin tests per allergen should be sufficient for this purpose. Occasionally, more quantitative information is needed than that provided by measuring the size of one skin reaction. In clinical practice such information is most often needed to help select a safe starting dose for immunotherapy. This dose is determined by serial testing with several dilutions of the treatment allergen or mixture of allergens to identify the lowest dilution that produces a positive skin reaction (the "end point"). Other than the special case of insect venom, titration is optional. If it is done, testing with one or two dilutions of the treatment extract is usually sufficient to determine the starting dose.

Another application of skin test titration is for allergen standardization. According to the degree of precision required, techniques vary from skin prick testing with tenfold dilutions (17) to intradermal tests with threefold dilutions. In the United States, some allergenic products are now standardized against reference extracts. Each reference standard is assigned a certain number of allergy units based on the dilution in a threefold titration that produces 50 mm (diameter) of skin erythema (18). The product can be standardized against the reference either by dilution skin testing or by an in-vitro assay that measures inhibition of the radioallergosorbent test (RAST) described in Section 3.1. Titration also offers one method of assessing the response to allergen immunotherapy and is a useful adjunct to the efficacy evaluation of immunotherapy conducted by clinical investigators (8), but is not essential for the care and follow-up evaluation of most patients who receive immunotherapy injections.

Unfortunately, skin test titration has the potential for overuse and subsequent inappropriate immunotherapy. It is used as part of a system of treatment that was originated by Hansel and developed by Rinkel in the 1950s. The basis of the system is serial testing using nine log₂ dilutions of all allergens to identify the "end points" that determine the starting immunotherapy doses. Testing is closely integrated with injection treatment in a highly structured program, and it is the treatment, not the testing, that is the most vulnerable to critical surveillance (19). As discussed earlier, the testing itself is a valid method for obtaining semiquantitative information about a person's sensitivity and for determining a safe beginning dose for immunotherapy. However, the system encourages the practitioner to overuse skin testing by doing up to nine different skin tests for each allergen, regardless of its appropriateness for immunotherapy. Furthermore, overzealous physicians are tempted to assign significance to reactions at the highest concentrations to justify the use of immunotherapy for patients who may not be allergic at all (5).

1.2 Delayed Type

Most skin testing for delayed, or T-cell mediated, sensitivity is done for one of two reasons. The first is to identify or exclude sensitivity to various microbial antigens. These antigens are used as an aid in recogniz-

ing exposure to infection such as tuberculosis, or to identify cell-mediated deficiency (anergy). Testing for clinical allergy to bacterial vaccines and antigens is of no proven value (3).

The second reason is to identify the cause or causes of allergic eczematous contact dermatitis. Most contact dermatitis is caused by irritants or by usually innocuous substances such as cosmetics, costume jewelry, dye-stuffs, or topical medications, which may become allergenic through continuous or repeated exposure. The diagnosis is often obvious, and the solution to the problem is simply to avoid contact with the offending material. Sometimes, testing is necessary to confirm the association or to identify one of several possible agents. The patch test is used for this purpose. A nonirritating concentration of each of the suspected materials is placed on an absorbent pad and is usually put on the skin of the back under occlusive nonallergenic tape. The sites are inspected at 48 and 72 hours, and earlier or later, whenever the patient notices any reaction. A positive reaction consists of pruritus, erythema, swelling, and vesiculation. The patch test is reliable and safe if properly done by well-qualified experts using appropriate materials (20). Essentially, it is a controlled, localized provocation of the original lesion. It is usually done only once on any given patient, using a relatively small number of substances. There are several variations in the use and interpretation of patch testing that are not particularly controversial and will not be discussed here.

2. Provocation Tests

Along with the patch test, most provocation tests are designed to reproduce the patient's main symptoms or signs by controlled exposure to the suspected substance. The substance may be delivered to the conjunctiva, nasal mucosa, bronchi, or gastrointestinal tract. Conjunctival challenge was used at one time as a supposedly safer alternative to skin testing with potent allergens such as horse serum, and some physicians used it for pollen testing. The method has been largely abandoned except for limited investigational use, because of the discomfort and limited amount of information that is gained compared with skin testing (10). For similar reasons, nasal provocation testing has limited usefulness except for clinical investigation into the mechanisms of nasal allergic reactions.

2.1 Bronchial Provocation Testing

Also called *inhalation challenge* testing, this procedure is a valuable research tool and in special cases has diagnostic usefulness as well. The testing is done by having the patient inhale low, gradually increasing concentrations of the suspected substance. The expiratory flow rate is measured periodically until a response occurs or the maximum nonirritating dose is reached. Depending on the problem, the patient may inhale a nebulized solution, or be exposed to dust or fumes in a special exposure chamber. In a typical test, the forced expiratory volume after one second (FEV₁) is mea-

sured every 5 minutes after each inhalation; a drop of the FEV₁ of 20% or greater is considered significant. In most tests, the maximum response will occur within 20 minutes. This response is termed the *early phase* reaction, and ventilatory function gradually returns to normal within 30 minutes. If the patient is monitored over a longer period, a second reduction in FEV₁ may appear 3 to 8 hours after challenge and may be associated with symptoms. This response is called the *late phase* reaction, which resolves slowly over many hours. Late phase reactions are now primarily of research interest—for example, to investigators who are studying the inflammatory aspects of asthma. In addition, the late phase has some diagnostic value, because certain persons may develop significant responses only in the late phase. Accordingly, any patient who is challenged should be monitored up to 8 hours after challenge, especially if no early phase response occurs.

The indications for inhalation challenge include identifying causative or provocative occupational or other allergens for which skin testing is not reliable, identifying new allergens for which skin or serum testing has not yet been validated, obtaining evidence to persuade a patient that an allergen is or is not asthmo-genic to that patient, and confirming the importance of unavoidable pollen or other inhalant allergens before committing a patient to immunotherapy (21, 22). These indications do not come up very often in most allergy practices. Most often, bronchial provocation testing is done by physicians who have both research and clinical interests in the procedure. Testing should be done only by experienced personnel in a facility with adequate emergency resuscitation equipment close by.

2.2 Oral Provocation Testing

Oral challenge sometimes helps in diagnosing allergy or intolerance to foods (including various food additives). It has been particularly valuable for validating other tests purported to be diagnostic of food allergy. As mentioned earlier, double-blind oral challenges done by Bock and colleagues (15) helped establish the reliability of skin prick testing for IgE-mediated food reactions. However, the verification of skin-testing methods is only one application of provocation testing. Rigorous evaluation of alleged or suspected food allergy by means of oral provocation is often needed to test various claims.

Because of the extraordinary number of medical and psychosocial variables involved, the combination of nutrition, diet, allergy, health, and illness has been particularly attractive to practitioners with inadequate scientific training and (at times) to unscrupulous entrepreneurs who have found that it is difficult for critics to obtain evidence to disprove their claims or to hold them accountable for their actions. This statement from a 1982 editorial in *Lancet* provides a useful perspective on the problem (23):

Nobody doubts that food allergy exists but the subject has acquired a dubious reputation. One reason is that the diagnosis has

been overworked to explain a great array of poorly understood clinical symptoms. Extravagant therapeutic claims have excited an antagonistic response, putting the subject under a cloud of controversy (which perhaps shelters a little quackery).

The first step toward establishing or excluding the diagnosis of food allergy is a careful review of the history. If symptoms are consistent with IgE-mediated reactions (anaphylaxis, acute urticaria and angioedema, eczema, asthma, rhinitis, nausea, vomiting, abdominal cramps, or diarrhea) and if the patient has noticed an association between certain symptoms and the ingestion of certain foods, then skin prick testing is likely to be helpful in identifying the offending food (15) and should be done first. Avoiding such food may resolve the problem.

If results of skin tests are negative or equivocal, or if the history is inconsistent with or atypical for IgE-mediated reactions, then oral challenge with the suspected foods may be necessary to identify what is now generally called *intolerance*, a phenomenon of uncertain pathogenesis that cannot be identified by any other testing method. Information from the history may be supplemented by repeatedly eliminating then reintroducing certain foods into the diet. If and when the patient becomes symptom-free or has minimal symptoms after elimination of all suspected foods, then it is customary to do an open challenge, adding one food at a time until symptoms are produced. Occasionally the result is inconclusive, especially if the patient has several fluctuating symptoms and is suggestive. A blinded challenge may then be necessary, and may also help in re-evaluating the patient who has been labeled as food-allergic on the basis of tests or procedures of doubtful validity.

Placebo-controlled, single-blinded or double-blinded challenges can be conducted by giving gelatin capsules containing various amounts of each suspected food. Freeze-dried foods, available from most outdoor recreation supply houses, are ideal for this purpose. Capsules hold 500 to 600 mg each, and results can be considered negative if no symptoms develop after ingestion of 10 to 12 capsules. If a severe reaction is suspected, then challenge should begin with 10 mg, followed by twofold increments every 60 to 90 minutes, until symptoms appear or the maximum dose of 8 to 10 grams is reached (15, 24, 25). The protocol must be flexible to account for the wide variations in the way symptoms and diet seem to be associated from patient to patient. The specificity of oral provocative challenge is fairly high, but its sensitivity is uncertain because of the difficulty in reproducing day-to-day dietary food exposure within the restrictions of blinded challenge procedures.

Oral challenge testing has a special role in the evaluation of patients with suspected sensitivity to nonsteroidal anti-inflammatory agents or to certain food additives, especially sulfiting agents. The usual adverse reaction is asthma, and skin testing is of no value. The oral route is used for bronchial provocation, allowing objective assessment. Challenge is done by starting

with a very small dose and gradually increasing the dose at 10- to 20-minute intervals while monitoring the FEV₁ as in inhalation provocation. The usual substances used for challenge are acetylsalicylic acid or potassium metabisulfite.

2.3 Drug Provocation

Occasionally a patient gives a history of allergy to a particular drug, treatment is essential, and there is no effective alternative drug. Except for macromolecular agents and the penicillins, skin testing to drugs is unreliable. If the risk of not treating exceeds the risk of a reaction, then careful challenge, orally if possible, may be justified, starting with 1% or less of the therapeutic dose, followed by tenfold increments at 15- to 60-minute intervals (depending on the route of administration). Drug fever and fixed eruptions are examples of low-risk reactions. Drug provocation should be conducted only by experts in a setting where emergency resuscitative equipment is at hand (26).

2.4 Other Provocation Tests

Some practitioners give small volumes of allergenic extracts, usually foods, one after another, intracutaneously, subcutaneously, or under the tongue, to provoke symptoms remote from the site of administration. This system supposedly detects sensitivity to the substance being tested by the appearance of "any systemic reaction" such as "headache, cough, nausea, mental or behavior change, etc" (27). As is the case with the Rinkel skin test titration system described earlier, cutaneous and sublingual symptom provocation is so closely tied to symptom "neutralization" procedures that a critique of the usefulness of the system for identifying significant allergens may be superfluous. The claim that cutaneous or sublingual provocation can identify specific allergens that are responsible for various somatic complaints has no foundation, either on the basis of generally accepted immunologic principles or results of properly controlled clinical trials (10, 19, 28-30). Furthermore, no measurable amount of allergen is absorbed after sublingual administration (19).

There are other unproven (31) and unacceptable food testing schemes that are used by some practitioners. One is the pulse test, which Arthur Coca, a well-known immunologist of half a century ago, advocated as an application of his theory of familial nonreaginic food allergy ("idioblapsis"). He claimed that an increase of the pulse rate after ingestion of a particular food indicated allergy (32). At the time, his observations could not be confirmed and the pulse test was all but forgotten, with little or no mention of it in the literature for several decades. Recently, the practice has been revived (33). Another method, kinesiology testing, involves muscle strength measurements after food ingestion or sublingual application of food extracts. There is no evidence in any peer-reviewed journal that either method is effective in identifying allergens.

3. In-Vitro Tests

3.1 Serologic Tests

Immunoglobulin E

The identification and characterization of IgE two decades ago was followed by research into the mechanisms of human allergic reactions as well as the development of two types of testing. One type is the measurement of total serum IgE by methods such as the paper radioimmunosorbent test (PRIST), in which IgE in the test serum reacts with anti-IgE bound to a solid phase (paper). The reaction is detected by adding ^{125}I labeled anti-IgE. The amount of radioactivity that remains on the paper is proportional to the total IgE (34). The most sensitive modification of PRIST can detect as little as 0.24 $\mu\text{g/L}$ (0.1 IU/mL). This degree of sensitivity is rarely necessary for clinical diagnosis, although it has several research applications. The mean total serum IgE in normal subjects is about 60 $\mu\text{g/L}$ with an upper 95% confidence limit of approximately 900 $\mu\text{g/L}$ (35). The mean serum IgE in patients with atopic disease falls within the normal range; so, in general, the total IgE is not a sensitive method for diagnosing the allergic state.

Extreme values of total IgE may be helpful. If the IgE is very low (less than 50 $\mu\text{g/L}$) it helps exclude atopic disease (but not IgE sensitivity to special allergens such as penicillin determinants or hymenoptera venoms). If it is over 900 $\mu\text{g/L}$, atopic disease is likely, but tests for sensitivity to specific allergens still need to be done if the information is to be useful. Very high levels (2000 to more than 60 000 $\mu\text{g/L}$) are found in asthma associated with severe atopic dermatitis, allergic bronchopulmonary aspergillosis, Buckley syndrome (staphylococcal infections with hyper-IgE), systemic parasitic infestations, and IgE myeloma (36).

The other development that followed the discovery of IgE was the measurement of allergen-specific antibody in the serum. The most widely used method is the radioallergosorbent test (RAST) that was developed in Uppsala, Sweden (37). This test is a solid-phase immunoassay in which small aliquots of the serum to be tested are added to each of the test allergens covalently bound to an insoluble carrier, such as a paper disk. The amount of specific IgE that reacts to antigen is then determined by adding ^{125}I labeled anti-IgE, washing the disk, and measuring the radioactivity bound to the disk. In a standard commercial assay, the amount of bound radioactivity in counts per minute is compared with a standard reference curve using five dilutions of pooled positive sera, and the test sera are scored as 0 (undetectable) 0/1 (very low), 1 (low), through 4 (very high). Alternatively, the result may be expressed in arbitrary units or as a percent of the reference serum. Such variables as the source and purity of anti-IgE, sample volume, extent of sample washing, incubation time, positive and negative controls (ideally, one for each allergen), interference by specific IgG and excessive IgE, and the arbitrary cut-off point in radioactive counts per minute between positive (clinically significant) and negative results must be taken into account. Results of the RAST cor-

relate well with skin and bronchial provocation tests, but the RAST is less sensitive. It probably has the same sensitivity as the P-K test. Up to 25% of patients with positive symptom-provocation tests will have negative results on RAST (35). A modified and more sensitive RAST has been developed that is supposed to detect antigen-specific IgE at radioactivity only 1.5 times that of negative controls. The method may not be controlled for variations in the nonspecific binding of some insolubilized allergens and positive results have been noted with some non-IgE containing fluids such as 5% human albumin (34, 35, 38).

The Usefulness of the Radioallergosorbent Test

No doubt exists about the value of RAST as a research technique. Furthermore, it has proved its worth for standardization of allergenic products by means of RAST inhibition, in which the product to be tested is used to inhibit the reaction between a standard allergen and specific IgE antibody. The extent to which RAST should be used in clinical diagnosis, however, has not yet been established. It is a serologic test, and as such, can be either less sensitive or less specific (depending on the methods and how the results are interpreted) than skin testing, in which not only the antibody, but the reactivity of the antibody-sensitized mast cell and the reactivity of the surrounding tissues to mast cell mediators are measured. The RAST is more expensive than skin testing, and the results are not immediately available. However, it is the test to use when skin testing is unreliable, as in patients with severe dermatographism or generalized dermatitis, and for patients who are unable to stop using antihistamines (35). Advocates of RAST claim that RAST is necessary for testing patients who are afraid of needles or uncooperative, and that RAST entails no risk for an adverse systemic reaction. These claims are dubious in that extreme fear of needles is rare, except for small children, and properly performed skin tests have an excellent safety record (39, 40).

Newer assay methods for allergen-specific serum IgE antibodies promise to be less expensive and to produce results in a shorter time than RAST. They use other solid phases such as agarose particles, cellulose threads, and microtiter plates; use radioactive, enzyme-linked (colorimetric) or fluorescent indicators; and are marketed with various catchy acronyms (34). By simplifying the measurement of specific serum IgE antibodies, such methods should be helpful adjuncts to clinical diagnosis. In general, their reliability is similar to the standard RAST. A comparison of the various in-vivo and in-vitro methods for measuring immediate hypersensitivity and their merits and drawbacks are given in Table 1.

The main problem in the use of promotion of in-vitro testing is that serologic methods may become so inexpensive and convenient that the laboratory, often far removed from the patient, will replace clinical judgment in assigning significance to various allergens and, worse, in making treatment decisions (41). The diagnosis of allergen-induced disease is made by the clinician taking an appropriate history; testing for IgE

Table 1. Relative Efficacy of Various Diagnostic Tests for Immediate Hypersensitivity, Based on Six Desirable Characteristics*

Variable	Skin Tests			Provocation Tests		Specific IgE (RAST)†
	Prick	Intradermal	Titration	Bronchial	Oral	
Sensitivity	4	5	Dose-related	Dose-related	Dose-related	3
Specificity	4	2	Dose-related	5	4	5
Safety	5	4	3	2	3	5
Reproducibility	4	4	4	4	3	5
Convenience‡	5	4	2	1	2	3
Economy (low cost per allergen)	5	4	2	1	2	3

* Characteristics are rated on a 5-point scale where: 5 = highest in class; 4 = acceptable for general use; 3 = intermediate, but important limitations; 2 = use is justified only in special cases; 1 = lowest, unacceptable except for occasional use in specialized centers. The ratings are based on the assumption that the procedures are done by well-qualified personnel using properly standardized reagents and adequate quality control.

† RAST = radioallergosorbent test.

‡ Convenience is a composite of efficiency and ease of testing; lack of patient discomfort; and availability of results.

antibodies is useful primarily for confirming or excluding the role of the suspected allergens.

Immunoglobulin G

For many years the detection of precipitating antibodies to allergens such as avian or fungal proteins has been an acceptable aid for the diagnosis of Farmer lung and other types of hypersensitivity pneumonitis. Measurements of specific nonprecipitating IgG antibodies to many other allergens are now done by commercial laboratories. These laboratories use standard radioimmunoassay techniques such as radioimmunoprecipitation, staphylococcal protein A binding, and IgG RAST. To date the only potential usefulness of specific IgG measurements is for monitoring the effectiveness of insect venom immunoassay. The clinical usefulness of measuring antigen-specific IgG subclasses is still under investigation (34). In particular, recent claims have appeared regarding the measurement of IgG₄ antibodies in allergy diagnosis, especially food allergy. However, there is no proof that the detection of such antibodies in the serum has any diagnostic value (42).

Immune Complexes

The measurement of circulating immune complexes may be of some value in diagnosing and following the progress of autoimmune and infectious disorders (vasculitis, systemic lupus erythematosus, hepatitis B, and bacterial endocarditis, for example), but it is of no proven value for the diagnosis of allergy to foods or any other allergens. In particular, there is no convincing evidence that either IgG or IgE food-immune complexes cause human disease (43).

Complement

The measurement of serum complement and its components is now done by most large clinical laboratories. Complement and its early-reactive components are depressed in some patients with acute glomerulonephritis, systemic lupus erythematosus, and vasculitis. Serum complement measurement may be helpful in the diagnosis and prognosis of these disorders and in following their activity. An occasional immunodeficiency disease may be traced to the lack of a functioning complement component. Hereditary angioedema,

associated with a functional deficiency in the inhibitor of activated C1, is associated with a low C4 in the serum. A similar condition may develop secondary to a lymphoproliferative or connective tissue disorder. The levels of C1 and C1q are also low in these conditions. Elevated complement levels have little diagnostic value; they are found in association with many acute and chronic illnesses, and complement values may fluctuate unpredictably. Serum complement changes in response to challenge with allergens such as foods have no diagnostic significance (36).

3.2 Cellular Tests

Histamine Release

The measurement of antigen-induced histamine release from peripheral blood basophils has been the subject of study for more than 40 years, and has been a valuable research tool, with sensitivity similar to the P-K test and RAST. However, it is expensive, time-consuming, and restricted to investigational use (36).

Lymphocyte Transformation

In the laboratory, cell-mediated reactions are studied by measuring lymphocyte responses in tissue culture. Sensitized lymphocytes undergo blastogenesis in culture and generate lymphokines on incubation with microbial or histocompatibility antigens. The in-vitro lymphocyte response is clinically useful in transplant compatibility screening and as part of the evaluation of immunodeficiency (36). It is not likely to provide any information on microbial immunity beyond that provided by the delayed skin test. Otherwise, much of the clinically related literature in recent years has consisted of reports on the use of in-vitro antigen-induced lymphocyte responses to diagnose drug allergy. A large proportion of allergic drug reactions are suspected of having a cell-mediated mechanism. Some reliable research-oriented laboratories have provided convincing data on blastogenesis and lymphokine production induced in vitro by a small number of drugs (26). However, few reports from one laboratory have been reproduced in other laboratories, and thus the application of lymphocyte response tests to the diagnosis of drug allergy should be considered experimental.

Lymphocyte Subsets

The measurement of circulating B cells and T-cell subsets is now a standard procedure in many clinical immunology laboratories, and has proved useful in evaluating immunodeficiency and lymphoproliferative disorders. In the research laboratory, interesting findings are appearing regarding the T-cell regulation of IgE synthesis in allergic states and during allergen immunotherapy (36, 42). However, the routine measurement of lymphocyte subsets as part of the evaluation of patients with suspected sensitivity to environmental allergens and chemicals is of no proven value (44).

Cytotoxic Testing

In the 1930s, the leukopenic index, a laboratory counterpart to the pulse test, came into use for diagnosing food allergy. It had a short life, but was superseded by claims, first appearing in the 1950s, that food allergy could be diagnosed by mixing aliquots of buffy coat cells from patients' blood with various food extracts and inspecting the cells for cytolysis, supposedly an indicator of food allergy. Leukocytotoxic testing has no rational scientific basis. On the contrary, the known immediate hypersensitivity reactions involving leukocytes involve the noncytotoxic release of mediators. Furthermore, in properly controlled trials of cytotoxic testing, several qualified clinical investigators have determined that the procedure is not reliable for identifying symptom-provoking allergens (28, 36, 45).

Acknowledgements: The authors thank the allergists, immunologists, internists, and otolaryngologists who reviewed drafts of this manuscript.

Requests for Reprints: Paul P. VanArsdel, Jr., MD, Department of Medicine, RM-13, University Hospital, Seattle, WA 98195.

Current Author Addresses: Dr. VanArsdel: Department of Medicine, RM-13, University Hospital, Seattle, WA 98195.

Dr. Larson: Department of Medicine, RG-20, University of Washington, Seattle, WA 98195.

References

1. VanArsdel PP Jr. Clinical causes of asthma. In: Weiss EB, Segal MS, Stein M, eds. *Bronchial Asthma: Mechanisms and Therapeutics*. 2nd ed. Boston: Little Brown & Co.; 1985:303-6.
2. Prausnitz V, Küstner H. Studien über die Ueberempfindlichkeit. *Centralbl Bakteriol Parasit Infektionskrankheiten*. 1921;86:160-9.
3. Farr RS, Chase RM Jr, Dees SC, Feldman HA, Huntley CC, Isaacson PE, et al. Bacterial vaccines and bacterial antigens with no U.S. standard of potency. *Fed Reg*. 1977;42:58266-8318.
4. Schaeffer M, Sisk LC. Allergenic extracts: a review of their safety and efficacy. *Ann Allergy*. 1984;52:2-14.
5. Lindblad JH, Farr RS. The incidence of positive intradermal reactions and the demonstration of skin sensitizing antibody to extracts of ragweed and dust in humans without history of rhinitis or asthma. *J Allergy*. 1961;32:392-401.
6. Hagy GW, Settiple GA. Prognosis of positive allergy skin tests in an asymptomatic population: a three-year follow-up of college students. *J Allergy*. 1971;48:200-11.
7. Barbee RA, Kaltenborn W, Lebowitz MD, Burrows B. Longitudinal changes in allergen skin reactivity in a community population sample. *J Allergy Clin Immunol*. 1987;79:16-24.
8. Valentine MD, Golden DBK. Diagnosis of insect allergy. In: Levine ML, Lockey RF, eds. *Monograph on Insect Allergy*. 2nd ed. Milwaukee, Wisconsin: American Academy of Allergy and Immunology; 1986:40.
9. Lutz CT, Bell CE, Wedner HJ, et al. Allergy testing of multiple patients should no longer be performed with a common syringe. *N Eng J Med*. 1984;310:1335-7.
10. Council on Scientific Affairs. In vivo diagnostic testing and immunotherapy for allergy: report 1 of the Allergy Panel. *JAMA*. 1987;258:1363-7, 1505-8.
11. Squire JR. The relationship between horse dandruff and horse serum antigens in asthma. *Clin Sci*. 1950;9:127-50.
12. Lessof MH, Buisseret PD, Merrett J, Merrett TG, Wraith DG. Assessing the value of skin prick tests. *Clin Allergy*. 1980;10:115-20.
13. Nelson HS. Diagnostic procedures in allergy. I. Allergy skin testing. *Ann Allergy*. 1983;51:411-8.
14. Perrin LF, Dechamps C, Deviller P, Joly P. Reproducibility of skin tests: a comparative study of the Pepys prick test and the Morrow-Brown needle and their correlation with the serum IgE level. *Clin Allergy*. 1984;14:581-8.
15. Bock SA, Buckley J, Holst A, May CD. Proper use of skin tests with food extracts in diagnosis of hypersensitivity to food in children. *Clin Allergy*. 1977;7:375-83.
16. VanArsdel PP Jr, Martonick GJ, Johnson LE, Sprenger JD, Altman LC, Henderson WR Jr. The value of skin testing for penicillin allergy diagnosis. *West J Med*. 1986;144:311-4.
17. Malling HJ. Quantitative skin prick testing: dose-response of histamine- and allergen-induced wheal reactions. *Allergy*. 1987;42:196-204.
18. Turkeltaub PC. In vivo methods of standardization. *Clin Rev Allergy*. 1986;4:371-87.
19. Hirsch SR. Proved and unproved procedures in the practice of allergy. In: Altman LC, ed. *Clinical Allergy and Immunology*. Boston: G.K. Hall Medical Publishers; 1985:457-66.
20. Fisher AA. *Contact Dermatitis*. 3rd ed. Philadelphia: Lea & Febiger; 1986:9-29.
21. Nelson HS. Diagnostic procedures in allergy. II. Bronchial provocation testing. *Ann Allergy*. 1983;51:488-92.
22. Cockcroft DW. Bronchial inhalation tests. II. Measurement of allergic (and occupational) bronchial responsiveness. *Ann Allergy*. 1987;59:89-98.
23. Food allergy and intolerance [Editorial]. *Lancet*. 1982;2:1344-5.
24. Metcalfe DD. Diagnostic procedures for immunologically-mediated food sensitivity. *Nutr Rev*. 1984;42:92-7.
25. Lessof MH, Wraith DG, Merrett TG, Merrett J, Buisseret PD. Food allergy and intolerance in 100 patients—local and systemic effects. *Q J Med*. 1980;195:259-71.
26. VanArsdel PP Jr. Drug hypersensitivity. In: Bierman CW, Pearlman DS, eds. *Allergic Diseases from Infancy to Adulthood*. 2nd ed. Philadelphia: WB Saunders; 1988:705-6.
27. Forman R. A critique of evaluation studies of sublingual and intracutaneous provocative tests for food allergy. *Med Hypotheses*. 1981;7:1019-27.
28. Van Metre TE Jr. Critique of controversial and unproven procedures for diagnosis and therapy of allergic disorders. *Pediatr Clin North Am*. 1983;30:807-17.
29. Jewett DL, Greenberg MR. Placebo responses in intradermal provocation testing with food extracts. *J Allergy Clin Immunol*. 1985;75:205.
30. Terr AI. Clinical ecology. *J Allergy Clin Immunol*. 1987;79:423-6.
31. Dong FM. *All About Food Allergy*. Philadelphia: Stickley; 1984:34-5.
32. Coca AF. *Familial Non-Reaginic Food Allergy*. Springfield: Charles C Thomas; 1942.
33. Radcliffe MJ. Diagnostic use of dietary regimes. In: Brostoff J, Challacombe SJ, eds. *Food Allergy and Intolerance*. London: Baillière Tindall; 1987.
34. Hamilton RG, Adkinson NF Jr. Clinical laboratory methods for the assessment and management of human allergic diseases. *Clin Lab Med*. 1986;6:117-38.
35. Yuninger JW. Clinical significance of IgE. In: Middleton E, Reed CE, Ellis EF, Adkinson NF Jr, Yuninger JW, eds. *Allergy, Principles and Practice*. 3rd ed. St Louis: CV Mosby; 1988:850-7.
36. Council on Scientific Affairs. In vitro testing for allergy: report II of the Allergy Panel. *JAMA*. 1987;258:1639-43.
37. Wide L, Bennich H, Johansson SG. Diagnosis of allergy by an in vitro test for allergen antibodies. *Lancet*. 1967;2:1105-7.
38. Santrach PJ, Parker JL, Jones RT, Yuninger JW. Diagnostic and therapeutic applications of a modified radioallergosorbent test and comparison with the conventional radioallergosorbent test. *J Allergy Clin Immunol*. 1981;67:97-104.
39. VanArsdel PP Jr, Sherman WB. The risk of inducing constitutional reaction in allergic patients. *J Allergy*. 1957;28:251-61.
40. Lockey RF, Benedict LM, Turkeltaub PC, Bukantz SC. Fatalities from immunotherapy (IT) and skin testing (ST). *J Allergy Clin Immunol*. 1987;79:660-77.
41. Anderson JA, Chai H, Claman HN, et al. Position statement: American Academy of Allergy and Immunology: the remote practice of allergy. *J Allergy Clin Immunol*. 1986;77:651-2.
42. Homburger HA. Diagnosis of allergy: in vitro testing. *CRC Crit Rev Lab Sci*. 1986;23:279-314.
43. Sheffer AL, Lieberman PL, Aaronson DW, et al. Measurement of circulating IgG and IgE food-immune complexes. *J Allergy Clin Immunol*. 1988;81:758-60.
44. Marshall E. Immune system theories on trial. *Science*. 1986;234:1490-2.
45. Grieco MH. Controversial practices in allergy. *JAMA*. 1982;247:3105-11.

