Evaluation of the Patient with Suspected Peanut Allergy: A Focused Evidence-based Guideline

Matthew Greenhawt,[1] Marcus Shaker,[2] Julie Wang,[3] John J Oppenheimer,[4] Scott Sicherer,[3] Corinne Keet,[5] Keri Swaggart[6], Matthew Rank,[7] Jay M Portnoy,[8] Jonathan Bernstein[9], Chitra Dinakar[10], David Golden[11], Carolyn Horner[12], David Lang[13], Eddy S. Lang[14],David Khan[15], Jay Lieberman[16], David Stukus[17], and Dana Wallace[18].

[1] Section of Allergy and Immunology, Department of Pediatrics, Children's Hospital Colorado, University of Colorado School of Medicine, Aurora, CO

[2] Dartmouth Geisel School of Medicine, Dartmouth-Hitchcock Medical Center, Section of Allergy and Immunology, Lebanon, NH

[3] Division of Pediatric Allergy and Immunology, Department of Pediatrics, Icahn School of Medicine at Mount Sinai and the Jaffe Food Allergy Institute, New York, NY

[4] Department of Internal Medicine, New Jersey Medical School, Morristown, NJ

[5]Division of Pediatric Allergy and Immunology, Johns Hopkins School of Medicine, Baltimore, MD

[6]Library Services, Children's Mercy Hospital, Kansas City, MO

[7]Division of Allergy, Asthma, and Clinical Immunology, Mayo Clinic, Scottsdale, AZ and Division of Pulmonology Phoenix Children's Hospital Phoenix, AZ

[8]Division of Allergy, Asthma & Immunology, Children's Mercy Hospital, Kansas City, MO

[9]Division of Immunology, Department of Internal Medicine, University of Cincinnati, Cincinnati, Ohio

[10]Division of Allergy and Asthma, Stanford University School of Medicine, Palo Alto, CA

[11]Department of Allergy-Clinical Immunology John Hopkins, Baltimore, MD

[12]Division of Allergy, Immunology, and Pulmonary Medicine, Washington University School of Medicine, St. Louis MO

[13]Department of Allergy and Clinical Immunology, Cleveland Clinic, Cleveland, OH

[14] Department of Emergency Medicine, Cumming School of Medicine, University of Calgary, Calgary, Alberta, Canada

[15]Division of Allergy & Immunology, Department of Medicine, University of Texas Southwestern Medical Center, Dallas, TX

[16]Division of Allergy and Immunology, The University of Tennessee, Memphis, TN

[17]Division of Allergy and Immunology, The Ohio State University College of Medicine, Columbus, OH

[18]Nova Southeastern University College of Allopathic Medicine, Fort Lauderdale, FL

Contact person

Peris Flagg

Joint Taskforce on Practice Parameters

American Academy of Allergy, Asthma & Immunology
555 E. Wells Street, Suite 1100
Milwaukee, WI 53202-3823
(414) 918-3124 (p)
pflagg@aaaai.org
www.aaaai.org

Peanut Allergy Diagnosis- a 2020 Practice Parameter Update, Systematic Review, and GRADE Analysis

Executive Summary

IgE mediated peanut allergy has an estimated prevalence of between 0.2-4.5%, depending on geographic area of the world and the methodology used for assessment.[1] While the prevalence in the US appears to have tripled in a recent 10-year period, in the UK the prevalence seems to have plateaued over a similar period, denoting regional heterogeneity in such trends.[1] Peanut allergy is associated with substantial economic and psychologic burden on families in that many suffer from poor empowerment, poor quality of life, and high anxiety related to the potential consequences of their child having an allergic reaction.[2,3] Peanut allergy is often a severe and usually a lifelong allergy that is a leading cause of food-related anaphylaxis.[1] There are emerging treatments approaching potential FDA approval for peanut allergy.[4] However, presently peanut allergy is managed through peanut avoidance, and by carrying emergency medication such as auto-injectable epinephrine to treat symptoms that may arise from unintended ingestion.[5]

Given this burden of disease and the consequences of diagnosis, it is important that peanut allergy be accurately diagnosed so that an appropriate treatment plan can be developed. However, a positive peanut test result is not always associated with clinical reactivity. This practice parameter addresses the diagnosis of IgE mediated peanut allergy both in children and adults as pertaining to 3 fundamental questions (see text box 1). This parameter exclusively discusses IgE mediated peanut allergy and all references herein pertain to IgE mediated food allergy to peanut only, and not to peanut as a potential trigger in eosinophilic esophagitis or non-IgE mediated food allergy such as food protein induced enterocolitis syndrome.

Diagnostic testing for peanut allergy is used to help make a diagnosis where there is suspicion of a peanut allergy based on the clinical history.[6] Failure to make a correct diagnosis can result in either unnecessary avoidance in a non-allergic person, or erroneous guidance that the patient can safely ingest peanut *ad libitum* when there is in fact an allergy—situations that are both problematic. A correct diagnosis facilitates peanut avoidance and counseling when the patient is at risk of potential life-threatening complications of peanut allergy, and therefore is advised to carry epinephrine for use in case

of symptomatic accidental ingestion. Alternatively, exclusion of peanut allergy allows peanut to be incorporated into the diet without concern, eliminating the burden of precautions and fear.[1] Testing is also used to monitor changes in baseline peanut sensitization since diagnosis, which may decrease (or increase) over time and may be associated with an increased likelihood that an allergic individual may be outgrowing their peanut allergy.[7,8] Although previous research in patients with established peanut allergy reported clinical diagnostic cut-off points for >95% chance of reaction and for <50% chance of reaction to oral food challenge, these are not necessarily predictive of clinical outcomes in all settings and patients, as they are highly dependent on the baseline prevalence of peanut allergy in the particular population.[1, 9-11]

The panel developed the key (PICO) questions to be addressed, and after systematic review of the literature (>1300 references searched), meta-analysis of the evidence, and GRADE analysis of the results, made recommendations - all of which were conditional in strength, with very low certainty of evidence. Thresholds for testing were at 3mm for SPT, and 0.35 KU_A/L for both whole peanut sIgE and component-specific peanut sIgE, based on the most widely reported levels evaluated in the literature. Extensive sensitivity analysis was performed to confirm the results.

The panel suggested that diagnostic testing for peanut allergy be used in patients with a high pre-test probability of peanut allergy, or prior to an oral food challenge for patients with moderate pre-test probability of peanut allergy, as a preference-sensitive choice, but not in patients with a low or very low pre-test probability of peanut allergy. If a single diagnostic test is to be used, testing for the Ara h 2 component would provide the most diagnostic accuracy as determined by the more optimal positive/negative likelihood ratio, provided this is available in the future as a stand-alone test and not ordered as a panel with other peanut components. The literature search did not provide patient-level data to determine the value of testing for peanut components in addition to skin prick test or sIgE to whole peanut to increase diagnostic accuracy, including isolated Ara h 2 in that context. The clinician should not use the results of a SPT, sIgE to whole peanut extract, or sIgE to peanut components to determine an allergy phenotype or to predict the severity of a future reaction (e.g., is the patient “anaphylactic” to peanut). Additional analysis of the health and economic benefits of the potential testing options showed that at multiple presumed prevalence of peanut allergy in the population, compared to use of peanut-specific Ara h 2 testing, the use of either whole peanut extract SPT or sIgE was associated with higher costs and lower health benefits (e.g. dominated analysis), making Ara h 2 the most cost-effective option

in this analysis until the specificity of Ara h 2 testing fell below 0.46. There remain important knowledge gaps and needs for well-designed studies to address these questions, as well as the need for patient-level data to be made available when reporting test sensitivity/specificity to enhance the ability to perform future meta-analysis that can explore different cut-off levels.

Question 1: Should diagnostic testing for peanut allergy be performed in adults and children with a history of suspected peanut allergy who are requesting evaluation for peanut allergy?

Recommendation 1a: We suggest in favor of diagnostic (skin prick or serum sIgE) testing for peanut allergy in patients with a 1) physician-judged high pre-test probability of peanut allergy, or 2) prior to an oral food challenge for patients with moderate pre-test probability of peanut allergy, for both of whom shared decision-making has been employed to arrive at the final decision**. Conditional recommendation; Certainty of evidence: very low**

Recommendation 1b: We suggest against diagnostic testing in patients where there is low or very low pre-test probability of peanut allergy. **Conditional recommendation; Certainty of evidence: very low**

Discussion: This question was not searched in a systematic manner as the content experts were unaware of any single research study that addressed this question. The workgroup did a Pubmed literature search that did not come up with any articles that address this question, which by default limits the certainty of evidence. The workgroup and JTFPP felt that it would be a waste of valuable resources to conduct a librarian-conducted formal literature search. However, expert evidence was collected both from the content experts and the JTFPP. Expert evidence differs from expert opinion, in that the former does not include a judgment on the evidence and offers a systematic and transparent appraisal of the evidence.[12] In their collective personal clinical experience, the guideline working group related that when evaluating their collective patient experiences, that diagnostic testing could be of value to confirm peanut allergy in high-risk individuals for which an oral challenge might not be advisable or agreed to by patients, but also acknowledged that in a patient presenting with a classical history the diagnosis could be made on the basis of history alone without further testing in some circumstances. The panel related that they suggested an oral food challenge when there was a moderate probability of peanut allergy but that a large proportion of their patients may prefer a diagnostic test prior to the oral food challenge. Similarly,

the collective personal experience of the panel was that diagnostic testing in patients with a low probability of peanut allergy (e.g., sibling has peanut allergy and patient has never ingested peanut) identified patients who were sensitized but not truly allergic. Unfortunately, many of these patients refused an oral food challenge and likely avoided peanut unnecessarily.

These recommendations are in alignment with previous expert guidelines and practice parameters[13-15] on food allergy diagnosis and management which provide similar consensus regarding the indications for testing for the presence of food sensitization, including peanut, in evaluating a possible diagnosis of food allergy. While screening of infants to foods prior to food introduction is discouraged, testing to peanut in infants at high-risk for peanut allergy (under the very prescribed context of those infants with either severe eczema and/or egg allergy) is now recommended prior to initial peanut introduction per the 2017 NIAID addendum guidelines.[16]

Question 2a: In the patient presenting for evaluation of suspected peanut allergy, which of the three tests—SPT, sIgE to whole peanut, or Ara h2 would provide the highest diagnostic accuracy as determined by the more optimal positive/negative likelihood ratio?

Question 2b: In a patient presenting for evaluation of suspected peanut allergy, does testing for peanut components in addition to either SPT or sIgE to whole peanut increase the diagnostic accuracy?

Recommendation 2a: We suggest in favor of Ara h2 diagnostic testing (over SPT or sIgE to whole peanut) in a patient presenting for evaluation of suspected peanut allergy for which a single diagnostic test is to be used, as Ara h2 would provide the best diagnostic accuracy as determined by virtue of more optimal positive/negative likelihood ratios. **Conditional recommendation. Certainty of evidence moderate.**

Recommendation 2b: We suggest against component testing in addition to either to skin prick test or sIgE to whole peanut to increase diagnostic accuracy. **Conditional recommendation**. **Certainty of evidence: very low**

Discussion: For GRADE analysis, Ara h 2 was compared to skin prick test and sIgE to whole peanut for the diagnosis of peanut allergy. (See Summary of GRADE Question below and review the Evidence to Recommendation Table for details) The literature search did not provide patient-level data to determine the value of testing for peanut components in addition to or reflexively with skin prick test or sIgE to whole peanut to increase diagnostic accuracy. In addition, expert evidence was not available to assist in answering this question. Thus, the use and value of components, including reflexive use of Ara h 2, remains a knowledge gap. There is an unclear utility for measuring sIgE to any other commercially available peanut components given the limited available data on performance of components beyond Ara h 2. Further research is needed to clarify the value of tandem testing, particular in regards to Ara h 2, Ara h 6, and Ara h 8.

Question 3: In the patient presenting for evaluation of suspected peanut allergy, can the results of a diagnostic test be used to predict the severity of a future allergic reaction?

Recommendation 3: We suggest against the clinician using the results of a SPT, sIgE to whole peanut extract, or sIgE to peanut components to determine the severity of an allergy phenotype or to predict the severity of a future reaction**. Conditional recommendation. Certainty of evidence: very low.**

Discussion: There was inadequate patient-level data to formulate a GRADE recommendation on the use of a diagnostic test for predicting the severity of a future allergy reaction to peanut.

Executive Summary References:

1. National Academies of Sciences E, Medicine. Finding a Path to Safety in Food Allergy: Assessment of the Global Burden, Causes, Prevention, Management, and Public Policy. Washington, DC: The National Academies Press; 2017.
2. Roy KM, Roberts MC. Peanut allergy in children: relationships to health-related quality of life, anxiety, and parental stress. Clin Pediatr (Phila) 2011;50:1045-51.
3. Antolin-Amerigo D, Manso L, Caminati M, de la Hoz Caballer B, Cerecedo I, Muriel A, et al. Quality of life in patients with food allergy. Clin Mol Allergy 2016;14:4.
4. Wood RA. Oral Immunotherapy for Food Allergy. J Investig Allergol Clin Immunol 2017;27:151-9.

5. Sicherer SH, Sampson HA. Food allergy: A review and update on epidemiology, pathogenesis, diagnosis, prevention, and management. J Allergy Clin Immunol 2018;141:41-58.
6. Bernstein IL, Li JT, Bernstein DI, Hamilton R, Spector SL, Tan R, et al. Allergy Diagnostic Testing: An Updated Practice Parameter. Annals of Allergy, Asthma & Immunology 2008;100:S1-S148.
7. Stiefel G, Anagnostou K, Boyle RJ, Brathwaite N, Ewan P, Fox AT, et al. BSACI guideline for the diagnosis and management of peanut and tree nut allergy. Clinical & Experimental Allergy 2017;47:719-39.
8. Sampson HA, Aceves S, Bock SA, James J, Jones S, Lang D, et al. Food allergy: A practice parameter update—2014. Journal of Allergy and Clinical Immunology 2014;134:1016-25.e43.
9. Peters RL, Allen KJ, Dharmage SC, Tang ML, Koplin JJ, Ponsonby AL, et al. Skin prick test responses and allergen-specific IgE levels as predictors of peanut, egg, and sesame allergy in infants. J Allergy Clin Immunol 2013;132:874-80.
10. Beyer K, Grabenhenrich L, Hartl M, Beder A, Kalb B, Ziegert M, et al. Predictive values of component-specific IgE for the outcome of peanut and hazelnut food challenges in children. Allergy 2015;70:90-8.
11. Sampson HA, Ho DG. Relationship between food-specific IgE concentrations and the risk of positive food challenges in children and adolescents. Journal of allergy and clinical immunology 1997;100:444-51.
12. Schünemann HJ, Zhang Y, Oxman AD; Expert Evidence in Guidelines Group. Distinguishing opinion from evidence in guidelines. BMJ. 2019 Jul 19;366:l4606. doi: 10.1136/bmj.l4606.
13. Boyce JA, Assa'ad A, Burks AW, Jones SM, Sampson HA, Wood RA, et al. Guidelines for the Diagnosis and Management of Food Allergy in the United States: Summary of the NIAID-Sponsored Expert Panel Report. J Allergy Clin Immunol 2010;126:1105-18.
14. Sampson HA, Aceves S, Bock SA, James J, Jones S, Lang D, et al. Food allergy: a practice parameter update-2014. J Allergy Clin Immunol 2014;134:1016-25 e43.
15. Muraro A, Werfel T, Hoffmann-Sommergruber K, Roberts G, Beyer K, Bindslev-Jensen C, et al. EAACI food allergy and anaphylaxis guidelines: diagnosis and management of food allergy. Allergy 2014;69:1008-25.
16. Togias A, Cooper SF, Acebal ML, Assa'ad A, Baker JR, Jr., Beck LA, et al. Addendum guidelines for the prevention of peanut allergy in the United States: Report of the National Institute of Allergy and Infectious Diseases-sponsored expert panel. J Allergy Clin Immunol 2017;139:29-44.

The following primer section helps provide background context on peanut allergy and the principles of how to apply diagnostic testing for peanut allergy. The next sections detail specific applications of diagnostic testing, determined through evidence-synthesis, meta-analysis, and systematic review to provide a clinical practice guideline for the clinician.

Prevalence of peanut allergy

In the general population, the prevalence of PA is approximately 1.5% when the diagnosis is based on OFC or highly convincing history, and 0.2% to 0.4% when it is based on OFC alone.[1] These values may differ based on age, race, ethnicity, and geography, but the evidence is not available to precisely determine what those differences are. Recent Australian data representative of the greater Victoria province in one year olds suggests the rate of peanut allergy could be as high as 3%, with as many as 23% of these cases resolving by age 4, and 31% by age 6.[2-5] US estimates range between 1.4%-4.5%, based on various indirect methods including phone surveys, internet surveys, and analysis of clinical history and epinephrine prescribing patterns.[6-8] As well, the prevalence of peanut allergy may change with age. Prevalence estimates also vary depending on how peanut allergy is defined. Many studies use peanut sensitization (at a particular level of detection) to define peanut allergy, while others accept a convincing history of a clinical reaction.[6,9,10] However, the criterion standard is an oral food challenge in which a clear outcome based on peanut ingestion is determined.[4]. Unsurprisingly, reported prevalence rates are higher in studies that include patients diagnosed based on either peanut sensitization and/or a reported convincing clinical history compared to estimates derived from patients diagnosed objectively through OFC. However, there may be some ethical and practical concern in performing OFC for the purpose of confirming prevalence rates using this criterion standard in such aforementioned individuals who already have a clinical diagnosis.[6,8] Understanding the prevalence rate of any allergy helps to determine the relative likelihood that any patient being evaluated could have the allergy, and sets the basis for interpreting any diagnostic test that may be able to infer likelihood of diagnosis through simple tools like Fagan nomograms.[11,12] Therefore, it is essential for a clinician to understand how and when performing specific diagnostic tests would provide the highest (or lowest) utility, to help gauge when such tests would be of value in clinical decision-making.

Making the Diagnosis

Available diagnostic tests for assessing peanut sensitization

Peanut specific IgE can be assessed with either a skin prick test (SPT) or a serologic in-vitro (blood) test. SPT assesses the presence of sIgE through formation of a wheal and erythema following percutaneous introduction of the target allergen. SPTs are based on extracts of whole peanut and therefore do not provide information about sensitization to individual peanut proteins (peanut components), though

extracts of recombinant components have been studied in research situations. Prick-to-prick testing with ingestible peanut products (e.g., peanut butter, powder, or kernels) as an alternative to testing with peanut extracts has been advocated by some, but the reproducibility, validity and reliability of this procedure is not established as a marker of sensitization, and this additional test in combination with the clinical history has uncertain value for clinical decision-making.[11]

A multitude of in-vitro tests for specific IgE are available using a variety of technologies. Modern-day serologic IgE tests rely on allergens that are attached to a solid phase substrate and detect IgE bound to those allergens using anti-human IgE antibodies conjugated to enzymes that create a colored (enzyme-linked immunosorbent assay or ELISA) or fluorescent (fluorescent enzyme immunoassay or FEIA) product. There also are technologies that measure the capture of specific IgE bound to allergen in liquid phase with subsequent detection using an appropriate enzyme-substrate. The amount of sIgE is determined by comparing the dilution curves of the unknown samples with a calibration curve based on samples with known sIgE.[11] Non-specific IgE binding resulting in false positive results (e.g., falsely indicating sensitization) is a potential risk when samples are assessed from patients that are known to have high total IgE levels, but is accounted for by the manufacturer in how the instruments are calibrated.[6] Generally, these tests are considered to be quantitative and to have a relatively low coefficient of variance (e.g., approximately 5%). Most commercially available tests for peanut-specific IgE measure sIgE directed at an extract of whole peanut, similar to what is used in skin testing. However, most allergens contain multiple epitopes, each of which may be associated with the ability to specifically bind IgE, and the potential for resulting distinct symptom patterns.[13] Patients may be sensitized to one or more components, which represent major allergens within peanut that IgE can bind to (such as the major allergens Ara h 1, Ara h 2, or Ara h 3; Ara=arachnic hypogeae, the Latin name for peanut, and major allegens are named based on their Latin names in the order of their discovery). There are now commercially available tests to measure select peanut components. Components are not available for skin testing outside of the research setting.[13]

Evaluation of Suspected Peanut Allergy

To properly use any allergy diagnostic test to evaluate for possible peanut allergy, the pre-test probability must be determined, which is accomplished through taking a comprehensive history.[11,14] Typically, patients present to a clinician for an evaluation of a suspected history of peanut allergy, usually having experienced symptoms (in some form) believed to be attributable to peanut ingestion, which represents a situation in which there is high pre-test probability. However, sometimes tests are run on individuals without such a history (possibly as part of a diagnostic testing panel), such as someone who has never eaten peanut before, or even in individuals who eat peanut and do not develop symptoms. As a general rule, persons who can eat peanut without developing symptoms are by definition not allergic and should not be tested for peanut allergy. The situation is a bit more nuanced when considering an individual for testing who has never before ingested peanut, or in someone where oropharyngeal symptoms most consistent with pollen food allergy syndrome present distinctly, in the absence of other typical IgE mediated symptoms. In general, the pre-test probability for allergy would be very low, so that even if the test were detecting sensitization, the post-test odds would remain low. However, there may be certain situations where a patient who has never before ingested peanut has other risk factors, such as moderate or severe eczema poorly responsive to therapy or a history of other food allergy, which may elevate the pre-test probability above that of the general population (but still lower than someone presenting with a history of a suspected reaction). In these scenarios, the clinician may desire to test these patients given the pre-test probability is potentially elevated or for more practical reasons such as if the test result will help the patient to make a decision whether they will introduce peanut. This is an example of preference-sensitive care, and requires delicate handling of the risks and benefits of all available options of how to manage detectable sensitization on testing with lower yet still elevated pre-test probability. With a detectable sensitization obtained in this context, performing an OFC (presuming both clinician and patient are willing) can be very helpful but needs to be balanced by how strongly the clinician and patient believe the positive test result indicates a high probability of allergy and the understanding of the risk and downstream consequences of a conflating sensitization and allergy.[14,15]

However, most cases do not present asymptomatically. In assessing the clinical history, close attention should be paid to the nature of the presenting symptoms (to make sure these are consistent with mast-cell mediator release characteristic of an IgE mediated reaction), and the timing of when these symptoms developed in association with known or suspected peanut ingestion. Symptoms typically

develop within minutes to up to about 2 hours if they are related to the peanut ingestion, and rarely develop outside this time window. Non-classical symptoms or time courses that fall outside this interval should decrease the suspicion of peanut allergy, though the clinician may have to consider the significance of an eruption/exacerbation of atopic dermatitis in a child potentially associated with peanut ingestion several hours after ingestion.[14,15] Diagnostic testing in the patient with a reasonable pre-test probability, established by eliciting a concerning or likely history of symptom development attributable to peanut ingestion, can then be used to help determine the likelihood of a clinical allergy.[11,12] This describes a high-utility setting of how such tests can be used. One exception of note is food protein induced enterocolitis syndrome (FPIES) to peanut. This is a non-IgE but immune-mediated reaction, which has a delayed onset presentation (typically 1-4 hours after ingestion), resulting in protracted vomiting to the point that lethargy and color change result, and in rare instances, bloody diarrhea may result at 6-12 hours. These symptoms represent this very distinct entity, which is hallmarked by isolated GI involvement. FPIES is a clinical diagnosis, and testing for the presence of IgE for peanut FPIES is not recommended. FPIES diagnosis and management is discussed elsewhere, and this document does not refer to peanut FPIES management.[16]

Potential Exceptions for Testing

A major possible exception are high-risk infants being considered for early peanut introduction. As specified in the 2017 NIAID Addendum Guidelines for the prevention of peanut allergy, a special case may be made for screening infants who present with moderate to severe atopic dermatitis in the first 4-6 months of life that is poorly controlled despite escalating skin care.[17] In formulating the Addendum Guidelines for the Prevention of Peanut Allergy, an expert panel appointed by the National Institutes of Allergy and Infectious Disease recommended that this presentation in these infants represents an elevated pre-test probability of some likelihood of "pre-existing" peanut allergy (based on data from the Learning Early About Peanut Allergy Study which used these particular risk factors). Therefore, in this highly specific subgroup the guidelines do recommend strong consideration that either peanut SPT or sIgE testing be obtained and interpreted before early peanut introduction in these infants. However, outside of this very circumscribed group, there are otherwise no formal recommendations that any individual should have peanut SPT or sIgE testing before peanut introduction specifically as a screening measure for risk-assessment.[17]

Historically, another potential exception involved testing children with moderate to severe atopic dermatitis to the common 8 food allergens (including peanut), even if these foods were never previously consumed. This practice reflected a concern that eczema is a precursor symptom of and a significant risk factor for developing food allergy, and represents a situation where the pre-test probability is potentially raised over that of the baseline general population to some degree. In these children, a diagnosis of allergy was typically made based on research that extrapolated positive predictive values taken from groups of children at referral centers with severe eczema who underwent oral food challenge.[18] In recent years, this practice has largely fallen out of favor as there has been better understanding of a) the limitations of sensitization as a determinant of clinical allergy, b) the pathogenesis of atopic dermatitis occurring independently and not as a marker pathognomonic for undiagnosed food allergy, c) the risks of prolonged allergen avoidance as a factor that may paradoxically increase the risk of food allergy development, and d) the observation that indiscriminant "screening creep" was occurring in children without risk factors or overt symptoms and the predictive values were being used to establish "diagnosis" out of their very tightly established context. [1] The underlying properties of the diagnostic tests themselves make their use as diagnostic screening measures perilous, given they are poorly specific and of optimal utility in the setting of high pre-test probability. Asymptomatic, clinically irrelevant peanut sensitization is common.

Interpreting peanut allergy sensitization

Allergy testing only confirms or refutes the presence of sensitization, requires "clinical correlation" not unlike a radiographic image, and does not independently diagnose allergic disease. Pre-test probability can be translated to post-test odds, using the positive or negative likelihood ratios associated with the sensitivity and specificity of these tests, which can then be used to provide a recommendation regarding diagnosis.[11,12] Thus the presence/absence of sensitization increases or decreases the estimated likelihood that a patient may experience a reaction following peanut ingestion. The final probability of reaction is dependent both on the pre-test probability and the characteristics of the diagnostic test. While this can be translated using a Fagan nomogram, [12] the process is rather intuitive in clinical practice in many situations. Individuals with a strong history (e.g., high pre-test probability) who are sensitized

above a critical threshold can be more confidently diagnosed with peanut allergy, and a person with a non-specific/weak history (e.g. low pre-test probability) and a negative or equivocally positive test indicating the presence of sensitization can be more confidently assessed as not having peanut allergy. In individuals with more questionable histories with a less clear pre-test probability, the test positive or negative likelihood ratio then becomes more crucial in influencing the direction of the decision-making, and ultimately diagnostic confidence may be low enough that an oral food challenge (OFC) still may be necessary to definitively establish diagnosis.[1,14,15,19]

Clinical Conundrums Related to Testing

As alluded to earlier, there are situations where the clinician may encounter a patient in whom testing was potentially inappropriately obtained, such as in a person with no risk-factors and no history of peanut ingestion leading to symptoms. These individuals may be peanut sensitized, but the sensitization is difficult to interpret given the lack of clinical data to determine context of the test value. Here we see two possible management choices. In clinical practice, many may follow prior data establishing positive predictive values (most representative of small populations of eczematous children undergoing OFC at a referral center)[18] for large skin tests or elevated peanut sIgE that may result in a potential misdiagnosis of peanut allergy leading to unnecessary avoidance. Alternatively, this could be viewed as a situation where a test was obtained with low pre-test probability, requiring OFC to provide diagnostic clarity.[20] Another conundrum is the use of so-called "alternative tests" for peanut allergy that are becoming popular, and are frequently utilized by non board-certified allergists or marketed directly to patients to order for use at home without provider involvement. Testing for peanut-specific IgG4 in either the symptomatic or non-symptomatic patient is not indicated, and no role for IgG4 levels in the current diagnostic paradigm exists.[21,22] The role of IgG4 is not well understood, but in studies of food oral immunotherapy and pollen/venom immunotherapy, IgG4 levels to the allergen in question have been noted to increase as the patient becomes desensitized. As such, no defined association between allergic reactivity and IgG4 levels exists. In addition, a multitude of other non-validated alternative tests are utilized by alternative medicine practitioners but have no role in the diagnosis of peanut allergy. This includes Mediator Release Testing, ALCAT testing, Nambudripad's Allergy Elimination Technique, muscle-provocation testing, electrodermal analysis, and hair/urine analysis.[21,22] Providers should be aware of these tests, as well as the

lack of evidence supporting use, as patients may either request such testing, or have already been subjected to them. Both the AAAAI and the ACAAI have discouraged use of these alternative tests.

Utility of the Oral Food Challenge (OFC)in Diagnosing Peanut Allergy

The OFC remains the criterion reference standard test to define peanut or any food allergy.[1,14,15] The OFC generally provides a definitive diagnosis as the outcome is apparent—under medical supervision to observe the outcome, either the person will tolerate ingestion or react. OFCs are rarely indeterminate, so long as the patient can cooperate and ingest the full challenge dose, or subjective symptoms can be avoided. While the double blind, placebo-controlled food challenge is considered the most objective form of OFC (and decreases the likelihood of subjective symptoms complicating interpreting the outcome), open OFC's are usually sufficient for clinical diagnosis and are more practical to conduct, though this has not been directly studied for comparison and represents expert opinion.[1] Inherent in the label "challenge", this implies the outcome is not known beforehand, and thus any challenge carries a risk of a potential allergic reaction, including anaphylaxis, so the clinician must be prepared to potentially treat, and the patient be made aware of such risks.[1,19] Detailed guidance on conducting OFCs in patients is provided elsewhere. [23,24] OFCs are considered both time- and resource-intensive by some, and require dedicated office space and provider expertise, which may make them less appealing to some providers to conduct.[25] However, this is a routine office-based procedure with a superb safety record in the hands of experienced providers.[1,23]

A decision to offer an OFC is complex and individualized, and providers approach this with a variable degree of expertise, comfort, and desire to offer the procedure.[25] OFC can be used to rule in as well as rule out a diagnosis. However with high pre-test probability, the necessity to offer diagnostic OFC may be low (e.g., when either the outcome is very likely to result in a reaction, or very likely to be tolerated).[1,18,23,26] This procedure becomes of greater importance when the probability of having had a reaction to peanut is poorly determinable based on pre-test probability, and testing does not provide much assistance in formulating post-test odds. In this context the OFC can provide an objective outcome to inform decision-making. However, while in such situations there may be obvious utility to perform an OFC, the decision to ultimately do so may depend on patient-specific and provider-specific factors like anxiety, vulnerability, desire to eat peanut as well as the clinical judgement and willingness of the

clinician to perform the procedure.[14,23,26] Patients and families that are particularly anxious about eating peanut might prefer to avoid peanut, even with a lower probability of reaction, rather than undergo OFC.

Overview of guideline development process

This practice parameter was developed using the GRADE (Grading of Recommendations, Assessment, Development, and Evaluation) approach. GRADE is a well-established methodology for developing evidence-based guidelines, detailed elsewhere.[27-29] In formulating the replies to four key questions we took into account the quality of evidence for treatment efficacy, combining this with patients' safety, achieving adherence, and cost. Table 1details the GRADE recommendations and evidence ratings. For more details of the GRADE process please see appendix 1.

In 2017, the Joint Taskforce on Practice Parameters submitted a concept for a peanut allergy clinical practice guideline (which replaces the former nomenclature used, practice parameter) to the AAAAI/ACAAI parent organizations. The JTFPP identified 4 liaisons to help identify content experts to form a working group. Historically, the practice parameters have been evidenced based documents, usually covering many aspects of an allergy-related topic, e.g., diagnostic testing. The initial concept of the peanut diagnostic guideline was of a limited guideline answering only a few questions but developed similar to the previous practice parameters. However, during late 2017 and 2018, the workgroup and JTFPP decided to use the GRADE process to develop this guideline. The workgroup conducted periodic calls to develop central questions to be answered through systematic reviews using the GRADE process, develop a search strategy to identify and review the relevant literature. The working group was divided into individual subgroups to evaluate the identified literature and draft the recommendations based upon the GRADE analysis, and following AMSTAR-2 criteria for systematic reviews.[30] A working draft was prepared by the workgroup, which was then reviewed and modified by the JTFPP. Both groups were provided the opportunity to comment, propose changes, and approve or disapprove each statement. Consensus was sought and reached for each recommendation's direction and strength. Actual or potential conflicts of interest were disclosed annually and transparency of discussion was maintained. A final draft was then approved by the JTFPP and sent to AAAAI and ACAAI appointed reviewers who were asked to comment on the statements and the rationale within free text fields. All these comments and suggestions

were discussed during an JTFPP teleconference. For each comment or suggestion, the JTF evaluated whether the statement needed to be adapted, again taking into account the balance between desirable and undesirable consequences of the alternative management strategies, the quality of the evidence, and the variability in values and preferences.

Concurrent with the AAAAI and ACAAI review, a working draft of the guideline was then posted on the AAAAI, ACAAI, and JTFPP websites for all members and the public at large to review. For each comment or suggestion, the JTF evaluated whether the statement needed to be adapted, again taking into account the balance between desirable and undesirable consequences of the alternative management strategies, the quality of the evidence, and the variability in values and preferences. The finalized draft was then sent to this journal for additional peer review before publication.

GRADE Methodology

Development of Searchable Questions

Prior to conducting a literature search, 4 pre-specified PICO (Population, Intervention, Comparator, Outcomes) format question were formulated by the workgroup and the JTF as per standard GRADE approach. [31] The population for study included published data for patients with known or highly suspected peanut allergy, who underwent oral food challenge (open or blinded) to establish/confirm a clinical outcome of peanut allergy in at least 50% of participants, where both serologic assessment of peanut allergen components (Ara h 1,2,3,6,8) and/or prick skin testing to whole peanut extract or sIgE testing to whole peanut were obtained as markers of peanut sensitization.

The questions developed were the following:

1. In adults and children with a history of suspected peanut allergy and requesting evaluation what are the indications to perform or not perform diagnostic test(s)?

Population: Adults and children presenting for the evaluation of suspected peanut allergy
Intervention: Perform a diagnostic test for peanut allergy based upon history provided
Comparator: Not perform a diagnostic test for peanut allergy based upon history provided
Outcomes: Accuracy of history in determining need for diagnostic testing for peanut allergy

2a. In the patient presenting for evaluation of suspected peanut allergy, should the provider use a skin prick test, a serum-specific IgE test, or both?

Population: Adults and children presenting for the evaluation of peanut allergy
Intervention: Using skin prick testing (SPT), serum specific IgE to whole peanut (sIgE) or both to determine peanut sensitization to assist in the diagnosis of peanut allergy
Comparator: Oral food challenge
Outcomes: Diagnostic accuracy of peanut allergy testing (true/false positive, true/false negative tests)

2b. In the patient presenting for evaluation of suspected peanut allergy, does testing peanut components in addition to SPT or sIgE whole peanut increase diagnostic accuracy?

Population: Adults and children presenting for the evaluation of peanut allergy
Intervention: Using peanut component testing, e.g., Ara h 2, in addition to SPT or sIgE whole peanut to determine peanut sensitization to assist in the diagnosis of peanut allergy
Comparator: Oral food challenge
Outcomes: Diagnostic accuracy of peanut allergy testing (true/false positive, true/false negative tests)

3. In the patient presenting for evaluation of suspected peanut allergy, do the results of diagnostic tests for peanut allergy, in addition to the patient history, help to predict the severity of a future allergic reaction to peanuts?

Population: Adults and children presenting for the evaluation of suspected peanut allergy
Intervention: Performing a diagnostic test(s) for peanut allergy to help predict the severity of a future allergic reaction to peanuts
Comparator: Predicting the severity of a future allergic reaction to peanuts based solely upon the history and without the use of a diagnostic test for peanut allergy
Outcomes: Accurate prediction of the severity of a future allergic reaction to peanuts

Literature Search and Study Eligibility

In conjunction with a medical librarian (KS), a detailed pre-specified search strategy was developed, with input from the working group, as well as based on recently published systematic reviews on peanut allergy diagnostic testing. Study selection was limited to human subjects of any age who were seeking evaluation for the diagnosis of peanut allergy, English language studies published or in press starting from 1946-2018. The finalized search parameters were then independently run on Medline (PubMed 1946-2018) and Embase (Elsevier 1947-2018) databases, with the results combined and filtered for duplicates. A total of 1,314 potential references were identified and transferred into Covidence for review by 4 taskforce members (MG, MS, JW, JO), where 127 studies were identified for full text review by the same 4 authors, resulting in a final selection of 89 studies for data extraction pertaining to searchable questions under GRADE format. (Figure 1a-d, overall PRISMA diagram and diagrams by individual searchable question; Appendix 1, literature search strategy). The search results were combined and culled for duplicate entries, then uploaded into Covidence, where a minimum of two study team members independently reviewed each study for eligibility for full-text review, to determine inclusion, with this process repeated to determine the final studies for data extraction. Conflicts regarding inclusion were resolved by a third study team member. Studies where OFC was not performed as part of the assessment accompanying the diagnostic testing were excluded (including cohort and observational studies based on patient-reported or chart-reported history of peanut allergy involving the use of the aforementioned diagnostic tests without OFC confirmation) but was inclusive of either prospective, retrospective, cross-sectional, or case-control methodologies from both pediatric and adult populations. The full-text versions of the final studies meeting inclusion were reviewed for data extraction of the measures of diagnostic accuracy including sensitivity, specificity, positive/negative predictive value, and the number of true positives, false positives, true negatives, and false negatives. No individual patient level data was sought. Individual study authors were contacted to provide additional data for the following reasons:

1) To clarify information pertaining to number of successful and non-successful challenges relative to a reported cut-off level of the test in question, where such data was not available or calculatable, so that sensitivity and specificity could be calculated (e.g., obtain the cells to inform true/false positive and true/false negative according to our pre-specified thresholds)
2) To request data not presented/analyzed in the selected paper according to the cut-off levels chosen as part of this review, to enable re-tallying of the true/false positive and true/false negative cases

3) To see if additional data regarding other searchable questions was potentially available, that had not been published

Studies selected for data extraction were excluded if the aforementioned measures of diagnostic testing accuracy were not directly reported in the manuscript; upon final review the population, use/application of the index test, use/application of the reference standard was deemed to not fit the pre-specified inclusion criteria; or the study team could not/did not provide the requested additional details for more tailored data to be reported per our extraction parameters upon being contacted to provide this information.

Outcomes and Data Synthesis

Based on the diagnostic test used, the extracted number of true positives, false positives, true negatives, and false negatives with respect to oral food challenge outcome were recorded into a MS Excel spreadsheet, as classified by a conservative cut-off level of these tests (for diagnosis, >0.35 KU_A/L for sIgE and Ara h 2 sIgE, ≥3mm for SPT; for severity >50 KU_A/L for sIgE, >2 KU_A/L Ara h 2 sIgE, ≥10 mm for SPT) relative to the oral food challenge performed in the study. To assess potential influence of Ara h 6 and Ara h 8 on diagnostic accuracy, pre-specified subgroup analyses were planned based on data availability. Meta-analysis of the pooled sensitivity, specificity, positive and negative likelihood ratios (with visual display of these ratios) on a Fagan Nomogram set to a range of potential lower (30%) and higher (70%) situational pre-test probabilities of a patient having peanut allergy. Data analysis was performed in Stata, version 15 using the MIDAS command (peto method, random effects model).[32] Study heterogeneity was reported by the I^2 statistic. Risk of bias was assessed using the QUADAS-2 tool. Publication bias was assessed using funnel plots when possible. [33] GRADEpro software was used to construct the evidence profiles and calculate the absolute effects.[34] Pre-specified sensitivity analyses were planned to explore inclusion only of trials with double blinded challenges as opposed to other challenge types, to assess the effects of geographical region of study, and pediatric vs. non-pediatric studies if permissible. Additional post-hoc sensitivity analyses were performed to verify impact of inclusion of any study on the estimates where there was elevated risk of bias based on patient selection and flow/timing, comparison of individual pooled test precision where SPT/sIgE, sIgE/Ara h 2, or all 3 tests were simultaneously performed, (which per the joint task force was prioritized as the top sensitivity analysis to

report despite this being post hoc, given it most directly answers the searched questions). Data were additionally synthesized narratively. The systematic review process followed AMSTAR2 criteria.[30] Lastly, cost-effectiveness analysis using simulated cohorts with Markov modeling over a 20-year horizon, from a societal perspective, was performed to assess simulated health and economic benefits of the use of the individual diagnostic tests (see supplemental methods).

A working protocol for the parameter and the systematic review was devised by the JTFPP liaisons and registered with PROSPERO.

Reaching workgroup consensus on statements and conclusions:

Where GRADE was not appropriate to answer a particular question, the workgroup employed a modified Delphi process for the determination of the "Strength of the recommendation" and the "Statement profile" for each question. The Delphi method is a structured, interactive, decision-making process used by a panel of experts to arrive at a consensus when there are differing views and perspectives. [35-37] For any statement or conclusion in which there was a difference of opinion, a modified Delphi method was used. Workgroup members provided anonymous answers via email to the JTFPP administrative director (AD) to the questions being considered. The AD provided via teleconference an anonymous summary of the experts' answers and reasons they provided for their responses. The workgroup members discussed all the answers and then were encouraged to modify their answers on the next round(s) of email voting and teleconferences until a consensus was reached.

Results

Question 1: Should diagnostic testing for peanut allergy be performed in adults and children with a history of suspected peanut allergy who are requesting evaluation for peanut allergy?

Recommendation 1a: We suggest in favor of diagnostic (skin prick or serum sIgE) testing for peanut allergy in patients with a 1) physician-judged high pre-test probability of peanut allergy, or 2) prior to an oral food challenge for patients with moderate pre-test probability of peanut allergy, with

whom shared decision-making has been employed to arrive at the final decision**. Conditional recommendation; Certainty of evidence: very low**

Recommendation 1b: We suggest against diagnostic testing in patients where there is low or very low pre-test probability of peanut allergy. **Conditional recommendation; Certainty of evidence: very low**

Agreement by the workgroup (By Delphi: 1a 9/9 agree; 1b 9/9 agree).

Quality of Evidence: This question was determined to not be searchable under GRADE format.

Evidence Summary

This question was not searched in a systematic manner as the content experts were unaware of any single research study that addressed this question. However, expert evidence was collected both from the content experts, the JTFPP, and the known prior literature most relevant to this topic. Expert evidence differs from expert opinion, in that the former does not include a judgment on the evidence and offers a systematic and transparent appraisal of the evidence.[38]

Discussion

Testing for peanut allergy is of the highest utility when there is a history of a known or suspected ingestion of peanut leading to symptoms of an IgE mediated reaction. The identification of individuals for whom testing is indicated requires careful consideration of the clinical history and of epidemiologic risk factors which may increase or decrease the odds of having peanut allergy (e.g., severe atopic dermatitis or another food allergy). Persons with no history of peanut ingestion or an unknown history of ingestion (without other potential risk factors for developing food allergy), or who asymptomatically ingest peanut with impunity should generally not be tested for peanut allergy.[14,15] The estimated pre-test probability of peanut allergy in these situations is very low, and in most circumstances detection of sensitization will not shift the post-test odds of diagnosis appreciably and will require peanut challenge to resolve the diagnosis. Peanut allergy testing itself is not diagnostic of peanut allergy, as asymptomatic sensitization is somewhat common.[1] Therefore, identifying individuals with a strong pre-test probability for peanut

allergy is imperative in the optimal use of diagnostic testing and making an accurate diagnosis of peanut allergy.

Apart from the high-risk infant meeting NIAID addendum 1 criteria, there are potential situations where some providers may ascribe a higher pre-test probability of peanut allergy to a child who has never eaten peanut, and feel that testing may be desired. These generally apply to peanut naïve individuals with other potential risk factors for developing food allergy (e.g., moderate to severe eczema and/or other food allergy), where the pre-test probability may be variably elevated but generally perceived as greater than that of the general population, though still lower than someone with a suspected reaction history. For example, consider the cases of the younger sibling of a peanut allergic child whose family is reluctant to introduce peanut; a child with milk, egg, tree nut or other food allergy; or the child with delayed peanut introduction for other reasons. The decision to test in these circumstances represents a preference-sensitive care option, and in the context of shared decision-making and a thorough explanation of the risks and benefits associated with the preference-sensitive care choices, testing for peanut sensitization may be a reasonable choice. This choice is subject to shared decision making with the patient, and consideration of the risks and benefits of the potential use of oral challenge to help confirm the test results, the magnitude of the degree to which the risk is appreciably different than that of the general population, as well as the potential for the likelihood and consequences of overdiagnosis resulting from detection of asymptomatic peanut sensitization if a challenge is not performed. No decision-aid for this has been developed, however, though this would be potentially useful.

To some degree, clinicians should be advised that they should be prepared to offer oral food challenge to patients where the pre-test probability is no higher than moderate, uncertainty remains, and the patient still desires testing. The risks and consequences of a diagnosis of varying potential accuracy or probability related to a potentially false positive detection of sensitization may or may not outweigh the potential benefit gained through an at-home introduction or an in-office OFC for some families. Table 2 details some considerations for these situations. Testing the younger sibling of a peanut allergic individual (who does not otherwise meet the addendum 1 high-risk criteria) before peanut introduction has not been shown to be cost-effective unless: a) the baseline prevalence of peanut allergy in younger siblings is >11%; b) that every peanut sensitized child undergoes an OFC to determine actual outcome; and c) the health utility detriment from the initial reaction to peanut was only experienced with at-home

introduction and not under an OFC in the office. Without OFC being performed, pre-testing was only cost-effective if the baseline prevalence of peanut allergy in younger siblings was >63%.[39]

More importantly, it is also crucial to consider the patient who presents to the allergist's office with a test indicating detection of peanut sensitization, but has never eaten peanut before. Here, the context (e.g. the presumed pre-test probability) under which the test denoting sensitization was obtained (and its potential interpretation) also requires careful consideration. This as well may represent a situation of a preference-sensitive choice where a role for shared decision-making arises, with consideration for the benefit of performing an OFC to better determine the outcome should be very carefully weighed against the risk of potential misdiagnosis (and recommended avoidance) from a falsely positive test. The presence of the detectable peanut sensitization itself cannot, however, be used as a condition of "elevated" pre-test probability.

Question 2a: In the patient presenting for evaluation of suspected peanut allergy, which of the three tests—SPT, sIgE to whole peanut, or Ara h2 would provide the highest diagnostic accuracy as determined by the more optimal positive/negative likelihood ratio?

Question 2b: In a patient presenting for evaluation of suspected peanut allergy, does testing for peanut components in addition to either SPT or sIgE to whole peanut increase the diagnostic accuracy?

Recommendation 2a: We suggest in favor of Ara h 2 diagnostic testing in a patient presenting for evaluation of suspected peanut allergy for which a single diagnostic test is to be used, as Ara h 2 would provide the best diagnostic accuracy as determined by virtue of more optimal positive/negative likelihood ratios. **Conditional recommendation. Certainty of evidence: moderate.**

Recommendation 2b: We suggest against component testing be sent in addition to either to skin prick test or sIgE to whole peanut to increase diagnostic accuracy. **Conditional recommendation. Certainty of evidence: very low**

Clinical Statement: For GRADE analysis, Ara h 2 was compared to skin prick test and sIgE to whole peanut for the diagnosis of peanut allergy. Providers can interchangeably use either SPT or serologic testing for whole peanut extract IgE, taking into account availability of the test, patient preference, safety, cost, and whether there are patient factors that preclude use of one or both tests. Both tests have high sensitivity but poor specificity in identifying oral food challenge reactive patients at cut-off levels of 3mm wheal size SPT or 0.35 KU_A/L peanut-specific IgE. No data were available regarding use of the tests in tandem or reflexively. In sensitivity analyses where both tests were available, there was minimal difference in the overall sensitivity/specificity between these modalities, and these were similar to the precision in the base analyses of each test individually. However, as a single stand-alone test, compared to either SPT or sIgE testing to whole peanut extract, Ara h 2 has the most optimal combination of positive and negative likelihood ratio, and has drastically enhanced specificity, likely decreasing the number of false positive cases where sensitization is detected. Despite the test characteristics, future research is needed to better clarify if Ara h 2 should be used as a stand-alone measure of peanut sensitization in the patient seeking evaluation for possible peanut allergy. In studies where Ara h 2 was evaluated with sIgE or where all 3 tests were evaluated, the precision advantage for Ara h 2 did not change. A potential risk associated with using Ara h 2 as a stand-alone test is that an allergic individual may be sensitized to other components but not to Ara h 2, though this may be balanced by superior test precision of this approach.

The literature search did not provide patient-level data to determine the value of testing for peanut components in addition to skin prick test or sIgE to whole peanut to increase diagnostic accuracy. Thus, the use and value of components, including reflexive use of Ara h 2, remains a knowledge gap. There is an unclear utility for measuring sIgE to any other commercially available peanut components (Ara h 1, Ara h 3, Ara h 6, Ara h 8, Ara h 9) if peanut sIgE is elevated or SPT >3mm (both indicating sensitization), given the limited available data on performance of components beyond Ara h 2.

Evidence Summary (Questions 2a and 2b):

For SPT and sIgE to whole peanut, from the 89 articles selected for final evidence synthesis, 56 directly pertained to this question. Of these, 32 had data available for extraction (5 studies had no data available, 10 authors did not respond to requests for data, and 9 studies had available data but could not

be analyzed due to zero-cell interactions in the 2x2 table). A total of 18 studies (n=2124 patients) were pooled for evidence synthesis for SPT[40-52] and 30 studies (n=3989 patients) for sIgE.[40-44,46,47,49,50,52-66] No literature was identified that detailed the simultaneous, tandem, or reflexive use of both SPT and sIgE to whole peanut. Figure 2a details the summary forest plot for the pooled sensitivity, specificity, and both positive and negative likelihood ratios for a prick skin test to whole peanut extract of 3mm or greater, and Figure 2b for peanut serum-specific IgE of 0.35 KU_A/L or higher. The summary measures for each test are presented in table 3. Heterogeneity across these studies was high. Figures 3 and 4 detail Fagan nomograms for a practical general example of how to roughly interpret the utility of these tests, set at a pre-specified pre-test probability of 2% (general population prevalence), 30% (low suspicion) and 70% (high suspicion). These nomograms show that the likelihood ratio for sensitization at 3mm or 0.35 KU_A/L at 2% or 30% pre-test probability do not translate to post-test odds >50%, but at the 70% pre-test probability this is raised to ~80%. Negative likelihood ratios do largely decrease post-test odds in all three scenarios. Based on these data, both SPT and sIgE to whole peanut can be used interchangeably, and this is a preference-sensitive choice given no discernable advantage in terms of test precision. There were no data noted that indicate using both tests together was disadvantageous. Both SPT and sIgE to whole peanut have similarly high sensitivity but poor specificity, with serologic testing having slightly higher specificity in identifying oral food challenge reactive patients at the assessed cut-off levels. Table 3 additionally includes sensitivity analysis for the individual sensitivity/specificity of SPT and sIgE assessed when both tests were assessed in the same study. The clinician should be advised of the inherent weaknesses of either of these tests having poor specificity, in that this may preclude to a higher rate of falsely positive detection of peanut sensitization.

For Ara h 2 component-specific IgE, from the 89 articles selected for final evidence synthesis, 41 directly pertained to this question. Of these, 24 had data available for extraction (11 authors did not respond to a request for additional data, 6 articles did not have data available). This resulted in a total of 24 studies (n=2289 patients) pooled for evidence synthesis.[42,43,46,47,53,55,56,60,67-73 49,50,52,61-63,74-76] The summary measures for Ara h 2 are presented in table 3. Figure 5 detail the summary forest plot for the pooled sensitivity and specificity, for Ara h 2 peanut serum-specific IgE of 0.35 KU_A/L or higher. Heterogeneity across these studies was high. Figure 6 details Fagan nomograms for the use of these tests, set at a pre-specified pre-test probability of 2% (population prevalence), 30% (low suspicion) and 70% (high suspicion). These nomograms show that the likelihood ratio for Ara h 2 sensitization at 0.35

739 KU_A/L at 2% or 30% pre-test probability translate to post-test odds of 10% and 70%, but at the 70% pre-
740 test probability translates to 89% post-test odds. Negative likelihood ratios do largely decrease post-test
741 odds in all three scenarios.

742 We were unable to find sufficient number of studies to analyze any other individual peanut
743 components or pool the use of component panels. Therefore, we can offer no comment regarding the role
744 or significance of evaluating these other components individually or in aggregate, or what the clinical
745 implications of their use may be. Similarly, there were no studies identified comparing reflexive use Ara
746 h 2 or any components after SPT or sIgE. There were no studies identified that evaluated the
747 comparative efficacy of Ara h 2 as a stand-alone test compared to any other component or whole peanut
748 PST or sIgE in their use for clinical decision-making. A potential advantage Ara h 2 relative to SPT and
749 sIgE to whole peanut is higher specificity, which may reduce the number of falsely positive cases of
750 sensitization identified, though a disadvantage is this could risk a falsely negative case if someone is
751 sensitized to other components but not Ara h 2. However, the high sensitivity and specificity of the test
752 may limit this risk. In studies where Ara h 2 was evaluated with sIgE or where all 3 tests were evaluated,
753 Ara h 2 consistently had slightly lower sensitivity but much higher specificity, and a more optimal
754 positive/negative likelihood ratio, comparatively. This is similar to the difference noted in the base case
755 where the tests were evaluated individually (Table 3).

756 *Quality of Evidence:* Tables 4a and 4b details the summary of GRADE evidence for both SPT and sIgE.
757 There is moderate certainty of evidence for use of either test, and the estimate was downgraded one point
758 for risk of bias. Table 5 details the certainty of evidence for the use of Ara h 2. There is moderate
759 certainty of evidence, and this estimate was downgraded one point for risk of bias.

760 *Discussion*

761 In practice, SPT and sIgE are often used interchangeably and at the preference of the ordering
762 clinician or the family. Many clinicians may use these tests in tandem with one another as well, though no
763 evidence exists to evaluate this practice. A 2009 systematic review by Chafen et al [77] noted no
764 statistically significant differences between the diagnostic utility of food-specific SPT and sIgE when
765 comparing their summary ROC curves. A 2015 systematic review by Klemans et al noted the sensitivity
766 of peanut SPT was 0.66-1 the specificity 0-0.95, and the positive and negative likelihood ratios between
767 1-3.91 and 0-0.65 respectively. For peanut sIgE, this had sensitivity between .8-1%, specificity between

0-0.63, and positive and negative likelihood ratios between 0.95-2.15 and 0-0.56 respectively.[78] Overall both SPT and sIgE to whole peanut have very similar test precision, with a very slight relative advantage in sensitivity (0.01) and specificity (0.05) for skin testing over sIgE testing. In the setting of the high-risk infant being evaluated for early peanut introduction, the guidelines specifically recommend SPT as the preferred modality when available, though non-allergists can elect to send peanut sIgE and refer patients for further evaluation or recommend at-home introduction in this population.[17] This recommendation is based on data from the LEAP study, suggesting that skin prick testing provided better classification of peanut allergic infants after peanut challenge than serologic testing. [79]

There is widespread availability of component testing and several publications have concluded that Ara h 2 may have unique diagnostic value, which has led to debate about whether clinician should routinely test for IgE to peanut components and base diagnostic decisions solely on these results.[78] In practice, the clinician has the option to request tests for peanut components in combination with whole peanut SPT and/or peanut specific IgE, or request tests for component testing as a stand-alone test. To date, no practice parameter or clinical practice guideline has advocated selective use of one or a panel of components over whole peanut SPT or sIgE, how components including just assessment of Ara h 2 could be used in tandem or reflexively with these tests, or specifically recommend how use of components definitively provides a diagnostic advantage.[1,14,15] There is limited study of other component testing that was found in this literature search. Ara h 6 sensitization is an emerging area of investigation,[80] and one study of Ara h 8 mono-sensitization suggested a potential role in discriminating asymptomatic peanut sensitization from allergy, more likely to have clinical relevance in geographic areas where birch pollen is endemic.[81,82] However, we found few studies that reported challenge-proven outcomes meeting our selection criteria for components apart from Ara h 2, and very limited studies that evaluated use of single vs. panels of peanut components. Thus, we are precluded from commenting any further on specific use of components such as Ara h 6 or Ara h 8, and their potential value in assisting the clinician in making a diagnosis of peanut allergy.

No studies were identified evaluating tandem use of SPT and sIgE to whole peanut. Many studies had both SPT and sIgE measured together, and the individual results are incorporated in the respective analyses. However, offer no recommendation to this tandem approach, perceived to be commonly done in practice. In studies where both SPT and sIgE were reported, the pooled sensitivity/specificity results were very similar to the base analyses, and reflective of those same small differences. Similarly, no

studies were identified evaluating reflexive or tandem use of Ara h 2 or any component with SPT and sIgE to whole peanut, and it is unclear how component testing would be optimally positioned in a clinician's arsenal. Future studies are required to determine if Ara h 2 should be tested as a stand-alone marker, if components should be tested reflexively after sensitization to whole peanut is denoted or even tested at all. Importantly, in the context of either very strong or very weak pre-test probability, it is debatable if components (including Ara h 2) offer any additional diagnostic leverage over whole peanut testing, or supersedes the OFC if there was any doubt. In such circumstances, even the good positive likelihood ratio associated with Ara h 2 would not likely change the clinical decision-making or provide more value than the OFC.

Ara h 2 may have more value vs. other testing options in the context of a questionable history and whole peanut sensitization given its higher specificity, in particular in areas with high birch (or birch cross-reactive) pollen. However, additional research is needed to more robustly evaluate such use, and we noted insufficient numbers of study specifically for this application. There is no universal cut-off value for any component (including Ara h 2) that can used to reliably predict peanut allergy--such levels vary considerably by geographic region, population tested, and possibly by age. [78 83] As was noted in question 1, there may be situations where a clinician may ascribe a higher pre-test probability to child who has never eaten peanut before (apart from those falling under NIAID Addendum 1 recommendations), and desire to obtain Ara h 2 component testing. Overall, use of Ara h 2 at present is limited in the capacity of a corroborating test, indicated when there is sufficient pre-test probability for peanut allergy, and not in the capacity of a screening test where there is no pre-test probability. This is demonstrated in the Fagan nomograms in figure 6 and supplemental figure 1, which may help illustrate practical general examples of how the test may be reasonably interpreted under different hypothetical pre-test probabilities.

There are several other considerations regarding test preference, including safety, cost, patient features that may drive the choice, availability and practice patterns. SPT is associated with an exceptionally rare risk of systemic reactions (0.077%, with 75% of cases attributable to food), though those doing skin testing should be prepared to potentially treat anaphylaxis. [84] There also are data demonstrating that there are more side effects from sIgE testing vs. SPT based on assessment in the NHANES study. The cost of SPT and sIgE tests varies among different offices and laboratories, but has been reported to be from 2-7 times less expensive per test for SPT (typically $3-5 per SPT and $10-20 per

allergen for sIgE test, including components, though components are presently available only as a full panel). (http://health.costhelper.com/allergy-testing.html) Certain patient-related factors may make SPT difficult to perform, such as inability to stop medications with anti-histamine activity, severe dermatographism, unstable asthma, patients who may be averse to or afraid of the procedure (such as young children) and hard to control eczema with extensive skin involvement.[11] However, since SPT can be done on the back or arm or may be possible on other unaffected areas of skin, it is often possible to do the test even with extensive eczema or delay this until the eczema flare has calmed down. The advantage of SPT is that it is a point of care test that can be rapidly performed in clinic, but a trained specialist generally perform this. There are few limitations to sIgE testing, and often multiple allergens can be assessed from 2-5 mL of blood obtained from routine venipuncture. The test is not point of care, however.[11] As was noted in question 1, there may be situations in which a clinician may ascribe a higher pre-test probability to a child who has never eaten peanut before (apart from those falling under NIAID Addendum 1 recommendations), and desire to obtain peanut PST or sIgE. The Fagan nomograms in figures 3-5 may help provide guidance for how the test may be reasonably interpreted in such a scenario.

Test thresholds of 3mm for SPT and 0.35KU_A/L for sIgE and Ara h 2 sIgE were chosen for analysis of this question. These represent sensitization levels at which a patient traditionally would be considered to have a test indicating allergic sensitization. These are the most widely published "cut-off" levels in the literature, though higher levels, including levels indicative of reported positive predictive values have also been reported, and more recently, lower levels of 0.1 KU_A/L are being commonly reported.[78,85] We considered different levels (both higher and lower) but disfavored such an approach as this would have reduced the number of citations that would have been available, and made the analysis even more dependent on the goodwill of authors sending us data reconfigured to our needs. A problem unique to the newer conventions of reporting to the technical lower limit of detection at 0.1 KU_A/L is that many studies otherwise eligible for inclusion in our search were performed before reporting to this lower standard was available, and would have limited our total numbers. More importantly, we are unaware of any literature indicating that sensitization between 0.1 and 0.34 KU_A/L is of clinical significance, as opposed to ample literature that clearly has defined sensitization >0.35 KU_A/L as significant.[1] Lastly, we did not attempt to provide a PPV for these cut-off levels. The PPV is dependent on a population prevalence of disease, which we do not know and did not assess. Instead, we report likelihood ratios and

provide example Fagan nomograms for how the test results could be interpreted at a clinic level, which is a more accurate and appropriate analysis.[86]

Question 3: In the patient presenting for evaluation of suspected peanut allergy, can the results of a diagnostic test be used to predict the severity of an allergic reaction?

Recommendation 3: We suggest against the clinician using the results of a SPT, sIgE to whole peanut extract, or sIgE to peanut components to determine an allergy phenotype or to predict the severity of a future reaction. **Conditional recommendation. Certainty of evidence: very low.**

Clinical statement:

There was inadequate patient-level data to formulate a GRADE recommendation on the use of a diagnostic test for predicting the severity of a future allergy reaction to peanut but a subset analysis did not demonstrate any benefit.

Evidence Summary:

From the 89 articles selected for final evidence synthesis, 31 directly pertained to this question. Of these, 16 had data available for extraction (12 authors did not respond to a request for additional data, 1 study did not have data available). A total of 18 studies were pooled for evidence synthesis (10 for Ara h 2 at 2 KU_A/L, n=845 patients;[42,49,50,52,53,56,61,73,87]13 for whole peanut sIgE at 50 KU_A/L, n=1051 patients;[42,44,49,50,52,56,66,87-90] 12 for SPT 10mm, n=737 patients[42,49-52,61,66,87,88,90]). The summary measures for each test are presented in table 3. Figures 7-9 details the summary forest plot for the pooled sensitivity and specificity for cut off levels for severe reactions for Ara h 2 peanut serum-specific IgE of 2 KU_A/L or higher, whole peanut sIgE at 50 KU_A/L, and for SPT 10mm. Due to both low sensitivity and specificity, with no individual measure greater than 0.68 for any of these analyses, likelihood ratios and Fagan nomograms were not reported. Heterogeneity across these studies was high. Based on these data, this analysis notes exceptionally poor sensitivity and specificity for these cut-off values, which differs from a similar analysis by Klemans et al in a 2015 systematic review where Ara h 2 as a marker of severity was concluded to have more potential. Klemans et al explored several different cut-off levels than we did in this analysis, though did so with far less studies included per cut-off level investigated.[78] Therefore, the results of this analysis should be interpreted as a significant caution to clinicians against

using the degree of sensitization to whole peanut (skin/blood) or peanut component (blood) as a surrogate to determine if someone will have a future severe reaction or has a "severe" reaction phenotype. This caution is pending further future studies of much higher quality, more consistently defining severity, with less selection bias, and with more patient level data for analysis. There were insufficient numbers of other studies to comment regarding the role or significance of evaluating these other components individually or in aggregate to determine if there is any test that may infer reaction severity.

Evidence Strength: Tables 6a-c details the certainty of evidence for the use of Ara h 2, sIgE, and SPT at these stated cut-off levels for the assessment of the severity of a reaction. There is very low certainty of evidence for all three of these measures and this estimate was downgraded one point for risk of bias and two points for inconsistency (based on wide CI's of the pooled studies and a different definition of severity among the studies).

Discussion

There is no relationship indicating that the degree of sensitization is predictive of the underlying severity of the reaction to peanut, using either skin or serologic markers, whole allergen or component. This includes any single test, component, or panel of tests. Importantly, the clinician is advised against making the interpretation that any level of sensitization—high or low—will predict if someone will have a severe reaction or not. Per our meta-analysis, there is no relationship with reaction severity from available data, criteria for severity, and reported cut-off levels. Severe reactions can still occur with low/lower sensitization levels. Multiple practice parameters, guidelines and systematic reviews have repeatedly emphasized these points.[1,14,15] A few individual peanut component-based studies have suggested some degree of association between the recognition of discrete levels of Ara h 2 and history of a severe allergic reaction, though a greater number of studies have noted no such association, and many of these have multiple biases.[78] At our chosen cut-off levels (Ara h 2 2 KU_A/L; PST 10mm, sIgE 50 KU_A/L), we affirm that no relationship exists, though if patient-level data were available for pooling, it is possible a relationship could exist. We caution that there is very serious risk of bias among even the few numbers of studies we included. In particular, many studies did not assess severity using Ara h 2, and small inclusion numbers may present a misleading estimate due to omission of data.

There is potential evidence that singular recognition of Ara h 8 sensitization (in the absence of other component recognition) may be a potential discriminator of pollen cross-sensitization in individuals residing in particular geographic areas who are likely to only experience oropharyngeal, transient itching from peanut ingestion (e.g., pollen food allergy syndrome).[13] However, we could not analyze this question due to low study numbers evaluating this relationship that met inclusion criteria (specifically that 50% of the population underwent OFC). Furthermore, while some expert opinions may support that Ara h 8 monosensitization is a potential indicator of pollen-food allergy syndrome and surrogate for low risk of a severe reaction, these findings lack definitive confirmation in this and prior meta-analysis.[78] Importantly, we found insufficient numbers of studies for components apart from Ara h 2 meeting our criteria to pool for analysis and cannot comment on the clinical utility of these tests without further rigorous study to validate this concept.

Regional geography may influence component sensitization patterns, in particular with the pollen cross-sensitized individuals, which complicate assessing the relationship between sensitization and severity. Two studies have shown differences in component recognition patterns in patients in northern Europe, southern Europe, and the US, as well as differing patterns among different regions in the US which may complicate the use of any particular component as a phenotypic discriminator.[81] For instance, in birch endemic areas, Ara h 8 may behave as a cross-sensitizing marker, and has been proposed to help identify such individuals from those recognizing other proteins in peanut. Ara h 9 could have relevance as a component associated with lipid transfer protein syndrome in certain areas of the world (with high potential to cause systemic reaction in sensitized individuals) whereas elsewhere it behaves similarly to Ara h 8 as a marker of tree pollen sensitization.[13] Therefore, it is unclear the degree to which severity of a reaction may be affected by such geographical differences influencing component recognition, and this area of component research remains promising, but at present represents a knowledge gap.

Importantly, there are issues of bias that must strongly be considered regarding the studies noting an association between sensitization levels and severity. Most of these studies suffer from multiple biases, the most concerning of which is patient selection from serum banks within retrospective cohorts, and lack of representativeness of the sample used for analysis. Many of these studies also lack clear comparison to a gold-standard, tended to be conducted only in certain aged samples, and lacked prospective use of an OFC complicating an objective determination of reaction severity. Study of severe reactions is further

hampered given a predilection to not challenge strongly sensitized individuals with a supporting clinical history, as well as ethical considerations to promptly treat reactions when individuals are challenged, which preclude determining how severe a reaction could be.

The cut-off levels chosen for this analysis were based on review of the literature, where we could include the maximal number of studies, and represent realistically large sensitization levels. For reasons discussed previously, we do not report to the lower limit of detection, other levels of sensitization, or attempted to derive a PPV for severe reactivity.

Sensitivity Analyses

In our protocol we pre-specified sensitivity analyses based on OFC type, geographical region of where the study was conducted, and patient age. We performed additional post-hoc sensitivity analyses for studies that had high risk of bias where both patient selection and flow/timing were noted to be issues. These results are shown in table 3,7, and supplemental figures 2 and 3.

Risk of Bias Assessment

Risk of bias was assessed using the QUADAS-2 assessment tool. [33] This noted some instances where high risk was noted pertaining to the studies for either risk of bias or applicability. The results of this are detailed in table 8. Sensitivity analyses for all 3 searchable questions were completed after removing studies judged to have high risk for bias based on patient selection and flow/timing of the testing and challenge but this did not alter the pooled sensitivity and specificity estimates to an appreciable or significant degree.

Analysis of Health and Economic Benefits of Peanut Diagnostic Strategies:

Cost-effectiveness of peanut allergy diagnostic options was evaluated with decision analysis informed by results of the meta-analysis of diagnostic operating characteristics of single ara h 2 sIgE, whole peanut sIgE, and skin prick testing (SPT) (Figures 10 and 11). Markov modeling was used in microsimulations of each testing strategy (n=100,000 per strategy). Model assumptions are outlined in Table 9. Age-adjusted all-cause mortality was included over a 20-year time horizon (sensitivity range 5-80 years) with a start age during infancy sensitivity range 0 years to 8 years), a 14% pre-test probability of peanut allergy (sensitivity range 5%-90%) , and an assumption that 20% (sensitivity range of 5%-

20%) of false positive diagnoses were refuted by accidental exposures over the model horizon in the base-case. Costs were expressed in 2019 dollars with future costs and life-years were equally discounted at 3%, and risks of reactions, costs, and utilities of peanut allergy burden of illness were incorporated.

In the base-case analysis at a pre-test probability of 14%, Ara h 2 dominated both whole peanut sIgE and whole peanut prick skin testing, producing greater health benefit in terms of quality-adjusted life years (QALY: Ara h 2 14.69, SD 1.32; SPT 14.36, SD 1.33; sIgE 14.29, SD 1.33. To illustrate the scale of the metric, a 0.1 difference in QALY represents ~36.5 days of life in a year traded in preference of a specific outcome). Ara h 2 screening produced cost savings of $13,960 and $11,530 when compared with whole peanut sIgE and SPT testing over a 20-year time horizon. Ara h 2 did result in a greater rate of peanut allergic reactions per patient screened (Ara h 2: 0.1725, SD 0.6169; SPT: 0.1555, SD 0.5784; whole peanut sIgE: 0.1581, SD 0.5836) but no significant difference in fatality rates (Table 10). At pretest probabilities of 3% and 75%, Ara h 2 continued to dominate analyses with cost saving (compared with SPT, whole peanut sIgE) of $13,065 (SPT), $15,797 (whole peanut sIgE) and $3,489 (SPT), $4,187 (sIgE), respectively. Peanut associated fatality was rare and not significantly different among testing strategies.

The analysis remained dominated in deterministic sensitivity analyses (Figure 12) provided Ara h 2 specificity remained above 0.46. If all patients with negative testing underwent supervised oral food challenge (14% pre-test probability), cost of Ara h 2 was $12,302 (SD, $22,233), SPT $23,853 (SD, $25,404), whole peanut sIgE $26,334 (SD,$25,359) producing respective benefits of 14.69 (SD, 1.32) QALY for Ara h 2, 14.37 (SD, 1.32) QALY for SPT, and 14.30 (SD, 1.31) QALY for whole peanut sIgE. In probabilistic sensitivity analysis (n=10,000) across fatality distributions demonstrated, the Ara h 2 strategy was the most cost-effective option in all iterations (willingness to pay (WTP) of $100,000/QALY).(Figure 13)

While we make no recommendation for or against the use of any component testing in question 3, this simulation, does suggest superior health and economic benefits would be associated with preferential use of Ara h 2 as a stand-alone diagnostic test, assuming these are used in populations similar to those pooled for analysis. Limitations of this analysis include a) use of the meta-analysis inputs, which have outcomes assessed at low cut-off values for sensitivity and specificity; b) lack of prospective validation of OFC proven outcomes when Ara h 2 is the only sensitization marker assessed; c) a knowledge gap in understanding the association of other component recognition in the absence of Ara h 2 recognition in

OFC proven cases of peanut allergy; and d) lack of commercial availability of Ara h 2 as an available stand-alone test. General limitations of the overall analysis are discussed in the next section.

General Limitations of this Analysis

There are multiple limitations to this analysis. Foremost, we were only able to address 4 questions, including one that was not searchable, in the scope of this analysis. This does not imply that there are other factors or issues within peanut allergy diagnostic testing that are less important. The JTFPP did limit the questions asked to 4, for pragmatic reasons to ensure we could produce a GRADE based parameter in the timeframe allotted which conformed to the bylaws set forth in 2016 by the AAAAI and ACAAI. These stated that no new parameter topics will be generated, and that all parameters going forward offer focused updates to formerly published documents using GRADE format. Therefore, this document updates the Diagnostic Testing parameter from 2008,[11] with a focus on the use of diagnostic testing for peanut allergy. GRADE is not the only system for evidence-based reviews, but is the chosen system for the JTFPP. GRADE has multiple noted limitations, including forced downgrading of certainty and strength of recommendation based on particular study attributes, and a general trend that the overall strength of recommendations are rarely strong.[27-29] Peanut components were not commercially available before the latter part of the 2000's and thus this may have introduced not-at-random factors about the types of patients studied in those compared to earlier studies when components were not available. Fairly low cut-off levels were chosen in the analysis for reasons detailed in the sub-sections, but this remains a limitation in that the relative precision of the test may perform differently at different levels.

We found a scarcity of available studies in our literature search that we found which met our OFC criteria and explored use of these tests at a general population level. Therefore, most included studies either involved a referral center cohort, or in many cases, a referral center cohort enriched for patients with known sensitization (skin and/or serologic IgE testing) as selection criteria before being offered OFC. In choosing the selection criteria and evaluating studies for final inclusion, it was felt that this was an acceptable approach given that the specialist clinician would generally be dealing with issues surrounding test interpretation in this population, and be less concerned with false negative rates from the general

population (which the pooled sensitivity and specificity may mis-estimate in this analysis). We have accounted for this by downgrading the risk of bias (on account of risk of bias from patient selection) category in the GRADE certainty of evidence table, which factors into the overall certainty of the recommendations. Additionally, the analyses involve pooling of studies for assessment of severity that did not all use the same severity criteria (they were similar enough to pool but the rankings reflected different criteria that have evolved over time) and most had wide confidence intervals, requiring us to downgrade 2 points for inconsistency.

The limitations of lack of studies evaluating a tandem or reflexive approach, or the robustness of studies pertaining to other components beyond Ara h 2 (necessary to allow for meta-analysis) have already been mentioned, as has the lack of a consistent objective grading criteria as well as the small number of studies evaluating reaction severity, as well as differences noted in the timing/flow and selection processes of each of these studies. This is accounted for in grading the certainty of evidence and risk of bias. As well, the aforementioned sensitivity analyses were done to further confirm if inclusion of those studies felt to be most at risk would alter the estimates, which they did not. We could not stratify by allergic co-morbidity (in particular presence of atopic dermatitis) or age with accuracy due to limited available data in the reporting which would allow for such stratifications to be made, though we did perform sensitivity analysis on challenge type, adult vs. pediatric studies, as well as by region of the world (Europe, North America) in which the data were observed. Statistically, the pooling of data are limited by high heterogeneity, with some included studies having high risk of bias.

Knowledge Gaps

Within in the scope of these questions, multiple gaps in the current knowledge base were identified that could not be resolved through our literature search and meta-analysis. These include, but are not limited to:

a) A lack of identified studies that systematically evaluate when someone should be tested for peanut allergy
b) A lack of identified studies that evaluate the tandem or reflexive use of whole peanut extract SPT and whole peanut sIgE in combination

c) A lack of identified studies that evaluate the tandem or reflexive use of whole peanut extract SPT and whole peanut sIgE in combination with peanut components
d) A lack of identified studies that evaluate the tandem or reflexive use of one or more peanut components
e) A lack of identified studies that evaluate Ara h 1, Ara h 3, Ara h 6, Ara h 8, and Ara h 9 performance, or if severity or reaction phenotypes are associated with recognition of these components
f) A lack of identified studies that consistently or systematically study reaction severity using unified criteria or cut-off markers, or evaluate this question at different cut-off levels
g) A lack of identified studies that study any of the searchable questions at a population level that are less enriched for already sensitized individuals as opposed to within more clustered clinical referral centers
h) A lack of identified studies that trace longitudinal outcomes and natural history of disease to better understand the full scope of the ramifications of diagnostic testing choices to inform best-practices
i) A lack of clear understanding and inconsistent use of diagnostic cut-off points for the use of these tests
j) A lack of consistent reporting at an individual level of allergic co-factors that may influence the performance of these diagnostic tests in relation to the food challenge outcome to assess the influence of such covariates

Text box 2 addresses a number of the key take-home messages and knowledge gaps.

Summary and Conclusions

In making a diagnosis of peanut allergy, it is important to clearly understand the indications for running a diagnostic test. Only patients with a history of peanut ingestion leading to symptom development benefit from peanut allergy diagnostic testing, and should be tested.[14,15,19] With the exception of patients who are not newborn infants under the age of 4-6 months of life who have either egg allergy or severe eczema,[17] there is no indication for any form of peanut allergy testing in someone who has not yet eaten peanut and subsequently developed symptoms of an allergic reaction. Testing only

determines the presence or absence of peanut sensitization and alone does not infer a diagnosis without a history to provide context as to what happens upon peanut ingestion.[11] Use of the tests in these contexts helps translate the pre-test probability of allergy (e.g. based on the history) into post-test odds of a peanut allergy diagnosis.[12] In some cases, an oral food challenge may be necessary to definitively rule in or rule out a diagnosis. In terms of choice of tests, when assessing for whole peanut sensitization, there is little practical difference between use of SPT or sIgE—both are highly sensitive but poorly specific, and may be prone to false positive detection of sensitization in certain contexts. Use of testing to the peanut component Ara h 2 has the best profile of high sensitivity, high specificity, and optimal positive/negative likelihood ratio, and is probably the most accurate single test that is available in terms of a test that could be sent with the lowest potential risk of false positive sensitization being detected. However, how this test should be used in the work up of the suspected peanut allergic patient remains unresolved and not prospectively validated in terms of clinical pathways as to how such properties could be leveraged. We do present evidence herein that shows that using Ara h 2 as a sole diagnostic test in the evaluation of peanut allergy could be cost effective, given the cost-savings at a societal level associated with a significant simulated reduction in the number of false positive cases, as one such possible application of how the test could be used. No whole peanut allergen or component test infers severity of a future reaction, or a reaction phenotype, and attempts to interpret these tests as such should be discouraged given no evidence of a relationship. (Table 11)

Benefits/Harms of Implementing the Guideline Recommendations

Potential Benefits

The potential benefit of this analysis is the appropriate management of patients with Peanut allergy. See the "Discussion" section for each question in the guideline document for benefits of tests. Cost-effectiveness analysis was undertaken to further explore such health benefits. Please refer to supplemental table 1, which details the evidence to recommendation process.

Potential Harms

The potential harms include adverse effects associated with incorrect diagnosis of peanut allergy. See the "Discussion" section for each question in the guideline document for adverse events of specific

interventions. Cost-effectiveness analysis was undertaken to further explore such health detriments. Please refer to supplemental table 1, which details the evidence to recommendation process.

Qualifying Statements

This clinical practice guideline was designed to facilitate informed decision-making on the diagnosis of children and adults with suspected peanut allergy. It was not intended to define a standard of care, and should not be construed as such. It should not be interpreted as a prescription for an exclusive course of management.

Implementation of the Guideline

Description of Implementation Strategy

This practice parameter will be published in XXX, and made available through direct hyperlink on the Joint Taskforce for Allergy Practice Parameters website. To help promote awareness of this new practice parameter and enhance knowledge translation, there are planned lectures at forthcoming national allergy meetings as well as at state/local allergy meetings.

Implementation Tools

A slide deck detailing the key findings in this practice parameter has been developed and is available on both the AAAAI and the ACAAI websites.

Date Released

(publication date) ####

Guideline Developer(s)

The Joint Task Force of Practice Parameters

Source(s) of Funding

American Academy of Allergy, Asthma, Immunology and the American College of Allergy, Asthma, and Immunology

Financial Disclosures/Conflicts of Interest

All members of the peanut diagnosis workgroup and the JTFPP were required to complete a detailed declaration of interest statement' including all current and future conflicts of interest as well as past conflicts of interest restricted to 2 years before joining the workgroup and/or JTFPP. It is felt that excluding all individuals with some degree of potential conflict of interest would prevent the assembly of a workgroup and JTFPP. The authors therefore allowed members of the workgroup and JTFPP to have past financial and/or intellectual conflicts of interest. No consequences were attached to the stated interests, but rather the authors insisted on transparency. All members of the workgroup and JTF were allowed to participate in all discussions and had equal weight in formulating the statements. All were allowed equal involvement in data extraction and writing the rationales.

The declaration of interest forms are available from **www.allergyparameters.org** and are updated on a regular basis.

Contributions of authors

(to revise)

References

1. National Academies of Sciences E, Medicine. Finding a Path to Safety in Food Allergy: Assessment of the Global Burden, Causes, Prevention, Management, and Public Policy. Washington, DC: The National Academies Press; 2017.
2. Peters RL, Allen KJ, Dharmage SC, Tang ML, Koplin JJ, Ponsonby AL, et al. Skin prick test responses and allergen-specific IgE levels as predictors of peanut, egg, and sesame allergy in infants. J Allergy Clin Immunol 2013;132:874-80.
3. Peters RL, Allen KJ, Dharmage SC, Koplin JJ, Dang T, Tilbrook KP, et al. Natural history of peanut allergy and predictors of resolution in the first 4 years of life: A population-based assessment. J Allergy Clin Immunol 2015;135:1257-66.e1-2.
4. Osborne NJ, Koplin JJ, Martin PE, Gurrin LC, Lowe AJ, Matheson MC, et al. Prevalence of challenge-proven IgE-mediated food allergy using population-based sampling and predetermined challenge criteria in infants. J Allergy Clin Immunol 2011;127:668-76.e1-2.
5. Peters R, Koplin J, Ponsonby A-L, Perret K, Dharmage SC, Allen K. The Natural History of Peanut and Egg Allergy and Predictors of Persistence: The Healthnuts Longitudinal Study, 6-Year-Old Follow-up. Journal of Allergy and Clinical Immunology 2019;143:AB421.

6. Gupta RS, Springston EE, Warrier MR, Smith B, Kumar R, Pongracic J, et al. The prevalence, severity, and distribution of childhood food allergy in the United States. Pediatrics 2011;128:e9-17.
7. Sicherer SH, Muñoz-Furlong A, Godbold JH, Sampson HA. US prevalence of self-reported peanut, tree nut, and sesame allergy: 11-year follow-up. Journal of Allergy and Clinical Immunology 2010;125:1322-6.
8. Bunyavanich S, Rifas-Shiman SL, Platts-Mills TAE, Workman L, Sordillo JE, Gillman MW, et al. Peanut allergy prevalence among school-age children in a US cohort not selected for any disease. Journal of Allergy and Clinical Immunology 2014;134:753-5.
9. McGowan EC, Keet CA. Prevalence of self-reported food allergy in the National Health and Nutrition Examination Survey (NHANES) 2007-2010. Journal of Allergy and Clinical Immunology 2013;132:1216-9.e5.
10. Liu AH, Jaramillo R, Sicherer SH, Wood RA, Bock SA, Burks AW, et al. National prevalence and risk factors for food allergy and relationship to asthma: Results from the National Health and Nutrition Examination Survey 2005-2006. Journal of Allergy and Clinical Immunology 2010;126:798-806.e14.
11. Bernstein IL, Li JT, Bernstein DI, Hamilton R, Spector SL, Tan R, et al. Allergy Diagnostic Testing: An Updated Practice Parameter. Annals of Allergy, Asthma & Immunology 2008;100:S1-S148.
12. Nelson A. An Interactive Workshop Reviewing Basic Biostatistics and Applying Bayes' Theorem to Diagnostic Testing and Clinical Decision-Making. MedEdPORTAL : the journal of teaching and learning resources 2018;14:10771-.
13. Sastre J. Molecular diagnosis in allergy. Clinical & Experimental Allergy 2010;40:1442-60.
14. Sampson HA, Aceves S, Bock SA, James J, Jones S, Lang D, et al. Food allergy: A practice parameter update—2014. Journal of Allergy and Clinical Immunology 2014;134:1016-25.e43.
15. Guidelines for the Diagnosis and Management of Food Allergy in the United States: Report of the NIAID-Sponsored Expert Panel. Journal of Allergy and Clinical Immunology 2010;126:S1-S58.
16. Nowak-Wegrzyn A, Chehade M, Groetch ME, Spergel JM, Wood RA, Allen K, et al. International consensus guidelines for the diagnosis and management of food protein-induced enterocolitis syndrome: Executive summary-Workgroup Report of the Adverse Reactions to Foods Committee, American Academy of Allergy, Asthma & Immunology. The Journal of allergy and clinical immunology 2017;139:1111-26 e4.
17. Togias A, Cooper SF, Acebal ML, Assa'ad A, Baker JR, Jr., Beck LA, et al. Addendum guidelines for the prevention of peanut allergy in the United States: Report of the National Institute of Allergy and Infectious Diseases-sponsored expert panel. J Allergy Clin Immunol 2017;139:29-44.
18. Sampson HA, Ho DG. Relationship between food-specific IgE concentrations and the risk of positive food challenges in children and adolescents. Journal of allergy and clinical immunology 1997;100:444-51.
19. Sicherer SH, Sampson HA. Food allergy: A review and update on epidemiology, pathogenesis, diagnosis, prevention, and management. J Allergy Clin Immunol 2018;141:41-58.
20. Fleischer DM, Bock SA, Spears GC, Wilson CG, Miyazawa NK, Gleason MC, et al. Oral Food Challenges in Children with a Diagnosis of Food Allergy. The Journal of Pediatrics 2011;158:578-83.e1.
21. Hammond C, Lieberman JA. Unproven Diagnostic Tests for Food Allergy. Immunology and Allergy Clinics of North America 2018;38:153-63.
22. Kelso JM. Unproven Diagnostic Tests for Adverse Reactions to Foods. The Journal of Allergy and Clinical Immunology: In Practice 2018;6:362-5.
23. Nowak-Węgrzyn A, Assa'ad AH, Bahna SL, Bock SA, Sicherer SH, Teuber SS. Work Group report: Oral food challenge testing. Journal of Allergy and Clinical Immunology 2009;123:S365-S83.

24. Bird JA, Leonard S, Groetch M, Assa'ad A, Cianferoni A, Clark A, et al. Conducting an Oral Food Challenge: An Update to the 2009 Adverse Reactions to Foods Committee Work Group Report. J Allergy Clin Immunol Pract 2020;8:75-90 e17.
25. Pongracic JA, Bock SA, Sicherer SH. Oral food challenge practices among allergists in the United States. Journal of Allergy and Clinical Immunology 2012;129:564-6.
26. Stiefel G, Anagnostou K, Boyle RJ, Brathwaite N, Ewan P, Fox AT, et al. BSACI guideline for the diagnosis and management of peanut and tree nut allergy. Clinical & Experimental Allergy 2017;47:719-39.
27. Brożek JL, Akl EA, Alonso-Coello P, Lang D, Jaeschke R, Williams JW, et al. Grading quality of evidence and strength of recommendations in clinical practice guidelines. Allergy 2009;64:669-77.
28. Brożek JL, Akl EA, Jaeschke R, Lang DM, Bossuyt P, Glasziou P, et al. Grading quality of evidence and strength of recommendations in clinical practice guidelines: Part 2 of 3. The GRADE approach to grading quality of evidence about diagnostic tests and strategies. Allergy 2009;64:1109-16.
29. Brożek JL, Akl EA, Compalati E, Kreis J, Terracciano L, Fiocchi A, et al. Grading quality of evidence and strength of recommendations in clinical practice guidelines Part 3 of 3. The GRADE approach to developing recommendations. Allergy 2011;66:588-95.
30. Shea BJ, Reeves BC, Wells G, Thuku M, Hamel C, Moran J, et al. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. BMJ 2017;358:j4008.
31. Schünemann H, Brożek J, Guyatt G, Oxman A, editors GRADE handbook for grading quality of evidence and strength of recommendations Updated October 2013 The GRADE Working Group, 2013 Available fromguidelinedevelopmentorg/handbook
32. Dwamena B, 2007. "MIDAS: Stata module for meta-analytical integration of diagnostic test accuracy studies," Statistical Software Components S456880, Boston College Department of Economics, revised 05 Feb 2009. Downloaded August 2, 2018.
33. Whiting PF, Rutjes AWS, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2: A Revised Tool for the Quality Assessment of Diagnostic Accuracy Studies. Annals of Internal Medicine 2011;155:529-36.
34. GRADEpro Guideline Development Tool [Software] McMaster University, 2015 (developed by Evidence Prime, Inc) Available from gradeproorg Downloaded August 1, 2018.
35. Dalkey N, Helmer O. An Experimental Application of the DELPHI Method to the Use of Experts. Management Science 1963;9:458-67.
36. Linstone HA, Turoff M. The Delphi method : techniques and applications. Reading, Mass.: Addison-Wesley Pub. Co., Advanced Book Program; 1975.
37. Brown B. Delphi Process: A methodology used for the elicitation of opinions of experts. In: Corporation TR, ed. Santa Monica, CA1968.
38. Schunemann HJ, Zhang Y, Oxman AD, Expert Evidence in Guidelines G. Distinguishing opinion from evidence in guidelines. BMJ 2019;366:l4606.
39. Shaker MS, Iglesia E, Greenhawt M. The Health and Economic Benefits of Approaches for Peanut Introduction in Infants with a Peanut Allergic Sibling. Allergy 2019.
40. Abrams EM, Becker AB. Oral food challenge outcomes in a pediatric tertiary care center. Allergy Asthma Clin Immunol 2017;13:43.
41. Begin P, Graham F, Killer K, Paradis J, Paradis L, Des Roches A. Introduction of peanuts in younger siblings of children with peanut allergy: a prospective, double-blinded assessment of risk, of diagnostic tests, and an analysis of patient preferences. Allergy 2016;71:1762-71.

42. Chinthrajah RS, Purington N, Andorf S, Rosa JS, Mukai K, Hamilton R, et al. Development of a tool predicting severity of allergic reaction during peanut challenge. Allergy 2018;121:69-76.e2.
43. Comberiati P, Colavita L, Minniti F, Paiola G, Capristo C, Incorvaia C, et al. Utility of Specific IgE to Ara h 2 in Italian Allergic and Tolerant Children Sensitized to Peanut. Int J Mol Cell Med 2016;5:160-6.
44. DunnGalvin A, Daly D, Cullinane C, Stenke E, Keeton D, Erlewyn-Lajeunesse M, et al. Highly accurate prediction of food challenge outcome using routinely available clinical data. J Allergy Clin Immunol 2011;127:633-9.e1-3.
45. Johannsen H, Nolan R, Pascoe EM, Cuthbert P, Noble V, Corderoy T, et al. Skin prick testing and peanut-specific IgE can predict peanut challenge outcomes in preschoolchildren with peanut sensitization. Clin Exp Allergy 2011;41:994-1000.
46. Klemans RJ, Broekman HC, Knol EF, Bruijnzeel-Koomen CA, Otten HG, Pasmans SG, et al. Ara h 2 is the best predictor for peanut allergy in adults. J Allergy Clin Immunol Pract 2013;1:632-8.e1.
47. Klemans RJ, Otte D, Knol M, Knol EF, Meijer Y, Gmelig-Meyling FH, et al. The diagnostic value of specific IgE to Ara h 2 to predict peanut allergy in children is comparable to a validated and updated diagnostic prediction model. J Allergy Clin Immunol 2013;131:157-63.
48. Ludman S, Ballabeni P, Eigenmann PA, Wassenberg J. Predicting positive food challenges in children sensitised to peanuts/tree nuts. Pediatr Allergy Immunol 2013;24:276-81.
49. Preece K, Bhatia R, Belcher J, Patchett K, McElduff P, Collison A, et al. The fraction of exhaled nitric oxide improves prediction of clinical allergic reaction to peanut challenge in children. Clin Exp Allergy 2014;44:371-80.
50. Rajput S, Sharma V, Hughes SM, Ewing CI, Arkwright PD. Allergy testing in predicting outcome of open food challenge to peanut. J Allergy Clin Immunol 2018;141:457-8.
51. Song GC, Wang XY, Wang Z, Ruan XL, Yang J, Zhu Z, et al. [Association between serum allergens and asthma in children]. Zhongguo Dang Dai Er Ke Za Zhi 2015;17:806-10.
52. van Erp FC, Knulst AC, Kentie PA, Pasmans SG, van der Ent CK, Meijer Y. Can we predict severe reactions during peanut challenges in children? Pediatr Allergy Immunol 2013;24:596-602.
53. Ballmer-Weber BK, Lidholm J, Fernandez-Rivas M, Seneviratne S, Hanschmann KM, Vogel L, et al. IgE recognition patterns in peanut allergy are age dependent: perspectives of the EuroPrevall study. Allergy 2015;70:391-407.
54. Beigelman A, Strunk RC, Garbutt JM, Schechtman KB, Jaenicke MW, Stein JS, et al. Clinical and laboratory factors associated with negative oral food challenges. Allergy Asthma Proc 2012;33:467-73.
55. Ebisawa M, Moverare R, Sato S, Borres MP, Ito K. The predictive relationship between peanut- and Ara h 2-specific serum IgE concentrations and peanut allergy. J Allergy Clin Immunol Pract 2015;3:131-2.e1.
56. Glaumann S, Nopp A, Johansson SG, Rudengren M, Borres MP, Nilsson C. Basophil allergen threshold sensitivity, CD-sens, IgE-sensitization and DBPCFC in peanut-sensitized children. Allergy 2012;67:242-7.
57. Guilloux L, Morisset M, Codreanu F, Parisot L, Moneret-Vautrin DA. Peanut allergy diagnosis in the context of grass pollen sensitization for 125 patients: roles of peanut and cross-reactive carbohydrate determinants specific IgE. Int Arch Allergy Immunol 2009;149:91-7.
58. Gupta RS, Lau CH, Hamilton RG, Donnell A, Newhall KK. Predicting outcomes of oral food challenges by using the allergen-specific IgE-total IgE ratio. J Allergy Clin Immunol Pract 2014;2:300-5.
59. Johansson EK, Bergstrom A, Kull I, Lind T, Soderhall C, van Hage M, et al. IgE sensitization in relation to preschool eczema and filaggrin mutation. J Allergy Clin Immunol 2017;140:1572-9.e5.

60. Klemans RJ, Liu X, Knulst AC, Knol MJ, Gmelig-Meyling F, Borst E, et al. IgE binding to peanut components by four different techniques: Ara h 2 is the most relevant in peanut allergic children and adults. Clin Exp Allergy 2013;43:967-74.
61. Leo SH, Dean JM, Jung B, Kuzeljevic B, Chan ES. Utility of Ara h 2 sIgE levels to predict peanut allergy in Canadian children. J Allergy Clin Immunol Pract 2015;3:968-9.
62. Lieberman JA, Glaumann S, Batelson S, Borres MP, Sampson HA, Nilsson C. The utility of peanut components in the diagnosis of IgE-mediated peanut allergy among distinct populations. J Allergy Clin Immunol Pract 2013;1:75-82.
63. Nicolaou N, Murray C, Belgrave D, Poorafshar M, Simpson A, Custovic A. Quantification of specific IgE to whole peanut extract and peanut components in prediction of peanut allergy. J Allergy Clin Immunol 2011;127:684-5.
64. Perry TT, Matsui EC, Kay Conover-Walker M, Wood RA. The relationship of allergen-specific IgE levels and oral food challenge outcome. J Allergy Clin Immunol 2004;114:144-9.
65. Wainstein BK, Yee A, Jelley D, Ziegler M, Ziegler JB. Combining skin prick, immediate skin application and specific-IgE testing in the diagnosis of peanut allergy in children. Pediatr Allergy Immunol 2007;18:231-9.
66. Wensing M, Penninks AH, Hefle SL, Koppelman SJ, Bruijnzeel-Koomen CA, Knulst AC. The distribution of individual threshold doses eliciting allergic reactions in a population with peanut allergy. Journal of allergy and clinical immunology 2002;110:915-20.
67. Bernard H, Paty E, Mondoulet L, Burks AW, Bannon GA, Wal JM, et al. Serological characteristics of peanut allergy in children. Allergy 2003;58:1285-92.
68. Beyer K, Grabenhenrich L, Hartl M, Beder A, Kalb B, Ziegert M, et al. Predictive values of component-specific IgE for the outcome of peanut and hazelnut food challenges in children. Allergy 2015;70:90-8.
69. Dang TD, Tang M, Choo S, Licciardi PV, Koplin JJ, Martin PE, et al. Increasing the accuracy of peanut allergy diagnosis by using Ara h 2. J Allergy Clin Immunol 2012;129:1056-63.
70. Ebisawa M, Moverare R, Sato S, Maruyama N, Borres MP, Komata T. Measurement of Ara h 1-, 2-, and 3-specific IgE antibodies is useful in diagnosis of peanut allergy in Japanese children. Pediatr Allergy Immunol 2012;23:573-81.
71. Eller E, Bindslev-Jensen C. Clinical value of component-resolved diagnostics in peanut-allergic patients. Allergy 2013;68:190-4.
72. Keet CA, Johnson K, Savage JH, Hamilton RG, Wood RA. Evaluation of Ara h2 IgE thresholds in the diagnosis of peanut allergy in a clinical population. J Allergy Clin Immunol Pract 2013;1:101-3.
73. Kukkonen AK, Pelkonen AS, Makinen-Kiljunen S, Voutilainen H, Makela MJ. Ara h 2 and Ara 6 are the best predictors of severe peanut allergy: a double-blind placebo-controlled study. Allergy 2015;70:1239-45.
74. Martinet J, Couderc L, Renosi F, Bobee V, Marguet C, Boyer O. Diagnostic Value of Antigen-Specific Immunoglobulin E Immunoassays against Ara h 2 and Ara h 8 Peanut Components in Child Food Allergy. Int Arch Allergy Immunol 2016;169:216-22.
75. Schots M, de Mol AC, Vermeer HJ, Roosen YM, Vriesman AW. Is Ara h 2 indeed the best predictor for peanut allergy in Dutch children? Diagnosis (Berl) 2016;3:31-5.
76. Suratannon N, Ngamphaiboon J, Wongpiyabovorn J, Puripokai P, Chatchatee P. Component-resolved diagnostics for the evaluation of peanut allergy in a low-prevalence area. Pediatr Allergy Immunol 2013;24:665-70.
77. Chafen JJ, Newberry SJ, Riedl MA, Bravata DM, Maglione M, Suttorp MJ, et al. Diagnosing and managing common food allergies: a systematic review. JAMA 2010;303:1848-56.

78. Klemans RJ, van Os-Medendorp H, Blankestijn M, Bruijnzeel-Koomen CA, Knol EF, Knulst AC. Diagnostic accuracy of specific IgE to components in diagnosing peanut allergy: a systematic review. Clin Exp Allergy 2015;45:720-30.
79. Du Toit G, Roberts G, Sayre PH, Bahnson HT, Radulovic S, Santos AF, et al. Randomized trial of peanut consumption in infants at risk for peanut allergy. The New England journal of medicine 2015;372:803-13.
80. Asarnoj A, Glaumann S, Elfstrom L, Lilja G, Lidholm J, Nilsson C, et al. Anaphylaxis to peanut in a patient predominantly sensitized to Ara h 6. Int Arch Allergy Immunol 2012;159:209-12.
81. Vereda A, van Hage M, Ahlstedt S, Ibanez MD, Cuesta-Herranz J, van Odijk J, et al. Peanut allergy: Clinical and immunologic differences among patients from 3 different geographic regions. J Allergy Clin Immunol 2011;127:603-7.
82. Asarnoj A, Nilsson C, Lidholm J, Glaumann S, Ostblom E, Hedlin G, et al. Peanut component Ara h 8 sensitization and tolerance to peanut. J Allergy Clin Immunol 2012;130:468-72.
83. Valcour A, Jones JE, Lidholm J, Borres MP, Hamilton RG. Sensitization profiles to peanut allergens across the United States. Annals of allergy, asthma & immunology : official publication of the American College of Allergy, Asthma, & Immunology 2017;119:262-6 e1.
84. Sellaturay P, Nasser S, Ewan P. The incidence and features of systemic reactions to skin prick tests. Ann Allergy Asthma Immunol 2015;115:229-33.
85. Flores Kim J. Diagnostic accuracy, risk assessment, and cost-effectiveness of component-resolved diagnostics for food allergy: A systematic review. Allergy 2018;73:1609-21.
86. Brenner H. Measures of differential diagnostic value of diagnostic procedures. Journal of Clinical Epidemiology 1996;49:1435-9.
87. Klemans RJ, Blom WM, van Erp FC, Masthoff LJ, Rubingh CM, van der Ent CK, et al. Objective eliciting doses of peanut-allergic adults and children can be combined for risk assessment purposes. Clin Exp Allergy 2015;45:1237-44.
88. Lewis SA, Grimshaw KE, Warner JO, Hourihane JO. The promiscuity of immunoglobulin E binding to peanut allergens, as determined by Western blotting, correlates with the severity of clinical symptoms. Clinical and experimental allergy 2005;35:767-73.
89. Peeters KA, Koppelman SJ, van Hoffen E, van der Tas CW, den Hartog Jager CF, Penninks AH, et al. Does skin prick test reactivity to purified allergens correlate with clinical severity of peanut allergy? Clinical and experimental allergy 2007;37:108-15.
90. Wainstein BK, Studdert J, Ziegler M, Ziegler JB. Prediction of anaphylaxis during peanut food challenge: usefulness of the peanut skin prick test (SPT) and specific IgE level. Pediatr Allergy Immunol 2010;21:603-11.

Supplemental Methods for the Analysis of Health and Economic Benefits of Peanut Diagnostic Strategies:

Cost-effectiveness of peanut allergy diagnostic options was evaluated with decision analysis informed by results of the meta-analysis of diagnostic operating characteristics of single ara h 2 sIgE, whole peanut sIgE, and skin prick testing (SPT) (Figure 10). Markov modeling was used in microsimulations of each testing strategy (n=100,000 per strategy). Model assumptions are outlined in Table 10. Age-adjusted all-cause mortality was included over a 20-year time horizon (sensitivity range 5-80 years) with a start age during infancy sensitivity range to 8 years), a 14% pre-test probability of peanut allergy (sensitivity range 5%-90%) , and an assumption that 20% (sensitivity range of 5%-20%) of false positive diagnoses were refuted by accidental exposures over the model horizon in the base-case. Future costs and life-years were equally discounted at 3%, and risks of reactions, costs, and utilities of peanut allergy burden of illness were incorporated.

1418

1419

Table 1: The GRADE System of Recommendations and Evidence Certainty

Strength of Recommendation		
	For the Patient	For the Clinician
Strong	Most individuals in this situation would prefer the recommended course of action and only a small proportion would not.	The attending provider should strongly consider the recommended course of action as a first-line management. Formal decision aids may have less of a role to help individuals make decisions consistent with their values and preferences.
Conditional	The majority of individuals in this situation would prefer the suggested course of action, but many would not.	Different choices may be appropriate for different patients. Decision aids may be useful in helping individuals in making decisions consistent with their values and preferences. Clinicians should expect to spend more time with patients when working towards a decision.
Certainty in estimates of effect / quality rating both for outcome and for an entire evidence base as it pertains to a PICO		
High	There is high confidence that the true effect lies close to that of the estimate of the effect.	
Moderate	There is moderate confidence in the effect estimate. The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.	
Low	There is limited confidence in the effect estimate. The true effect may be substantially different from the estimate of the effect.	
Very Low	There is very little confidence in the effect estimate. The true effect is likely to be substantially different from the estimate of effect	

Table 2: Situations of Low to Moderate Pre-Test Probability for Peanut Allergy Where Testing May be a Preference-Sensitive Care Option to Offer in the Evaluation of a Patient[a]

Situations Where A Clinician Might Be Considering Testing for Peanut Allergy[b]	Pros for Testing	Cons for Testing
• A young child >1yr but <3 yr with multiple asthma hospitalizations, on chronic inhaled steroids, with known milk allergy who has not yet tried peanut	• Possible elevated risk for an additional food allergy in someone who already has one food allergy • Parents may not introduce peanut without a positive test, leading to additional risk from delayed introduction	• While the risk could be elevated over baseline, it is unclear if the absolute risk is elevated more than the low probability scenario of a 30% pre-test probability where a positive test was not shown to appreciably shift the post-test odds
• A young child >1yr but <3 yr old without eczema with prior anaphylaxis to one or more foods, but who has not yet tried peanut	• Possible elevated risk for an additional food allergy in someone who already has one food allergy • Parents may not introduce peanut without a positive test, leading to additional risk from delayed introduction	• While the risk could be elevated over baseline, it is unclear if the absolute risk is elevated more than the low probability scenario of a 30% pre-test probability where a positive test was not shown to appreciably shift the post-test odds
• A child in the first year of life with eczema suspected to be flared by one legume, and anaphylaxis to hummus who has not yet tried peanut	• Possible elevated risk for an additional food allergy in someone who already has one food allergy • Parents may not introduce peanut without a positive test, leading to additional risk from delayed introduction	• While the risk could be elevated over baseline, it is unclear if the absolute risk is elevated more than the low probability scenario of a 30% pre-test probability where a positive test was not shown to appreciably shift the post-test odds • By NIAID addendum criteria, the eczema does not make this child "high-risk"
• A 6 month old child with mild eczema tolerating a milk based formula, who has not tried egg or	• Parents may not introduce peanut without a positive test, based on the experience with	• While the risk could be elevated over baseline, it is unclear if the absolute risk is elevated more than the low probability scenario of a

peanut. Their older sibling has milk, egg, and peanut allergy	the older child, leading to additional risk from delayed introduction • Some clinicians ascribe to older literature that has suggested the younger sibling may be at some degree of increased risk of developing peanut allergy, though such literature did not account for the highly important factor of delayed introduction.	30% pre-test probability where a positive test was not shown to appreciably shift the post-test odds • By NIAID addendum criteria, the eczema does not make this child "high-risk" • Recent data has shown that testing the younger sibling is not cost effective until the prevalence of peanut allergy in siblings is shown to be >14% AND all such screened children also undergo an oral food challenge to provide a definitive outcome.

[a]See textbox 3 for explanation of what high, moderate, and low pre-test probability represent in the context of evaluating peanut allergy.

[b]These are hypothetical examples of situations that the workgroup members felt could represent potential scenarios that a clinician may evaluate under the context of a preference-sensitive care option. The choice of specific allergens, ages, and comorbidities are for illustration purposes only. Other allergens, ages, and comorbidities may represent possible presentations for consideration.

Table 3: Summary Statistics with 95% Confidence Intervals for SPT, sIgE, Ara h 2 Peanut Diagnostic Testing and Assessment of Reaction Severity

Diagnostic Test	Outcome	Sensitivity	Specificity	Positive Likelihood Ratio	Negative Likelihood Ratio
SPT 3mm	Diagnosis	0.97 (0.93-0.99)	0.46 (0.29-0.65)	1.82 (1.29-2.57)	0.05 (0.02-0.18)
sIgE 0.35 kU/L	Diagnosis	0.95 (0.91-0.97)	0.38 (0.28-0.48)	1.52 (1.3-1.77)	0.14 (0.08-0.24)
Ara h 2 sIgE 0.35 kU/L	Diagnosis	0.86 (0.81-0.89)	0.84 (0.79-0.89)	5.5 (3.99-7.56)	0.17 (0.13-0.23)
Ara h 2 sIgE 2 kU/L	Severe reaction	0.78 (0.69-0.85)	0.45 (0.28-0.63)	1.4 (1.08-1.83)	0.5 (0.37-0.66)
sIgE 50 kU/L	Severe reaction	0.39 (0.26-0.53)	0.89 (0.75-0.95)	3.4 (1.57-2.03)	0.69 (0.56-0.84)
SPT 10mm	Severe reaction	0.37 (0.22-0.55)	0.62 (0.44-0.77)	0.98 (0.71-1.35)	1 (0.84-1.22)
Sensitivity Analyses					
SPT 3mm[a]	SPT/sIgE Assessed in Same Study	0.98 (0.92-0.99)	0.5 (0.31-0.69)	1.94 (1.32-2.86)	0.04 (0.01-0.15)
sIgE 0.35 kU/L[a]	SPT/sIgE Assessed in Same Study	0.94 (0.9-0.97)	0.46 (0.32-0.6)	1.75 (1.35-2.26)	0.13 (0.07-0.21)
sIgE 0.35 kU/L[a]	sIgE/Ara h 2 Assessed in Same Study	0.95 (0.93-0.97)	0.3 (0.21-0.41)	1.36 (1.19-1.56)	0.47 (0.26-0.87)
Ara h 2 sIgE 0.35 kU/L[a]	sIgE/Ara h 2 Assessed in Same Study	0.85 (0.79-0.9)	0.86 (0.79-0.9)	5.87 (4.02-8.58)	0.18 (0.12-0.25)
SPT 3mm[a]	SPT/sIgE/Ara h 2 Assessed in Same Study	0.98 (0.89-1)	0.39 (0.22-0.6)	1.63 (1.19-2.23)	0.04 (0.01-0.25)
sIgE 0.35 kU/L[a]	SPT/sIgE/Ara h 2 Assessed in Same Study	0.95 (0.91-0.97)	0.4 (0.3-0.5)	1.58 (1.35-1.84)	0.12 (0.07-0.22)
Ara h 2 sIgE 0.35 kU/L[a]	SPT/sIgE/Ara h 2 Assessed in Same Study	0.83 (0.74-0.9)	0.79 (0.73-0.85)	4.03 (3.11-5.21)	0.21 (0.14-0.32)

[a]Test sensitivity and specificity are being reported for pooled studies for the particular individual test evaluated in the setting where multiple tests were run simultaneously in patients undergoing oral food challenge. Please refer to table 7 for reporting of additional sensitivity analyses.

Table 4a: GRADE Table of Evidence Certainty, Skin Prick Testing

Question: Should peanut skin prick testing at a threshold of 3mm wheal size be used to diagnose peanut allergy in patients with known or suspected peanut allergy?

Total number of studies/patients entered into the analysis: 18 studies, 2124 patients

Bibliography: Abrahms 2017, Begin 2017, Bernard2003, Chinthrajah 2018, Comberiati 2016, Dang 2012, DunnGalvin 2001, Johannsen 2016; Klemans Broekman 2013; Klemans Otte 2013, Leo 2015, Ludman 2013, Preece 2014, Rajput 2018, Rance 2003, Sampson 2017, Song 2015, Van Erp 2013.

Sensitivity	0.97 (95% CI: 0.93 to 0.90)
Specificity	0.46 (95% CI: 0.29 to 0.65)

Prevalences	2%	30%	70%

Outcome	№ of studies (№ of patients)	Study design	Factors that may decrease certainty of evidence					Effect per 1,000 patients tested (95% CI)			Test accuracy CoE
			Risk of bias	Indirectness	Inconsistency	Imprecision	Publication bias	pre-test probability of 2%	pre-test probability of 30%	pre-test probability of 70%	
True positives (patients with peanut allergy)	18 studies 961 patients	cross-sectional (cohort type accuracy study)	serious[a]	not serious[b]	not serious	not serious	none	19 (18 to 19)	291 (270 to 279)	679 (630 to 651)	⊕⊕⊕◯ MODERATE
False negatives (patients incorrectly classified as not having peanut allergy)								1 (1 to 2)	9 (21 to 30)	21 (49 to 70)	
True negatives (patients without peanut allergy)	18 studies 1163 patients	cross-sectional (cohort type accuracy study)	serious[a]	not serious[b]	not serious	not serious	none	451 (284 to 637)	322 (203 to 455)	138 (87 to 195)	⊕⊕⊕◯ MODERATE
False positives (patients incorrectly classified as having peanut allergy)								529 (343 to 696)	378 (245 to 497)	162 (105 to 213)	

Explanations

a. Multiple studies had potential for selection bias due to non-consecutive, non-randomized, or otherwise unexplained selective enrollment of the study population within the potentially eligible cohort. Multiple studies with issues relative to the flow/timing of when index diagnostic test performed relative to the reference oral food challenge

b. I^2 for sensitivity was 90.1% and for specificity was 93%

Table 4b: GRADE Table of Evidence Certainty, Serum IgE Testing

Question: Should peanut serologic IgE testing at a threshold of >0.35 KU_A/L be used to diagnose peanut allergy in patients with suspected peanut allergy?

Total number of studies/patients entered into the analysis: 30 studies, 3983 patients

Bibliography: Abrahms 2017, Balmer Weber 2015, Begin 2017,Beigelman 2012, Bernard 2003, Beyer 2015, Chinthrajah 2018, Comberiati 2016, Dang 2012, DunnGalvin 2001, Ebisawa 2012, Ebisawa 2015, Eller 2013, Johannsen 2016; Klemans Broekman 2013; Klemans Otte 2013, Leo 2015, Lieberman 2013, Ludman 2013, Martinet 2016, Nicolaou 2011, Preece 2014, Rajput 2018, Rance 2003, Sampson 2017, Song 2015, Van Erp 2013, Wainstein 2007

Sensitivity	0.95 (95% CI: 0.91 to 0.97)
Specificity	0.38 (95% CI: 0.28 to 0.48)

Prevalences	2%	30%	70%

Outcome	№ of studies (№ of patients)	Study design	Factors that may decrease certainty of evidence					Effect per 1,000 patients tested (95% CI)			Test accuracy CoE
			Risk of bias	Indirectness	Inconsistency	Imprecision	Publication bias	pre-test probability of 2%	pre-test probability of 30%	pre-test probability of 70%	
True positives (patients with peanut allergy)	30 studies 2046 patients	cross-sectional (cohort type accuracy study)	serious [a]	not serious[b]	not serious	not serious	none	19 (18 to 19)	285 (273 to 291)	665 (637 to 679)	⊕⊕⊕◯ MODERATE
False negatives (patients incorrectly classified as not having peanut allergy)								1 (1 to 2)	15 (9 to 27)	35 (21 to 63)	
True negatives (patients without peanut allergy)	30 studies 1937 patients	cross-sectional (cohort type accuracy study)	serious [a]	not serious[b]	not serious	not serious	none	372 (274 to 470)	266 (196 to 336)	114 (84 to 144)	⊕⊕⊕◯ MODERATE
False positives (patients incorrectly classified as having peanut allergy)								608 (510 to 706)	434 (364 to 504)	186 (156 to 216)	

Explanations

a. Multiple studies had potential for selection bias due to non-consecutive, non-randomized, or otherwise unexplained selective enrollment of the study population within the potentially eligible cohort. Multiple studies with issues relative to the flow/timing of when index diagnostic test performed relative to the reference oral food challenge

b. I^2 for sensitivity was 95.9% and for specificity was 92.8%

Table 5: GRADE Table of Evidence Certainty, Ara h 2 sIgE Testing

Question: Should Ara h 2 specific IgE at a threshold of >0.35 KU_A/L be used to diagnose peanut allergy in patients with suspected peanut allergy?

Total number of studies/patients entered into the analysis: 24 studies, 2289 patients

Bibliography: Balmer Weber 2015, Bernard 2003, Beyer 2015, Chinthrajah 2018, Comberiati 2016, Dang 2012, Ebisawa 2012, Ebisawa 2015, Eller 2013, Glaumann 2012, Keet 2013,; Klemans Broekman 2013; Klemans Otte 2013, Kukkonen 2015, Leo 2015, Lieberman 2013, Martinet 2016, Nicolaou 2011, Preece 2014, Rajput 2018, Rance 2003, Schots 2016, Suratannon 2013 Van Erp 2013

Sensitivity	0.86 (95% CI: 0.81 to 0.89)
Specificity	0.84 (95% CI: 0.79 to 0.89)

Prevalences	2%	30%	70%

Outcome	№ of studies (№ of patients)	Study design	Factors that may decrease certainty of evidence					Effect per 1,000 patients tested (95% CI)			Test accuracy CoE
			Risk of bias	Indirectness	Inconsistency	Imprecision	Publication bias	pre-test probability of 2%	pre-test probability of 30%	pre-test probability of 70%	
True positives (patients with peanut allergy)	24 studies 1336 patients	cross-sectional (cohort type accuracy study)	serious [a]	not serious[b]	not serious	not serious	none	17 (16 to 18)	258 (243 to 267)	602 (567 to 623)	⊕⊕⊕◯ MODERATE
False negatives (patients incorrectly classified as not having peanut allergy)								3 (2 to 4)	42 (33 to 57)	98 (77 to 133)	
True negatives (patients without peanut allergy)	24 studies 953 patients	cross-sectional (cohort type accuracy study)	serious [a]	not serious[b]	not serious	not serious	none	823 (774 to 872)	588 (553 to 623)	252 (237 to 267)	⊕⊕⊕◯ MODERATE
False positives (patients incorrectly classified as having peanut allergy)								157 (108 to 206)	112 (77 to 147)	48 (33 to 63)	

Explanations

a. Multiple studies had potential for selection bias due to non-consecutive, non-randomized, or otherwise unexplained selective enrollment of the study population within the potentially eligible cohort. Multiple studies with issues relative to the flow/timing of when index diagnostic test performed relative to the reference oral food challenge

b. I^2 for sensitivity was 81.4 and specificity was 69.7

Table 6a: GRADE Table of Evidence Certainty, Ara h 2 sIgE to Assess Reaction Severity

Question: Should Ara h 2 specific IgE at a threshold of >2 KU_A/Lbe used to diagnose severe peanut allergy in patients with suspected peanut allergy?

Total number of studies/patients entered into the analysis: 10 studies, 845 patients

Bibliography: Balmer Weber 2015, Chinthrajah 2018, Dang 2012, Glaumann 2012, Klemans Broekman 2013; Kukkonen 2015, Leo 2015, Preece 2014, Rajput 2018, Van Erp 2013

Sensitivity	0.78 (95% CI: 0.69 to 0.85)
Specificity	0.45 (95% CI: 0.28 to 0.63)

Prevalences	2%	30%	70%

Outcome	№ of studies (№ of patients)	Study design	Factors that may decrease certainty of evidence					Effect per 1,000 patients tested (95% CI)			Test accuracy CoE
			Risk of bias	Indirectness	Inconsistency	Imprecision	Publication bias	pre-test probability of 2%	pre-test probability of 30%	pre-test probability of 70%	
True positives (patients with severe peanut allergy)	10 studies 308 patients	cross-sectional (cohort type accuracy study)	serious [a]	not serious[b]	very serious [c,d]	not serious	none	16 (14 to 17)	234 (207 to 255)	546 (483 to 595)	⊕◯◯◯ VERY LOW
False negatives (patients incorrectly classified as not having severe peanut allergy)								4 (3 to 6)	66 (45 to 93)	154 (105 to 217)	
True negatives (patients without severe peanut allergy)	10 studies 380 patients	cross-sectional (cohort type accuracy study)	serious [a]	not serious [b]	very serious [c,d]	not serious	none	441 (274 to 617)	315 (196 to 441)	135 (84 to 189)	⊕◯◯◯ VERY LOW
False positives (patients incorrectly classified as having severe peanut allergy)								539 (363 to 706)	385 (259 to 504)	165 (111 to 216)	

Explanations

a. Multiple studies had potential for selection bias due to non-consecutive, non-randomized, or otherwise unexplained selective enrollment of the study population within the potentially eligible cohort. Multiple studies with issues relative to the flow/timing of when index diagnostic test performed relative to the reference oral food challenge

b. I^2 for sensitivity was 68.7% and for specificity was 91.6%

c. The heterogeneity for the estimate was very high

d. The criteria to assess severity was not uniform among all studies included

Table 6b: GRADE Table of Evidence Certainty, Peanut sIgE to Assess Reaction Severity

Question: Should peanut serologic IgE testing at a threshold of >50 KU_A/Lbe used to diagnose severe peanut allergy in patients with suspected peanut allergy?

Total number of studies/patients entered into the analysis: 13 studies, 1051 patients

Bibliography: Chinthrajah 2018, Dang 2012, DunnGalvin 2001, Glaumann 2012, Klemans Broekman 2013; Lewis 2005, Peeters 2007, Preece 2014, Rajput 2018, Song 2015, Van Erp 2013, Wainstein 2007, Wensing 2002

Sensitivity	0.39 (95% CI: 0.26 to 0.53)
Specificity	0.89 (95% CI: 0.75 to 0.95)

Prevalences	2%	30%	70%

Outcome	№ of studies (№ of patients)	Study design	Factors that may decrease certainty of evidence					Effect per 1,000 patients tested (95% CI)			Test accuracy CoE
			Risk of bias	Indirectness	Inconsistency	Imprecision	Publication bias	pre-test probability of 2%	pre-test probability of 30%	pre-test probability of 70%	
True positives (patients with severe peanut allergy)	13 studies 256 patients	cross-sectional (cohort type accuracy study)	serious [a]	not serious[b]	very serious [c,d]	not serious	none	8 (5 to 11)	117 (78 to 159)	273 (182 to 371)	⊕◯◯◯ VERY LOW
False negatives (patients incorrectly classified as not having severe peanut allergy)								12 (9 to 15)	183 (141 to 222)	427 (329 to 518)	
True negatives (patients without severe peanut allergy)	13 studies 795 patients	cross-sectional (cohort type accuracy study)	serious [a]	not serious[b]	very serious [c,d]	not serious	none	872 (735 to 931)	623 (525 to 665)	267 (225 to 285)	⊕◯◯◯ VERY LOW
False positives (patients incorrectly classified as having severe peanut allergy)								108’ (49 to 245)	77 (35 to 175)	33 (15 to 75)	

Explanations

a. Multiple studies had potential for selection bias due to non-consecutive, non-randomized, or otherwise unexplained selective enrollment of the study population within the potentially eligible cohort. Multiple studies with issues relative to the flow/timing of when index diagnostic test performed relative to the reference oral food challenge

b. I^2 for sensitivity was 75.7% and for specificity was 90.9%

cThe criteria to assess severity was not uniform among all studies

d. The heterogeneity for the estimate was very high

Table 6c: GRADE Table of Evidence Certainty, Peanut sIgE to Assess Reaction Severity

Question: Should peanut skin prick testing at a threshold of 10mm wheal size be used to diagnose severe peanut allergy in patients with suspected peanut allergy?

Total number of studies/patients entered into the analysis: 12 studies, 737 patients

Bibliography: Chinthrajah 2018, Dang 2012, DunnGalvin 2001, Klemans Broekman 2013; Leo 2015, Lewis 2005; Preece 2014, Rajput 2018, Song 2015, Van Erp 2013, Wainstein 2010, Wensing 2002

Sensitivity	0.37 (95% CI: 0.22 to 0.55)
Specificity	0.62 (95% CI: 0.44 to 0.77)

Prevalences	2%	30%	70%

Outcome	№ of studies (№ of patients)	Study design	Factors that may decrease certainty of evidence					Effect per 1,000 patients tested (95% CI)			Test accuracy CoE
			Risk of bias	Indirectness	Inconsistency	Imprecision	Publication bias	pre-test probability of 2%	pre-test probability of 30%	pre-test probability of 70%	
True positives (patients with severe peanut allergy)	12 studies 166 patients	cross-sectional (cohort type accuracy study)	serious [a]	not serious[b]	very serious [c,d]	not serious	none	7 (4 to 11)	111 (66 to 165)	259 (154 to 385)	⊕◯◯◯ VERY LOW
False negatives (patients incorrectly classified as not having severe peanut allergy)								13 (9 to 16)	189 (135 to 234)	441 (315 to 546)	
True negatives (patients without severe peanut allergy)	12 studies 571 patients	cross-sectional (cohort type accuracy study)	serious [a]	not serious[b]	very serious [c,d]	not serious	none	608 (431 to 755)	434 (308 to 539)	186 (132 to 231)	⊕◯◯◯ VERY LOW
False positives (patients incorrectly classified as having severe peanut allergy)								372 (225 to 549)	266 (161 to 392)	114 (69 to 168)	

Explanations

a. Multiple studies had potential for selection bias due to non-consecutive, non-randomized, or otherwise unexplained selective enrollment of the study population within the potentially eligible cohort. Multiple studies with issues relative to the flow/timing of when index diagnostic test performed relative to the reference oral food challenge

b. I^2 for sensitivity was 64% nd for specificity was 87.9%

c. The criteria to assess severity was not uniform among all studies included

d. The heterogeneity for the estimate was very high

Table 7: Additional Sensitivity Analyses

Test	Outcome	Analyses	Sensitivity	Specificity	Positive LR	Negative LR
SPT 3mm	Diagnosis	Exclusion of studies with high risk of bias	0.96	0.48	1.85	0.08
		Pediatric studies only	0.97	0.52	2.02	0.06
		Open OFC studies only	0.96	0.53	2.04	0.08
		DBPCFC studies only	0.99	0.38	1.60	0.03
		European studies only	0.98	0.56	2.23	0.04
		Non-European studies only	0.97	0.32	1.43	0.09
sIgE >0.35	Diagnosis	Exclusion of studies with high risk of bias	0.96	0.44	1.71	0.09
		Pediatric studies only	0.94	0.41	1.59	0.15
		Open OFC studies only	0.94	0.4	1.57	0.15
		DBPCFC studies only	0.97	0.42	1.67	0.07
		European studies only	0.95	0.38	1.53	0.13
		Non-European studies only	0.95	0.37	1.51	0.14
Ara h 2 sIgE >0.35	Diagnosis	Exclusion of studies with high risk of bias	0.86	0.81	4.53	0.17
		Pediatric studies only	0.85	0.85	5.67	0.18
		Open OFC studies only	0.85	0.85	5.67	0.18
		DBPCFC studies only	0.87	0.83	5.12	0.16
		European studies only	0.88	0.85	5.87	0.14
		Non-European studies only	0.83	0.84	5.19	0.20
Ara h 2 sIgE >2	Severity	Exclusion of studies with high risk of bias	0.75	0.42	1.29	0.60
		Pediatric studies only	0.72	0.49	1.41	0.57
		Open OFC only	0.64	0.43	1.12	0.84
		DBPCFC only	0.8	0.44	1.43	0.45
		European studies only	0.77	0.43	1.35	0.53
		Non-European studies only	0.71	0.44	1.27	0.66
sIgE >50	Severity	Exclusion of studies with high risk of bias	0.36	0.88	3.00	0.73
		Pediatric studies only	0.38	0.92	4.75	0.67
		Open OFC only	0.29	0.97	9.67	0.73
		DBPCFC studies only#	0.47	0.71	1.62	0.75
		European studies only	0.38	0.86	2.71	0.72
		Non-European studies only	0.44	0.92	5.50	0.61
SPT 10mm	Severity	Exclusion of studies with high risk of bias	0.41	0.57	0.95	1.04
		Pediatric studies only	0.29	0.71	1.00	1.00
		Open OFC studies only	0.26	0.69	0.84	1.07
		DBPCFC studies only#	0.62	0.41	1.05	0.93
		European studies only	0.39	0.67	1.18	0.91
		Non-European studies only	0.36	0.59	0.88	1.08

Table 8: Risk of Bias Assessment

Study	Year	Bias				Applicability		
		Patient Selection	Index Test	Reference Standard	Flow and Timing	Patient Selection	Index Test	Reference Standard
Abrahms	2017							
Balmer Weber	2015							
Begin	2017							
Beigelman	2012							
Bernard	2003							
Beyer	2015							
Chinthrajah	2018							
Comberiati	2016							
Dang	2012							
DunnGalvin	2011							
Ebisawa	2015							
Ebisawa	2012							
Eller	2013							
Glaumann	2012							
Guilloux	2009							
Gupta	2014							
Johannsen	2016							
Keet	2013							
Klemans Blom	2015							
Klemans Broekman	2013							
Klemans Liu	2013							
Klemans Otte	2013							
Kukkonen	2015							
Leo	2015							
Lewis	2005							
Lieberman	2013							
Ludman	2013							
Martinet	2016							
Nicolaou	2011							
Peeters	2007							
Perry	2004							
Preece	2014							
Rajput	2018							
Rance	2003							
Schots	2016							
Song	2015							
Suratannon	2013							
Van Erp	2013							
Wainstein	2007							
Wainstein	2010							
Wensing	2002							

Red: high risk Yellow: Unclear risk Green: low risk

Table 9: Simulation Model Inputs

Variable	Model Reference (sensitivity range)	Source
US Life Table	National Vital Statistics Reports, April 2017	Arias E, Heron M, Xu J. United States Life Tables, 2013. National Vital Statistics Reports 2017; 66(3): 1-64.
Testing characteristics	Skin prick testing: Sn: 0.97 (0.86-0.98); Sp 0.46 (0.17-0.67) Ara h 2: Sn 0.86 (0.72 – 0.90); Sp 0.84 (0.65-0.87) Whole peanut sIgE: 0.95 (0.89-0.97); Sp 0.38 (0.23-0.49)	Meta-analysis
Food allergy fatality	5-19 years: 3.25 per million person years (0.3 – 30) 20 years and older: 1.81 per million person years (1.81-18.1)	Umasunthar T. Leonardi-Bee, Hodes M, et al. Incidence of fatal food anaphylaxis in people with food allergy: a systematic review and meta-analysis. Clinical & Experimental Allergy. 2013; 43: 1333-1341.
Rate of accidental peanut exposure and symptoms in peanut allergic persons	7% per year (5%-45%)	Neuman-Sunshine D, Eckman J, Keet C, Matsui E, Peng R, et al. The natural history of peanut allergy. Ann Allergy Asthma Immunol. 2012; 108: 326-331.
Rate of emergency room visit for severe symptoms in peanut allergic persons	1% per year (0.5%-35%)	Neuman-Sunshine D, Eckman J, Keet C, Matsui E, Peng R, et al. The natural history of peanut allergy. Ann Allergy Asthma Immunol. 2012; 108: 326-331.
Hospitalization following emergency room visit for anaphylaxis	35% (5%-45%)	Robinson M, Greenhawt M, Stukus D. Factors associted with epinephrine administration for anaphylaxis in children before arrvial to the emergency department. Ann Allergy Asthma Immunol. 2017; 119: 164-169.
Primary care visits (mean incremental annual cost for food allergy diagnosis)	$102 ($94-$105)	Gupta R, Holdford D, Bilaver L, Dyer A, Holl JL, Meltzer D. The economic impact of childhood food allergy in the United States. JAMA Pediatr 2013;167:1026-31; US Department of Labor, Bureau of Labor Statistics. Available from www.bls.gov. =
Allergist visits for food allergy (mean incremental annual cost for food allergy diagnosis)	$151 ($140 - $152)	Gupta R, Holdford D, Bilaver L, Dyer A, Holl JL, Meltzer D. The economic impact of childhood food allergy in the United States. JAMA Pediatr 2013;167:1026-31; US Department of Labor, Bureau of Labor Statistics. Available from www.bls.gov.
Nutritionist visits for food allergy (per year)	$17 ($15 - $18)	Gupta R, Holdford D, Bilaver L, Dyer A, Holl JL, Meltzer D. The economic impact of childhood food allergy in the United States. JAMA Pediatr 2013;167:1026-31; US Department of Labor, Bureau of Labor Statistics. Available from www.bls.gov.

Alternative provider visits for food allergy (per year)	$25 ($22 - $27)	Gupta R, Holdford D, Bilaver L, Dyer A, Holl JL, Meltzer D. The economic impact of childhood food allergy in the United States. JAMA Pediatr 2013;167:1026-31; US Department of Labor, Bureau of Labor Statistics. Available from www.bls.gov.
Incremental annual grocery costs (living with food allergy)	$315 ($290-330)	Gupta R, Holdford D, Bilaver L, Dyer A, Holl JL, Meltzer D. The economic impact of childhood food allergy in the United States. JAMA Pediatr 2013;167:1026-31; US Department of Labor, Bureau of Labor Statistics. Available from www.bls.gov.
Job-related opportunity costs from food allergy (per year)	$2,637 ($0 - $2,697)	Gupta R, Holdford D, Bilaver L, Dyer A, Holl JL, Meltzer D. The economic impact of childhood food allergy in the United States. JAMA Pediatr 2013;167:1026-31; US Department of Labor, Bureau of Labor Statistics. Available from www.bls.gov.
Personal epinephrine auto-injector	$726 ($100-$800)	Shaker M, Bean K, Verdi M. Economic evaluation of epinephrine autoinjectors for peanut allergy. Ann Allergy Asthma Immunol. 2017; 119(2): 160-163. US Department of Labor, Bureau of Labor Statistics. CPI Inflation calculator. Accessed at https://data.bls.gov on 9/2/18.
sIgE / ara h 2 IgE testing	$17 per test ($10-$117)	Healthcare Bluebook. www.healthcarebluebook.com. Accessed 11/22/18
Skin test cost	$24 ($10-$40)	Physician Fee schedule. Available from http://www.cms.gov/. Accessed 10/3/17.
Hospitalization	$5,991 ($5,732-$6,066)	Patel DA, Holdford DA, Edwards E, Carroll NV. Estimating the economic burden of food-induced allergic reactions and anaphylaxis in the United States. J Allergy Clin Immunol 2011;128:110-5 e5 US Department of Labor, Bureau of Labor Statistics. CPI Inflation calculator. Accessed at https://data.bls.gov on 9/2/18.
ED visit	$702 ($689-$710)	Patel DA, Holdford DA, Edwards E, Carroll NV. Estimating the economic burden of food-induced allergic reactions and anaphylaxis in the United States. J Allergy Clin Immunol 2011;128:110-5 e5 US Department of Labor, Bureau of Labor Statistics. CPI Inflation calculator. Accessed at https://data.bls.gov on 9/2/18.

Supervised oral food challenge	$123 ($110-600)	Doctors office visits billing and charges. Dartmouth-Hitchcock. http://www.dartmouth-hitchcock.org. Accessed March 10, 2017 US Department of Labor, Bureau of Labor Statistics. CPI Inflation calculator. Accessed at https://data.bls.gov on 4/21/19
Start age	0 years (0 years to 8 years)	
Negative health state influence for food allergy and food anaphylaxis	-0.09 (-0.02 - -0.11)	Carroll AE, Downs SM. Improving decision analyses: parent preferences (utility values) for pediatric health outcomes. J Pediatr 2009;155:21-5, 5 e1-5.
Cycle length	1 year	
Time Horizon	20 years (5 years – 80 years)	
Peanut allergy pre-test probability	14% (3% - 90%)	
Annual discount rate	0.03 (0-0.03)	
Probability of spontaneous tolerance	22% (0-22%)	
Probability of identifying false positive test over model horizon	20% (5%-20% over 5-20 years)	

Table 10: Cost-Effectiveness Comparisons of Use of Peanut SPT, sIgE, and Ara h 2 sIgE Testing

	Cost	QALY	Net Monetary Benefit	Total Rxn	Anaphylaxis	Anaphylaxis Fatality
3% peanut allergy pre-test probability						
Skin Prick Test						
Mean	$20,734.48	14.43	-$20,734.48	0.0341	0.0047	0.0000
Std Deviation	$23,902.82	1.36	$23,902.82	0.2833	0.0745	0.0000
Ara h 2						
Mean	$7,669.24	14.79	-$7,669.24	0.0379	0.0049	0.0000
Std Deviation	$17,355.08	1.33	$17,355.08	0.3008	0.0783	0.0000
Whole peanut sIgE						
Mean	$23,466.54	14.35	-$23,466.54	0.0345	0.0048	0.0000
Std Deviation	$24,165.42	1.35	$24,165.42	0.2852	0.0756	0.0000
14% peanut allergy pre-test probability						
Skin Prick Test						
Mean	$23,859.49	14.36	-$23,859.49	0.1555	0.0213	0.0000
Std Deviation	$25,361.09	1.33	$25,361.09	0.5784	0.1574	0.0000
Ara h 2						
Mean	$12,329.23	14.69	-$12,329.23	0.1725	0.0223	0.0000
Std Deviation	$22,237.68	1.32	$22,237.68	0.6169	0.1614	0.0000
Whole peanut sIgE						
Mean	$26,289.04	14.29	-$26,289.04	0.1581	0.0212	0.0000
Std Deviation	$25,304.83	1.32	$25,304.83	0.5836	0.1574	0.0000
75% peanut allergy pre-test probability						
Skin Prick Test						
Mean	$41,680.67	13.99	-$41,680.67	0.8479	0.1200	0.0000
Std Deviation	$25,973.46	1.24	$25,973.46	1.1182	0.3571	0.0000
Ara h 2						
Mean	$38,191.62	14.09	-$38,191.62	0.9273	0.1206	0.0000
Std Deviation	$27,947.58	1.28	$27,947.58	1.1690	0.3588	0.0000

Whole peanut sIgE						
Mean	$42,378.21	13.97	-$42,378.21	0.8632	0.1205	0.0000
Std Deviation	$25,494.62	1.24	$25,494.62	1.1286	0.3579	0.0000

Table 11: Summary Recommendations in Evaluating the Patient with Suspected Peanut Allergy

Question	Recommendation	Evidence Certainty	Risk of Bias
Should diagnostic testing for peanut allergy be performed in adults and children with a history of suspected peanut allergy who are requesting evaluation for peanut allergy?	We suggest in favor of diagnostic (skin prick or serum sIgE) testing for peanut allergy in patients with a 1) physician-judged high pre-test probability of peanut allergy, or 2) prior to an oral food challenge for patients with moderate pre-test probability of peanut allergy, with whom shared decision-making has been employed to arrive at the final decision. We suggest against diagnostic testing in patients where there is low or very low pre-test probability of peanut allergy.	Very Low	Not Rated
In the patient presenting for evaluation of suspected peanut allergy, which of the three tests—SPT, sIgE to whole peanut, or Ara h2 would provide the highest diagnostic accuracy as determined by the more optimal positive/negative likelihood ratio?	We suggest in favor of Ara h2 diagnostic testing in a patient presenting for evaluation of suspected peanut allergy for which a single diagnostic test is to be used, as Ara h2 would provide the best diagnostic accuracy as determined by virtue of more optimal positive/negative likelihood ratios.	Moderate	High
In a patient presenting for evaluation of suspected peanut allergy, does testing for peanut components in addition to either SPT or sIgE to whole peanut increase the diagnostic accuracy?	We suggest against component testing in addition to either to skin prick test or sIgE to whole peanut to increase diagnostic accuracy.	Very Low	High
In the patient presenting for evaluation of suspected peanut allergy, can the results of a diagnostic test be used to predict the severity of a future allergic reaction?	We suggest against the clinician using the results of a SPT, sIgE to whole peanut extract, or sIgE to peanut components to determine the severity of a previous reaction and/or allergy phenotype or to predict the severity of a future reaction.	Very Low	High

Text Box 1: GRADE Questions Evaluated in this Practice Parameter

PICO Questions: GRADE Analysis of Diagnostic Testing in the Diagnosis of Peanut Allergy

1. Should diagnostic testing for peanut allergy be performed in adults and children with a history of suspected peanut allergy who are requesting evaluation for peanut allergy?

2a. In the patient presenting for evaluation of suspected peanut allergy, which of the three tests—SPT, sIgE to whole peanut, or Ara h2 would provide the highest diagnostic accuracy as determined by the more optimal positive/negative likelihood ratio?

2b. In a patient presenting for evaluation of suspected peanut allergy, does testing for peanut components in addition to either SPT or sIgE to whole peanut increase the diagnostic accuracy?

3. In the patient presenting for evaluation of suspected peanut allergy, can the results of a diagnostic test be used to predict the severity of a future allergic reaction?

Text Box 2: Defining Allergic Sensitization and a "Positive Test"

Allergic sensitization is denoted by the presence of detectable allergen-specific IgE, either through a serologic assay or through skin prick testing. All tests for sensitization have a threshold where the test is considered to be positive, as well as either a detection limit or a reporting limit. For skin prick testing, the most commonly reported convention for where a test is considered "positive" for the presence of allergen specific IgE is when the allergen-specific test is 3mm of wheal diameter greater than that of a simultaneously placed glycerinated saline control. As discussed in the 2008 Diagnostic Testing Practice Parameter (www.allergyparameters.org) different testing devices produce some degree of variation in the size range of negative controls, as does variation related to the tester. Wheal size is recommended to be measured as the average length of the two longest bisecting planes, though many clinics may elect to measure the longest single plane.

For serum-specific IgE tests using fluorescent enzymatic immunoassay (FEIA) detection, the instruments have both detection limits and reporting limits that have influenced test results. However, each instrument has particular reporting and detection ranges, and these differ between commercial tests. The technical detection limit for these machines is typically 0.1 KU_A/L, and antibody levels above this threshold are reported as they are detected, to an upper reporting limit of 100 KU_A/L. Quantification of levels >100 KU_A/L is possible through sample dilution. For many years, the reporting limit was conventionally set at < 0.35 KU_A/L, though in recent years, this has been replaced by the detection limit of 0.1 KU_A/L. Using the older convention of the 0.35 KU_A/L reporting limit, "positive" sensitization was considered to be 0.36 KU_A/L or higher. With the newer convention of using the 0.1 KU_A/L detection limit as the reporting limit, "positive" sensitization would therefore be 0.11 KU_A/L. This creates a conundrum of how to interpret sensitization between 0.11 KU_A/L and 0.35 KU_A/L, which prior to the change in reporting convention would have fallen into the "negative" range. It is debatable that such sensitization is clinically relevant, or that many clinicians would only consider sensitization above 0.7 KU_A/L as clinically relevant. Nonetheless, studies may report positive sensitization at 0.11 KU_A/L in a binary fashion. One additional classification that is seen are classes representing sextiles of IgE quantity detected between the upper and lower reporting limits. These are arbitrary conventions that date back to the quartiles originally described for Radioallosorbent Testing, adjusted for the FEIA method. Class 0 represents levels below the reporting limit, and class 1 typically starts at the reporting lower limit, ranging to class 6 representing the highest levels detectable which are reported. These class designations have no clinical relevance in and of themselves, and no reference to class designations is made in this document.

In this document, if the term positive is used, in relation to either form of test it is in this sense that this refers to positive detection of sensitization (e.g. a positive test). Unequivocally, positive detection of sensitization is unrelated to a positive clinical diagnosis of allergy. A positive diagnosis is predicated on both a demonstrated clinical history of allergy and the presence of detectable sensitization, or in very circumscribed instances, very high levels of sensitization in infants with very particular pre-existing risk factors who have never ingested peanut previously.

Text Box 3: Examples of Pre-test Probability in Determining if Diagnostic Testing is Indicted

1. ***High pre-test probability*** should be considered as a situation where there was ingestion of peanut and typical IgE mediated symptoms of an allergic reaction resulted, either directly observed or reported; or an infant meeting NIAID early peanut introduction high-risk criteria prior to peanut introduction. Testing is of the highest utility in these scenarios and peanut sensitization above a certain threshold is of high likelihood to be associated with the highest post-test odds of a diagnosis of peanut allergy.
2. ***Moderate pre-test probability*** should be considered as a situation where there is less clarity that peanut was ingested and resulted in IgE mediated symptoms, but some consideration for this in explaining an allergic reaction under evaluation. In some instances it may represent situations where the patient has not previously consumed peanut but could be considered at a risk greater than the general population for peanut allergy based on the presence of certain types of other food allergies, certain atopic comorbidities (e.g., severe eczema), or certain children outside the first year of life with delayed peanut introduction. Testing is of unclear utility in these situations, and not necessarily associated with post-test odds that clarify clinical decision making. An oral food challenge may be required to definitively establish a diagnosis when there is peanut sensitization above a certain threshold.
3. ***Low pre-test probability*** should be considered a situation where there is very little uncertainty that the person is peanut tolerant (e.g. eats peanut without becoming symptomatic), that peanut was unrelated to the allergic reaction being evaluated (e.g. it is clear that a single allergen other than peanut likely caused the aforementioned reaction and the product was peanut-free, or peanut is being tested solely because it is part of a multi-allergen panel and there is no specific independent concern for peanut allergy itself), family history of peanut allergy or allergic disease, general curiosity about what someone could speculatively be "allergic to", or for an infant meeting addendum 2 or 3 criteria for NIAID early peanut introduction guidelines prior to peanut introduction. In some instances it may represent situations where the patient has not previously consumed peanut but the clinician may have concern that the patient is at a risk greater than the general population for peanut allergy based on the presence of certain types of other food allergies or concern for cross-reactivity, certain atopic comorbidities (e.g., mild or moderate eczema), or certain children outside the first year of life with delayed peanut introduction but who have no baseline risk factors. Testing in these situations is of exceptionally limited to no utility whatsoever, is not associated with any shift of post-test odds over baseline, and is not indicated. An oral food challenge is likely required to establish that the peanut sensitization detected is clinically irrelevant.

Text Box 4: Key Questions in Peanut Allergy Diagnostic Testing

- **Are there any clinical indications to obtain peanut allergy testing for a patient who is eating peanut without immediate onset or reproducible symptoms?**

  In general, no. However, rare exceptions to this include part of the evaluation of patients with eosinophilic esophagitis where dietary elimination is considered as a treatment option, which is a highly specific context with very particular (non-IgE mediated) symptoms, which is beyond the scope of this document. (Section xx, page xx)

- **Which test should be ordered in the evaluation of patients who have never ingested peanut, i.e. prior to early introduction for at risk infants?**

  Peanut skin prick and serum IgE testing is poorly specific and in general should not be used as a screening tool for someone who has never eaten peanut before and developed symptoms. When used as part of the early introduction guidelines for infants less than 6 months of age who have severe eczema and/or egg allergy, both skin prick and serum peanut IgE tests can be utilized. There is no current role for component testing in this context. (Section xx, page xx)

- **Are there cut-off levels for peanut skin prick or serum IgE testing that diagnoses peanut allergy?**

  A universal cut-off level does not exist. These are technically difficult to generate, given that these are based on accurately knowing the population prevalence of peanut allergy. Cut-off levels are only relative probabilities that are imperfect and have an error rate that will potentially misclassify individuals. When prevalence of disease is not known, the likelihood ratio is a more applicable test. This tells the likelihood of a positive test in someone with the disease compared to the likelihood of a positive test in someone without disease, and can help convert the pre-test probability that someone has the disease to a post-test odds using a Fagan nomogram. Thus, as stand-alone measures, neither skin prick nor serum IgE test results can be interpreted as diagnostic for peanut allergy. (Section xx, page xx)

- **Should peanut allergy testing be considered in children with moderate-to-severe atopic dermatitis?**

  Atopic dermatitis is caused by changes in the epidermal skin barrier and is generally not due to food allergy, though children with persistent and refractory moderate-to-severe atopic dermatitis may be at higher risk of developing food allergy. Peanut allergy testing should not be a standard part of the evaluation for any patient with atopic dermatitis. However, in a very small subset of infants and young children with severe, treatment- refractory atopic dermatitis may benefit from select food allergy, including peanut allergy testing if the clinical history suggests peanut has not yet been introduced, or there is suspicion that peanut ingestion is temporally associated with flares. (Section xx, page xx)

- **Should children with a family history of peanut allergy in another sibling be evaluated for peanut allergy prior to this being introduced?**

  Screening of younger siblings for peanut allergy should not be routinely performed, and there is no evidence that such individuals are at higher risk for developing peanut allergy based just on the sibling history alone. To facilitate timely introduction and prevent delay, there could be consideration for a role for testing when parents are overly anxious about introducing peanut and will not introduce peanut to their child through any other means. However, such testing must be interpreted properly and a positive result not be considered diagnostic for

peanut allergy. In these situations, either skin prick or serum IgE testing may be utilized. Data exists to show that this practice is not cost-effective until there is a much higher baseline prevalence of peanut allergy in the population, and then only cost-effective if sensitized children undergo challenge rather than avoid peanut based on strong sensitization. There is no indication to utilize component testing in this context. (Section xx, page xx)

- **Are all patients with detectable Ara h 2 clinically allergic to peanuts?**

  No. Detectable isolated sensitization to Ara h 2 is not diagnostic for peanut allergy, and a diagnosis can only be made where the individual is sensitized in the context of a known or suspected reaction after eating peanut. There are no well-established cut off levels for Ara h 2 at this time that indicate the presence of allergy versus sensitization. However, when compared to whole peanut skin prick and sIgE tests, Ara h 2 testing has vastly increased specificity, though this is still largely dependent on the context in which any testing is indicated. Patients may have detectable Ara h 2 but exhibit no clinical reactivity upon ingestion of peanut. (Section xx, page xx)

- **Does component testing predict the severity of future reactions?**

  No test, including components, has good sensitivity or specificity to indicate the severity of a future reaction. Component testing may have a potential role to help identify sensitization patterns that indicate recognition of cross sensitization with pollen allergens as opposed to more primary allergens unique to peanut, though the clinical significance of this is still to be defined. (Section xx, page xx)

- **When should component testing be ordered as the initial diagnostic test?**

  The role of component testing is evolving, and it is unclear how and when these tests should be used. Comparatively, testing for Ara h 2 compared to whole peanut skin prick and sIgE testing does have significantly higher specificity, which may translate to a lower likelihood of a false positive diagnosis if testing is run the right context. Moreover, in this context, use of Ara h 2 as a stand-alone test is highly cost-effective. However, there is a present knowledge gap if Ara h 2 should be the initial test ordered. (Section xx, page xx)

QUESTION

In patients presenting for evaluation of suspected peanut allergy, which of the three tests—Skin prick test, sIgE to whole peanut, or Ara h2 would provide the most diagnostic accuracy as determined by the more optimal positive/negative likelihood ratio?	
POPULATION:	Adults and children presenting for evaluation of peanut allergy
INTERVENTION:	Using peanut skin prick testing (SPT), serum specific IgE to whole peanut (sIgE), or Ara h 2 serum specific IgE (Ara h 2 sIgE) to determine peanut sensitization to assist in the diagnosis peanut allergy
COMPARATOR	Oral Food challenge
OUTCOMES	Diagnostic accuracy of peanut allergy testing as determined by the more optimal positive/negative likelihood ratio.
PURPOSE OF THE TEST:	**TO DETECT SENSITIZATION TO PEANUT PROTEIN**
ROLE OF THE TEST:	**DETECTABLE OR NON-DETECTABLE SENSITIZATION CAN BE USED TO HELP INCREASE OR DECREASE THE LIKELIHOOD OF PEANUT ALLERGY BASED ON THE PRESENTING PATIENT HISTORY**
LINKED-RECOMMENDATIONS	**AD LIBITUM PEANUT INGESTION, SUPERVISED ORAL FOOD CHALLENGE TO PEANUT, PEANUT AVOIDANCE WITH/WITHOUT TREATMENT**
ANTICIPATED OUTCOMES:	Appropriate selection of the test to improve the likelihood of correct diagnosis
SETTING:	Patients presenting a to an allergist or a primary care provider for evaluation of suspected peanut allergy
PERSPECTIVE:	Patients and clinicians want to know the best diagnostic test to perform to help confirm the patients' history of suspected peanut allergy. Clinicians want to know when an oral food challenge should be performed, when it is safe to advise a patient to eat peanut, and when peanut should be avoided due to risk of an allergic reaction and consider seeking treatments
BACKGROUND:	Peanut allergy affects between 1.4% to 4.5% of the US population. This can be a potentially severe and life-long condition associated with reduced health and economic outcomes. Soon to be approved treatments can offer limited protection to a small amount of peanut but no therapy can cure the condition, but being on treatment still implies the patient is peanut allergic and must otherwise avoid intended peanut ingestion and carry emergency medication. Approximately 20%-34% will outgrow their peanut allergy. With the advent of available treatment options, it is imperative to understand how to use available diagnostic tests and interpret their results to aid in making an accurate diagnosis of peanut allergy.
SUBGROUPS:	Persons with a severe allergic reaction occurring during an observed oral food challenge; Persons with low, medium, and high pre-test probability of a suspected peanut allergy
CONFLICT OF INTERESTS:	See main document

ASSESSMENT

Problem

Is the problem a priority?

JUDGEMENT	RESEARCH EVIDENCE	ADDITIONAL CONSIDERATIONS
○ No ○ Probably no ○ Probably yes ● Yes ○ Varies ○ Don't know	The following studies support that peanut allergy, among other food allergies, is a major public health issue for children and adults in westernized countries. National Academies of Sciences E, Medicine. Finding a Path to Safety in Food Allergy: Assessment of the Global Burden, Causes, Prevention, Management, and Public Policy. Washington, DC: The National Academies Press; 2017.	The precise prevalence of peanut allergy is uncertain, given variation in the methods used to determine prevalence, and practice variation where detectable sensitization may be considered as clinical allergy without a history of symptoms arising from peanut ingestion in some circumstances. This may complicate using peanut allergy prevalence as an estimation of

	Sicherer SH, Sampson HA. Food allergy: A review and update on epidemiology, pathogenesis, diagnosis, prevention, and management. The Journal of allergy and clinical immunology 2018;141:41-58. Sampson HA, Aceves S, Bock SA, James J, Jones S, Lang D, et al. Food allergy: A practice parameter update—2014. Journal of Allergy and Clinical Immunology 2014;134:1016-25.e43. Osborne NJ, Koplin JJ, Martin PE, Gurrin LC, Lowe AJ, Matheson MC, et al. Prevalence of challenge-proven IgE-mediated food allergy using population-based sampling and predetermined challenge criteria in infants. The Journal of allergy and clinical immunology 2011;127:668-76.e1-2. Gupta RS, Springston EE, Warrier MR, Smith B, Kumar R, Pongracic J, et al. The prevalence, severity, and distribution of childhood food allergy in the United States. Pediatrics 2011;128:e9-17. Chafen JJ, Newberry SJ, Riedl MA, Bravata DM, Maglione M, Suttorp MJ, et al. Diagnosing and managing common food allergies: a systematic review. JAMA 2010;303:1848-56 Klemans RJ, van Os-Medendorp H, Blankestijn M, Bruijnzeel-Koomen CA, Knol EF, Knulst AC. Diagnostic accuracy of specific IgE to components in diagnosing peanut allergy: a systematic review. Clinical and experimental allergy : journal of the British Society for Allergy and Clinical Immunology 2015;45:720-30. Gupta RS, Warren CM, Smith BM, et al. The public health impact of parent-reported childhood food allergies in the United States. Pediatrics. 2018;142:e20181235.	pre-test probability difficult in certain contexts. The FDA has fast-tracked the development of commercial therapies to address the growing prevalence of peanut allergy. Overdiagnosis and unwarranted practice variation can create a significant healthcare burden and family quality of life. All stakeholders desire a fast, reliable diagnostic tool.

Test accuracy

How accurate is the test?

JUDGEMENT	RESEARCH EVIDENCE	ADDITIONAL CONSIDERATIONS
○ Very inaccurate ○ Inaccurate ○ Accurate ○ Very accurate ● Varies ○ Don't know	Pooled Sensitivity/Specificity positive/negative likelihood ratio (with 95% CI): Both SPT and whole peanut sIgE have high pooled sensitivity but relatively poor specificity for the diagnosis of peanut allergy proven by oral food challenge. Ara h 2 sIgE has somewhat reduced sensitivity to SPT or sIgE, but has enhanced specificity relative to these tests, and the most optimal positive/negative likelihood ration combination. Despite the individual test precision, the interpretation of the test of choice is highly dependent on an adequate suspicion of significant pre-test probability, reflected by a reasonable history that the patient had ingested peanut and demonstrated symptoms characteristic of an IgE mediated reaction. Using thresholds evaluated in the present meta-analysis, all three tests are suboptimal screening measures due to poor specificity and a high likelihood of detecting asymptomatic sensitization, potentially resulting in a false positive diagnosis. Irrespective of the test used, there are limited situations where a positive result alone relays adequate post-test odds of a peanut allergy diagnosis without the need to do a follow up oral food challenge. Based on the current analysis, in situations of low to moderate pre-test probability, detectable Ara h 2 sIgE translates to higher post-test odds of peanut allergy, compared to SPT and whole peanut sIgE. ***Where there is moderate to high pre-test probability, choice of test is less likely to influence the post-test odds, as best illustrated by the Fagan nomograms in figures3, 4 and 6.***	Skin prick testing is the traditional test of choice of the board-certified allergist and otolaryngologist with focused allergy sub-training. It is a point-of-care test that is easy to perform, exceptionally safe, inexpensive, and reliable under contexts where there is a reasonable suspicion for allergy. The advantage of this test is that the clinician can detect if sensitization is present or absent during the visit, though ambient dermatographism can affect interpretation. Serologic tests are usually performed outside of the clinical encounter. In the patient where there is strong clinical suspicion for peanut allergy, detecting sensitization through the SPT at the time of the encounter can help make the diagnosis in real-time, and allow for the patient to be counseled on avoidance and anaphylaxis management. There may be consideration that given the enhanced likelihood ratio combination that Ara h 2 is the most optimal confirmatory test to be sent after detection of sensitization on skin prick testing.

Test	Sensitivity	Specificity	Positive Likelihood	Negative Likelihood
SPT 3mm	0.97 (0.93-0.99)	0.46 (0.29-0.65)	1.82 (1.29-2.57)	0.05 (0.02-0.18)
sIgE 0.35 kU/L	0.95 (0.91-0.97)	0.38 (0.28-0.48)	1.52 (1.3-1.77)	0.14 (0.08-0.24)
Ara h 2 sIgE 0.35 kU/L	0.86 (0.81-0.89)	0.84 (0.79-0.89)	5.5 (3.99-7.56)	0.17 (0.13-0.23)

Desirable Effects

How substantial are the desirable anticipated effects?

JUDGEMENT	RESEARCH EVIDENCE	ADDITIONAL CONSIDERATIONS
○ Trivial ○ Small ● Moderate ○ Large	In setting of high pre-test probability, detection of peanut sensitization with any of the 3 tests can significantly increase the post-test odds of a peanut allergy diagnosis as shown in Figures 4 and 5.	The main advantage to the SPT over serologic IgE tests is that this is a point-of-care test that can help facilitate a diagnosis being made during the encounter. No test is a substitute or surrogate for taking a good history.

○ Varies ○ Don't know		

Undesirable Effects

How substantial are the undesirable anticipated effects?

JUDGEMENT	RESEARCH EVIDENCE	ADDITIONAL CONSIDERATIONS
○ Large ○ Moderate ○ Small ○ Trivial ● Varies ○ Don't know	In settings of low to moderate pre-test probability, detection of sensitization with any of the three tests still translates to a similar low to moderate post-test odds (figures 3, 4, and 6), and considering these results indicative of peanut allergy may significantly risk a false positive diagnosis.	**Clinician Vantage:** The level of sensitization above the positive threshold cannot be used to predict the risks of a future reaction. Likewise, test sensitization below the positive threshold in the setting of a history suggestive of high risk, cannot exclude peanut allergy. Test results, whether positive or negative, may still require an oral food challenge be performed to clarify the diagnosis **Patient vantage:** patients may have variable preferences regarding having a false positive diagnosis than a false negative diagnosis, and therefore patients may prefer an oral food challenge after the test results are known, in particular when considering entering into possible treatment for peanut allergy The clinician should be aware of the role for shared decision making and the need for decision-aids to help patients consider their options and to make the most appropriate decisions.

Certainty of the evidence of test accuracy

What is the overall certainty of the evidence of test accuracy?

JUDGEMENT	RESEARCH EVIDENCE	ADDITIONAL CONSIDERATIONS
○ Very low ○ Low ● Moderate ○ High ○ No included studies	18 studies were pooled for evidence synthesis for use of SPT at a threshold of 3mm, with sensitivity of 97% and specificity of 46%. 30 studies were pooled for evidence synthesis for peanut sIgE at a threshold of >0.35 KU_A/L, with sensitivity of 95% and specificity of 38%. 24 studies were pooled for Ara h 2 sIgE >0.35 KU_A/L, with sensitivity of 0.86 and specificity of 0.84. There was high heterogeneity among the pooled studies, and serious risk of bias, but no serious risk of indirectness, imprecision, or inconsistency. Sensitivity analysis where studies with high risk for both patient selection and flow/timing were removed had similar pooled sensitivity and specificity for all three tests. Overall there is moderate certainty in the evidence for each of the 3 tests. (Please see tables 4 and 5).	Where there is high pre-test probability, detection of peanut sensitization using any of the 3 tests can greatly increase the post-test odds of a peanut allergy diagnosis. Absence of sensitization in such patients can be helpful in lowering the odds that peanut allergy is present. The choice of which test to use is also not crucial in this setting. The Fagan nomograms in figures 4 and 5 demonstrate how the likelihood ratios translate to post-test odds in these situations, and based on these post-test odds some clinicians may feel an oral food challenge is still necessary to confirm the diagnosis. Ara h 2 may perform better than SPT or sIgE where the pre-test probability is low to moderate, but is unlikely to allow the clinician and patient to be provided with the degree of certainty to where an oral food challenge would be unnecessary to confirm a diagnosis.

Certainty of the evidence of test's effects

What is the overall certainty of the evidence for any critical or important direct benefits, adverse effects or burden of the test?

JUDGEMENT	RESEARCH EVIDENCE	ADDITIONAL CONSIDERATIONS
○ Very low ○ Low ○ Moderate	There are no included studies that detail the overall certainty or importance of direct benefits, adverse effects or burden of the tests.	Please refer to the explanation in the above box. While one may question why patients with low suspicion for peanut allergy require testing, there may be a role for shared decision making

○ High ● No included studies		where the risks and benefits of potential overdiagnosis vs. misdiagnosis are clearly explained, given some patients may clearly prefer a test be run, notwithstanding the pre-test probability.

Certainty of the evidence of management's effects

What is the overall certainty of the evidence of effects of the management that is guided by the test results?

JUDGEMENT	RESEARCH EVIDENCE	ADDITIONAL CONSIDERATIONS
○ Very low ● Low ○ Moderate ○ High ○ No included studies	Identification of the trigger of a previous episode of anaphylaxis can lead to a reduction in the risk of future anaphylactic events. Treatment options based on a positive diagnosis of peanut allergy include avoidance and carriage of epinephrine. Additionally, for some patients there may be an opportunity for treatments that desensitize the patient to the point of being able to tolerate a low threshold dose of peanut. However, we have very low certainty in the evidence that by making a diagnosis of peanut allergy that the above described options provide an unequivocal benefit for the patient. National Academies of Sciences E, Medicine. Finding a Path to Safety in Food Allergy: Assessment of the Global Burden, Causes, Prevention, Management, and Public Policy. Washington, DC: The National Academies Press; 2017. Sicherer SH, Sampson HA. Food allergy: A review and update on epidemiology, pathogenesis, diagnosis, prevention, and management. The Journal of allergy and clinical immunology 2018;141:41-58. Sampson HA, Aceves S, Bock SA, James J, Jones S, Lang D, et al. Food allergy: A practice parameter update—2014. Journal of Allergy and Clinical Immunology 2014;134:1016-25.e43. Chu DK, Wood RA, French S, et al. Oral immunotherapy for peanut allergy (PACE): a systematic review and meta-analysis of efficacy and safety. Lancet. 2019;393:2222-2232. Robinson M, Greenhawt M, Stukus D. Factors associated with epinephrine administration for anaphylaxis in children before arrival to the emergency department. Ann Allergy Asthma Immunol. 2017;119:164-169. https://icer-review.org/wp-content/uploads/2018/12/ICER_PeanutAllergy_Final_Report_071019.pdf Shaker M, Greenhawt M. Estimation of health and economic benefits of commercial peanut immunotherapy products: a cost-effectiveness analysis. JAMA Netw Open. 2019;2:e193242.	

Certainty of the evidence of test result/management

How certain is the link between test results and management decisions?

JUDGEMENT	RESEARCH EVIDENCE	ADDITIONAL CONSIDERATIONS
○ Very low ● Low ○ Moderate ○ High ○ No included studies	As reflected in the Fagan nomograms in figures 3, 4, and 6, an oral food challenge may often still be necessary to provide a definitive diagnosis and management strategy despite a positive test result, given that the systematic review suggests that even with very high pre-test probability, the post-test odds do not eclipse 90% (coming closest for the use of Ara h 2). Moreover, even with no detectable sensitization, the post-test odds are still 2-3%.	

Certainty of effects

What is the overall certainty of the evidence of effects of the test?

JUDGEMENT	RESEARCH EVIDENCE	ADDITIONAL CONSIDERATIONS

○ Very low ●Low ○ Moderate ○ High ○ No included studies	All three of the diagnostic tests for peanut allergy have high sensitivity though detection of peanut specific IgE does not always translate in to post-test odds sufficient enough to support a diagnosis of peanut allergy without a confirmatory oral food challenge. Conversely, the absence of detectable sensitization using any of these three tests should translate to very low post-test odds of a diagnosis of peanut allergy. Therefore, despite which test is used, in many cases an oral food challenge will still be indicated, and thus there is low certainty in the effects of the test in providing benefit for the patient from this perspective. Once a diagnosis is made through either testing or oral food challenge, there is low certainty in the benefits of the available treatment options. Chu DK, Wood RA, French S, et al. Oral immunotherapy for peanut allergy (PACE): a systematic review and meta-analysis of efficacy and safety. Lancet. 2019;393:2222-2232. https://icer-review.org/wp-content/uploads/2018/12/ICER_PeanutAllergy_Final_Report_071019.pdf Shaker M, Greenhawt M. Estimation of health and economic benefits of commercial peanut immunotherapy products: a cost-effectiveness analysis. JAMA Netw Open. 2019;2:e193242. National Academies of Sciences E, Medicine. Finding a Path to Safety in Food Allergy: Assessment of the Global Burden, Causes, Prevention, Management, and Public Policy. Washington, DC: The National Academies Press; 2017.	The use of these diagnostic tests at the stated thresholds (SPT 3 mm or greater, sIgE >0.35 KU_A/L, Ara h 2 > 2 KU_A/L) is most helpful in situations of high (>70%) pre-test probability in shifting the post-test odds appreciably, and can provide moderate to high certainty of a diagnosis. Testing should be undertaken with extreme caution in patients with low pre-test probability.

Values

Is there important uncertainty about or variability in how much people value the main outcomes?

JUDGEMENT	RESEARCH EVIDENCE	ADDITIONAL CONSIDERATIONS
○ Important uncertainty or variability ○ Possibly important uncertainty or variability ● Probably no important uncertainty or variability ○ No important uncertainty or variability	Quality of life and qualitative research has indicated a negative effect based upon a poor understanding of the implications of diagnostic testing in terms of the severity and prognosis of the patient's allergy. Kao LM, Greenhawt MJ, Warren CM, et al. Parental and parent-perceived child interest in clinical trials for food allergen immunotherapy. Ann Allergy Asthma Immunol. 2018;120:331-333.e1. Lewis MO, Brown-Whitehorn TF, Cianferoni A, Rooney C, Spergel JM. Peanut-allergic patient experiences after epicutaneous immunotherapy: peanut consumption and impact on QoL. Ann Allergy Asthma Immunol. 2019; 123:101-103. Ward C, Greenhawt M. Differences in caregiver food allergy quality of life between tertiary care, specialty clinic, and caregiver-reported food allergic populations. J Allergy Clin Immunol Pract. 2016;4:257-264. Waggoner MR. Parsing the peanut panic: the social life of a contested food allergy epidemic. Soc Sci Med. 2013;90:49-55. Greenhawt M, Marsh R, Gilbert H, Sicherer S, DunnGalvin A, Matlock D. Understanding caregiver goals, benefits, and acceptable risks of peanut allergy therapies. Ann Allergy Asthma Immunol. 2018;121:575-579. Greenhawt M, DunnGalvin A. Preliminary psychometric analyses and clinical performance of a caregiver self-efficacy scale for food allergy self-management. Ann Allergy Asthma Immunol. 2018;120:73-79. Greenhawt M, Dunn Galvin A, Chalil JM, Prinz M, Rogers M, Green TD. Patient and Caregiver Burden of Peanut Allergy: An Ethnographic Study. Presented at the 2019 EAACI Pediatric Allergy and Asthma Management Conference, Florence, Italy, October 18, 2019.	

Both patients and clinicians highly value an accurate diagnosis, but may be concerned about the undesirable effects highlighted above. There is emerging evidence that uncertainty of what diagnostic test results imply at the time of diagnosis may have detrimental effects on patients and their families.

There are no published data on the values and preferences of patients and families regarding performing diagnostic testing for food allergy. Specifically there are no data regarding the potential harms of a false-positive test result as compared with the potential harms of a missed diagnosis (false-negative test result), or how the future implications of the erroneous diagnosis may be handled. This could encompass a scenario where a false negative test results in no diagnosis being given, but the individual later eats a peanut containing item and has a reaction, or alternatively (and more likely), the scenario of someone diagnosed as peanut allergic based on positive testing (without a history of ingestion), who later "outgrows" the allergy and may be resentful of the possibility of a false positive diagnosis.

Balance of effects

Does the balance between desirable and undesirable effects favor the intervention or the comparison?		
JUDGEMENT	**RESEARCH EVIDENCE**	**ADDITIONAL CONSIDERATIONS**
○ Favors the comparison ○ Probably favors the comparison ○ Does not favor either the intervention or the comparison ○ Probably favors the intervention ○ Favors the intervention ● Varies ○ Don't know	The oral food challenge is the most accurate, definitive assessment of peanut allergy. However, in situations where there is high pre-test probability for peanut allergy, the three diagnostic tests can greatly assist in increasing (sensitization detected) or significantly decreasing (no sensitization detected) the post-test odds of having peanut allergy, and confirmatory oral food challenge may not always be required. Outside of a strong stated preference where there is low pre-test probability, the comparator test (oral food challenge) has more desirable effects than the intervention diagnostic tests, and can be used to avoid diagnostic misclassification. Ward C, Greenhawt M. Differences in caregiver food allergy quality of life between tertiary care, specialty clinic, and caregiver-reported food allergic populations. J Allergy Clin Immunol Pract. 2016;4:257-264. National Academies of Sciences E, Medicine. Finding a Path to Safety in Food Allergy: Assessment of the Global Burden, Causes, Prevention, Management, and Public Policy. Washington, DC: The National Academies Press; 2017. Sampson HA, Aceves S, Bock SA, James J, Jones S, Lang D, et al. Food allergy: A practice parameter update—2014. Journal of Allergy and Clinical Immunology 2014;134:1016-25.e43. Franxman T, Howe L, Teich E, Greenhawt M. Oral food challenge improves food allergy quality of life. J Allergy Clin Immunol Pract: 2015; 3: 50-56 Kansen, HM, Le, T-M, Meijer, Y, et al. The impact of oral food challenges for food allergy on quality of life: A systematic review. Pediatr Allergy Immunol. 2018; 29: 527– 537. https://doi.org/10.1111/pai.12905	The risks of a false positive test are significant and may lead to prolonged unnecessary avoidance and costs, as well as potential stigma related to being classified as being peanut allergic. Particularly at young ages, over-diagnosis by isolated positive tests of sensitization may also lead to a lost opportunity to establish peanut tolerance.

Resources required

How large are the resource requirements (costs)?

JUDGEMENT	**RESEARCH EVIDENCE**	**ADDITIONAL CONSIDERATIONS**
○ Large costs ○ Moderate costs ○ Negligible costs and savings ○ Moderate savings ○ Large savings ● Varies ○ Don't know	There are no studies that directly investigated the resources requirements	All the possible interventions and the comparator tests do require resources in terms of both direct and indirect costs. These costs and cost burdens may vary depending on the healthcare system in question but are likely already nested into the cost of normal practice operation. Newer management options based on test results may have additional costs that have not been studied. Costs may vary based on the particular healthcare system and geography but these largely fall into overlapping ranges across the US. Skin testing (CPT code 95004) may have more variability in terms of cost and reimbursement than serologic IgE testing (CPT code 86003) based on a selected sample of US cities in different parts of the country, detailed below: Lebanon, NH: 95004 code $9-28; 86003 code $15-$98 New York City: 95004 code $9-28; 86003 code $15-$98 Winston-Salem, NC: 95004 code $8-23; 86003 code $15-$98 Miami, FL: 95004 code $8-25; 86003 code $15-$98 Kansas City, MO: 95004 code $8-25; 86003 code $15-$98 Duluth, MN: 95004 code $9-25; 86003 code $15-$98

		Denver, CO: 95004 code $8-23; 86003 code $15-$98 Eugene, OR: 95004 code $9-28; 86003 code $15-$98 Los Angeles, CA: 95004 code $8-25; 86003 code $15-$98

Certainty of evidence of required resources

What is the certainty of the evidence of resource requirements (costs)?

JUDGEMENT	RESEARCH EVIDENCE	ADDITIONAL CONSIDERATIONS
○ Very low ○ Low ○ Moderate ○ High ●No included studies	There are no studies that directly assessed the certainty of evidence of resource requirement.	There would not be any anticipated new resources needed to support the use of any of these tests that are not already established and in use in clinical practice. There may be additional resources required to offer Ara h 2 as a stand-alone test, as opposed to a full component panel. Initially more expensive but then cheaper later. Operating costs vary from region to region and depend on practice location, personnel experience, and practice volume. While 95% of practicing allergists offer oral food challenges, only 17% perform more than 10 per month, which could complicate access to confirm diagnostic test results.

Cost effectiveness

Does the cost-effectiveness of the intervention favor the intervention or the comparison?

JUDGEMENT	RESEARCH EVIDENCE	ADDITIONAL CONSIDERATIONS
○ Favors the comparison ○ Probably favors the comparison ○ Does not favor either the intervention or the comparison ○ Probably favors the intervention ○ Favors the intervention ● Varies ○ No included studies	The cost-effectiveness of diagnostic testing varies based on which test is chosen. A cost-benefit analysis as part of this document shows that use of skin prick testing as opposed to use of Ara h 2 testing is not cost-effective and is associated with higher societal costs related to the risk of false positive results, leading to a patient who is not truly peanut allergic being managed as such. Skin prick testing remains associated with higher costs and lower benefits as a choice of test (e.g. "dominated" in economic terms) in the analysis until the specificity of Ara h 2 decreases significantly from the values identified in the meta-analysis. Deterministic sensitivity analysis did not reveal other factors related to assessing a patient for peanut allergy with diagnostic testing, that, if changed, could make this test more cost-effective than Ara h 2.	Skin prick testing to peanut has lower specificity than Ara h 2 testing, and will result in more falsely positive diagnoses identified, resulting in lower QALY accumulation. However, with the marginal increase in sensitivity, SPT would result in a slightly lower rate of peanut allergic reactions.

Equity

What would be the impact on health equity?

JUDGEMENT	RESEARCH EVIDENCE	ADDITIONAL CONSIDERATIONS
○ Reduced ○ Probably reduced ○ Probably no impact ○ Probably increased ○ Increased ○ Varies ● Don't know	There are no studies that directly assessed the impact on equity.	Serologic testing is more widely available and less dependent on allergy specialists which may improve equity potentially, whereas skin testing is the opposite. Certain states have different reimbursement rules/rates for skin vs. serologic testing, which could reduce equity if certain of these tests are not available, based on location or insurance.

Acceptability

Is the intervention acceptable to key stakeholders?

JUDGEMENT	RESEARCH EVIDENCE	ADDITIONAL CONSIDERATIONS
○ No ○ Probably no ○ Probably yes ○ Yes ○ Varies ● Don't know	There are not studies that directly assess if the intervention is acceptable to key stakeholders	**Clinician vantage:** Multiple prior practice parameters have echoed these findings; however, there is well-known practice variation with respect to indication for testing, and interpretation of tests in certain contexts. The clinician may not accept or follow guidelines that advise against their current practices, their training, or their comfort level with decision-making. **Patient/Advocate Vantage:** patients may have variable preferences regarding how diagnostic testing is used, and as stated above may differentially value having a false positive diagnosis than a false negative diagnosis, or having to undergo an oral challenge to confirm a diagnosis. There could be a role for shared decision making and a decision-aid to help patients consider their options.

Feasibility

Is the intervention feasible to implement?

JUDGEMENT	RESEARCH EVIDENCE	ADDITIONAL CONSIDERATIONS
○ No ○ Probably no ○ Probably yes ○ Yes ○ Varies ● Don't know	There are no studies that evaluated the feasibility of implementing these findings.	This should be feasible to implement but implementation could be limited by lack of availability of Ara h 2 as a stand-alone test. Variable reimbursement of allergy testing services may also limit access to care and implementation. An even more problematic implementation would be if there are sufficient resources at all allergy practices to support an increased need for subsequent oral food challenge to confirm diagnosis when indicated. Not all allergy practices offer oral food challenges and most primary care providers would not be conducting oral food challenges.

SUMMARY OF JUDGEMENTS

	JUDGEMENT						
PROBLEM	No	Probably no	Probably yes	**Yes**		Varies	Don't know
TEST ACCURACY	Very inaccurate	**I**naccurate	Accurate	Very accurate		**Varies**	Don't know
DESIRABLE EFFECTS	Trivial	Small	**Moderate**	Large		Varies	Don't know
UNDESIRABLE EFFECTS	Large	Moderate	Small	Trivial		**Varies**	Don't know
CERTAINTY OF THE EVIDENCE OF TEST ACCURACY	Very low	Low	**Moderate**	High			No included studies
CERTAINTY OF THE EVIDENCE OF TEST'S EFFECTS	Very low	Low	**Moderate**	High			No included studies
CERTAINTY OF THE EVIDENCE OF MANAGEMENT'S EFFECTS	Very low	**Low**	Moderate	High			No included studies
CERTAINTY OF THE EVIDENCE OF TEST RESULT/MANAGEMENT	Very low	**Low**	Moderate	High			No included studies
CERTAINTY OF EFFECTS	Very low	Low	**Moderate**	High			No included studies
VALUES	Important uncertainty or variability	Possibly important uncertainty or variability	**Probably no important uncertainty or variability**	No important uncertainty or variability			
BALANCE OF EFFECTS	Favors the comparison	Probably favors the comparison	Does not favor either the intervention or the comparison	**Probably favors the intervention**	Favors the intervention	Varies	Don't know
RESOURCES REQUIRED	Large costs	Moderate costs	Negligible costs and savings	Moderate savings	Large savings	**Varies**	Don't know

	JUDGEMENT						
CERTAINTY OF EVIDENCE OF REQUIRED RESOURCES	Very low	Low	**Moderate**	High			No included studies
COST EFFECTIVENESS	Favors the comparison	Probably favors the comparison	Does not favor either the intervention or the comparison	Probably favors the intervention	Favors the intervention	**Varies**	No included studies
EQUITY	Reduced	Probably reduced	Probably no impact	Probably increased	Increased	Varies	**Don't know**
ACCEPTABILITY	No	Probably no	Probably yes	Yes		Varies	**Don't know**
FEASIBILITY	No	Probably no	Probably yes	Yes		Varies	**Don't know**

TYPE OF RECOMMENDATION

Strong recommendation against the intervention	Conditional recommendation against the intervention	Conditional recommendation for either the intervention or the comparison	**Conditional recommendation for the intervention**	Strong recommendation for the intervention
○	○	○	●	○

CONCLUSIONS

Recommendation

Recommendations 1a: We suggest in favor of diagnostic (skin prick or serum sIgE) testing for peanut allergy in patients with a 1) physician-judged high pre-test probability of peanut allergy, or 2) prior to an oral food challenge for patients with moderate pre-test probability of peanut allergy, with whom shared decision-making has been employed to arrive at the final decision. Conditional recommendation; Certainty of evidence: very low

Recommendation 1b: We suggest against diagnostic testing in patients with a low or very low pre-test probability of peanut allergy. Conditional recommendation; Certainty of evidence: very low

Recommendation 2a: We suggest in favor of Ara h2 diagnostic testing in a patient presenting for evaluation of suspected peanut allergy for which a single diagnostic test is to be used, as Ara h2 would provide the best diagnostic accuracy as determined by the more optimal positive/negative likelihood ratio. Conditional recommendation. Certainty of evidence: moderate

Recommendation 2b: We suggest against component testing in addition to either skin prick test or sIgE to whole peanut to increase diagnostic accuracy. Conditional recommendation. Certainty of evidence: moderate.

Recommendation 3: We suggest against the clinician using the results of a SPT, sIgE to whole peanut extract, or sIgE to peanut components to determine the severity of a previous reaction and/or allergy phenotype or to predict the severity of a future reaction, Conditional recommendation. Certainty of evidence: very low.

Technical remarks:
It is critical to consider diagnostic test performance in the context of the pre-test probability of peanut allergy. The clinician should recognize the circumstances where one or more of the peanut diagnostic tests may not translate to a clinically meaningful improved post-test odds of peanut allergy. Except in cases of high pre-test probability, it is likely that an oral food challenge will be needed to establish the diagnosis of peanut allergy, regardless of the results of the selected diagnosis test(s).

Certain tests may be more appropriate than others in particular situations. We suggest that the choice of SPT, sIgE, or Ara h 2 sIgE is not critical in circumstances where there is high pre-test probability of peanut allergy.

While testing of patients with low pre-test probability is not generally recommended, if the decision is made to test in these circumstances, from a test precision standpoint, use of Ara h 2 rather than SPT or sIgE can help decrease misclassification of patients as peanut allergic, leading to less harm through falsely positive diagnosis. When testing individuals with low pre-test probability, it is recommended that an oral food challenge still be performed to validate the clinical significance of the detection of sensitization, given that the low pre-test probability in the setting of detectable sensitization translates to only moderate post-test odds of a diagnosis.

Justification

Overall justification
In patients with a high pre-test probability for peanut allergy, SPT, sIgE, and Ara h 2 sIgE are highly sensitive and reliable tests that can be considered for routine use in the diagnosis of peanut allergy.
Detailed justification
Test accuracy
These are tests with high sensitivity
Desirable Effects
Detection of sensitization in an individual with likely or suspected peanut allergy will aid considerably in confirming the diagnosis. Choice of test in circumstances where there is high pre-test probability is not critical. The absence of sensitization is helpful in ruling out the diagnosis (although in many cases, oral food challenge will still be necessary).

Subgroup considerations

Severity of reaction was investigated as a potential subgroup. These tests do not perform well to identify individuals for potentially severe reactions at the dichotomous thresholds investigated. Data are limited that may better inform if these tests have higher or lower value within other particular subgroups. In infants meeting high-risk criteria for early peanut introduction, SPT is often used; however, incorporation of Ara h 2 might result in a lower rate of over-diagnosis. Unfortunately, evidence comparing these tests is this particular population is lacking.

Implementation considerations

These are tests that are already in routine use or routinely available for use; however, testing for Ara h 2 as a single component would be needed to implement routine use of this test. In many instances, the clinician is already using these tests in tandem, potentially. If the clinician starts with the SPT in the office setting, using the Fagan nomogram in figure 3, the post-test odds could then reasonably be used as the pre-test odds for choice of a confirmatory test, represented by figure 3 for sIgE or figure 6 for Ara h 2. In this setting, given higher specificity and higher positive likelihood ratio, Ara h 2 may be the better choice of a confirmatory test if it could be obtained as a stand-alone measure.

Monitoring and evaluation

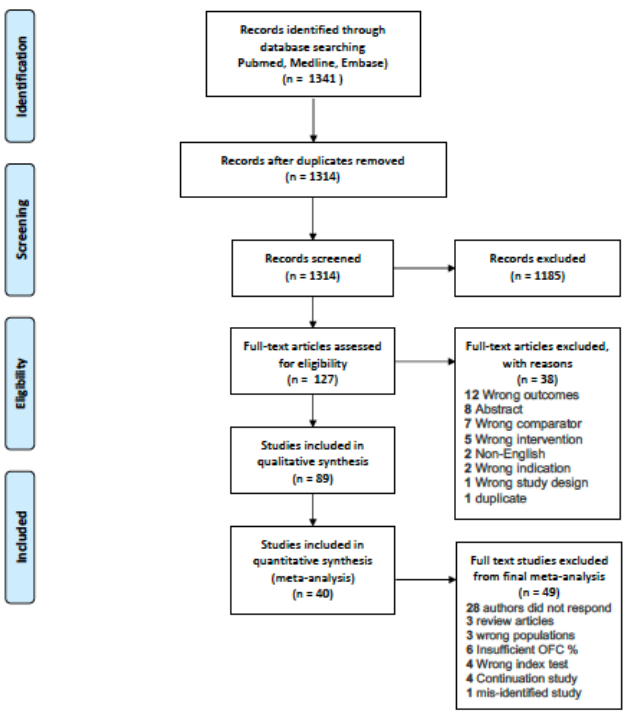
Additional meta-analysis at different cut-off points may help inform decision making, in particular for the severity subgroup or use of tests in sequence/tandem. We would recommend to journal editors that there be a requirement for future reporting of studies investigating diagnostic test precision in relation to food challenge outcome that raw data be included as a supplemental text denoting the challenge outcome, the numeric quantity of the test, the sequence of testing run if multiple tests were assessed simultaneously, and any data on severity of the reaction. This would allow for a repository to be created that would greatly assist with updating practice guidelines. For study authors to make such deidentified data available, it would enable more direct assessment of test performance as a continuous variable, which would allow for different diagnostic thresholds could be directly assessed and compared for the purposes of meta-analysis and systematic review, as opposed to having to rely on dichotomous assessment of pre-selected thresholds and potential back calculation of sensitivity/specificity. These factors serve as distinct limitations with regard to this particular document.

Research priorities

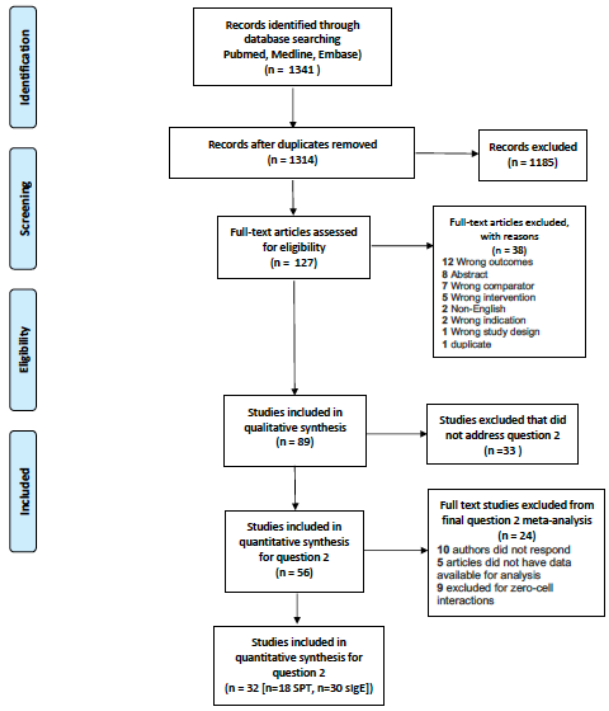
Additional studies in more unselected populations, and at a population level are needed. Future research studies reporting diagnostic sensitivity and specificity should report the true/false positive and true/false negative patient level results to assist in future meta-analyses where cut off levels would be easier to assess. If these data were available, it would have permitted analysis of the sensitization levels as a continuous variable rather than a dichotomous variable and potentially allowed better comparison of tests used sequentially or in tandem. Better data are needed to help inform what defines low, moderate, or high pre-test probability in a patient being assessed for possible peanut allergy, as well as to understand how clinicians and patients may perceive risk.

Figure 1: PRISMA Diagrams

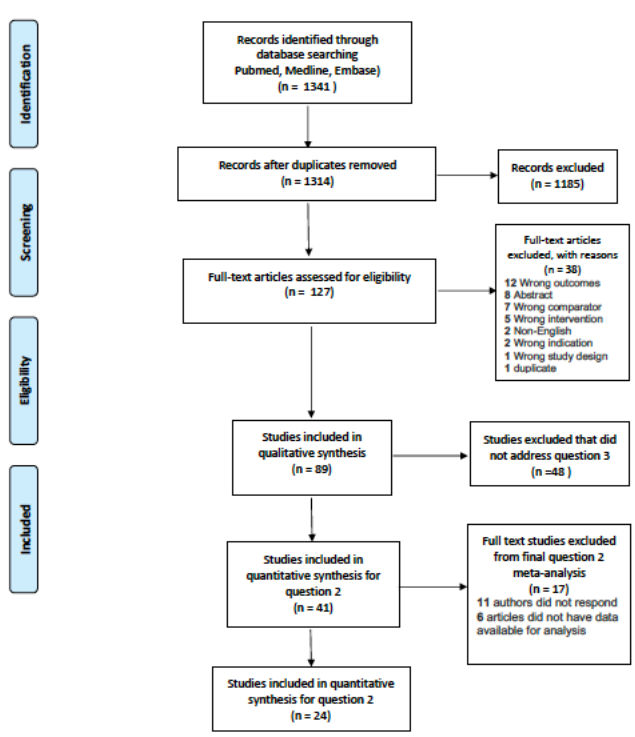
A.



B.



C.



D.

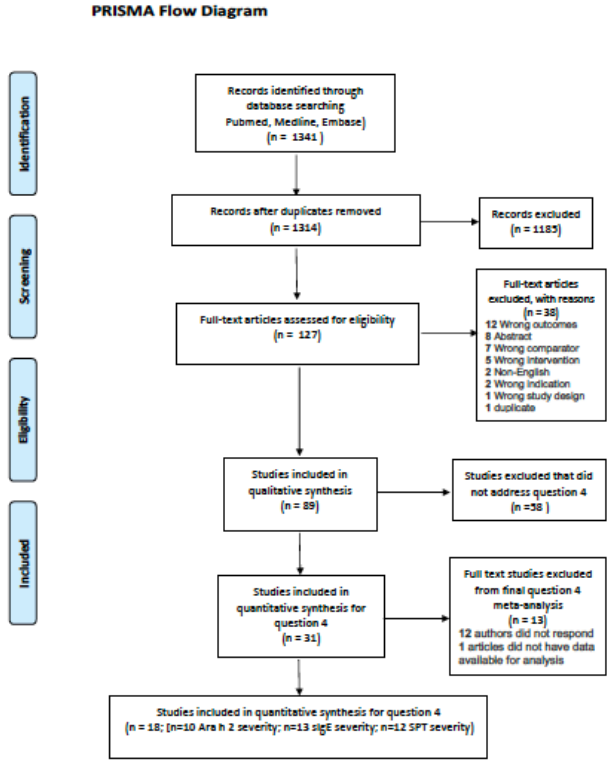
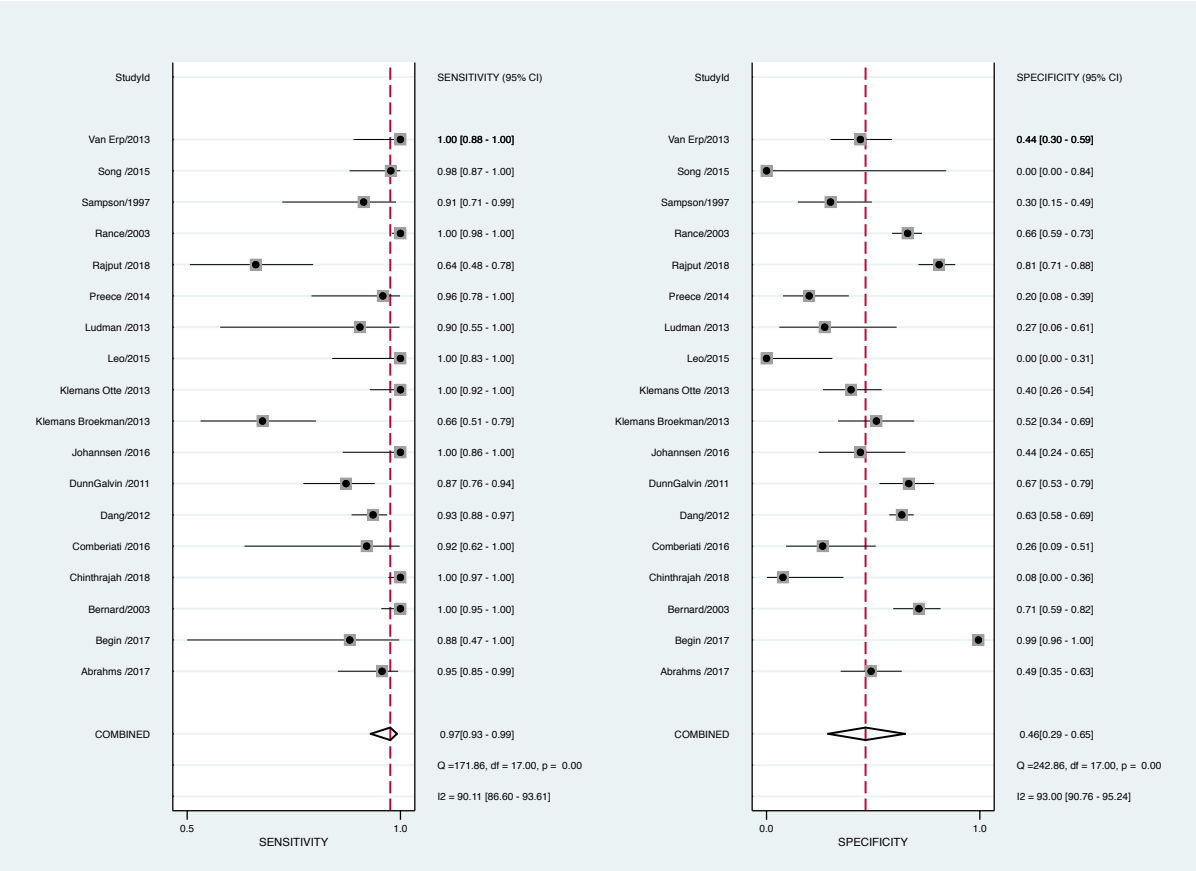


Figure 2: Summary Forest Plots for Sensitivity and Specificity of Skin Prick Testing at 3mm and sIgE testing at 0.35KU_A/L

a



b

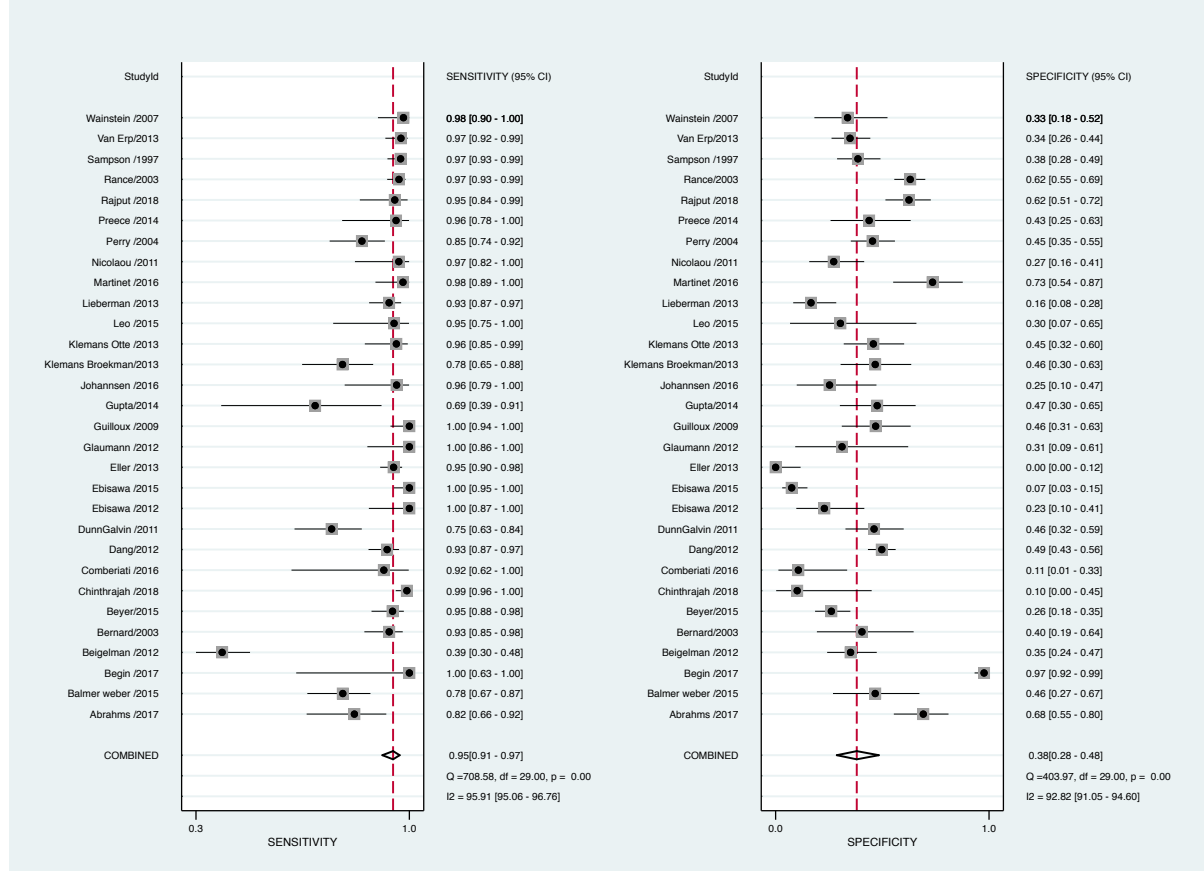
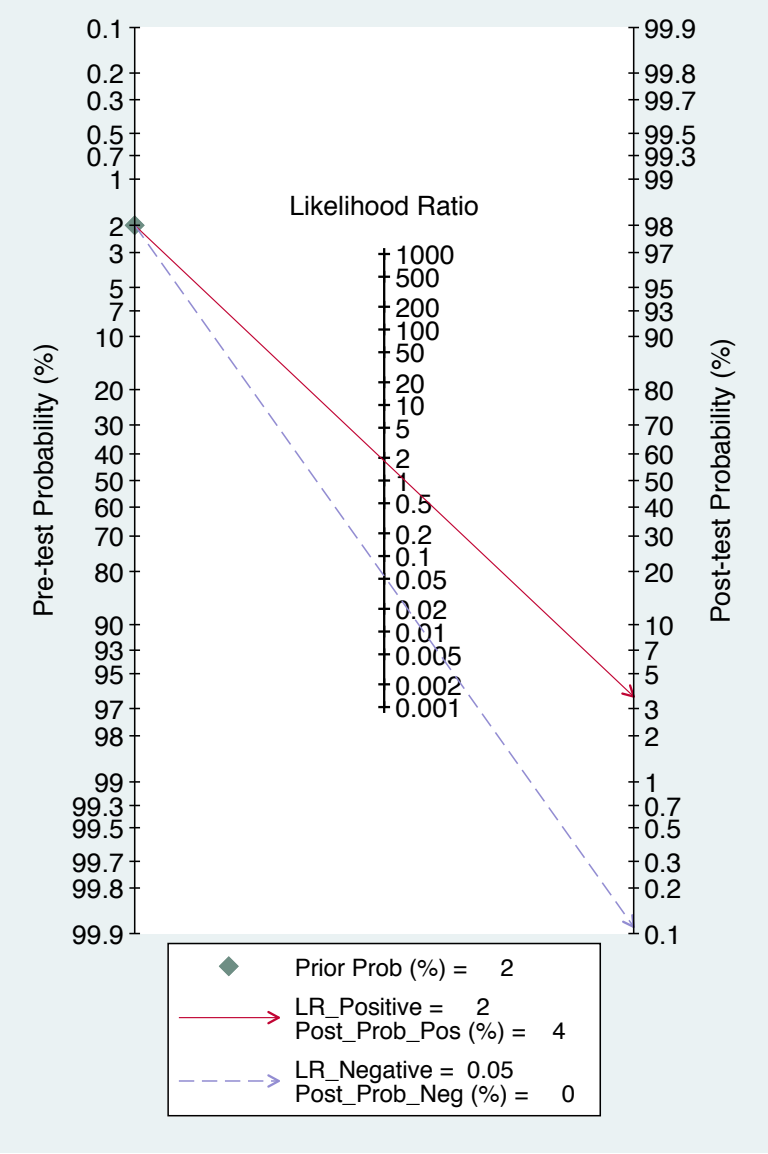
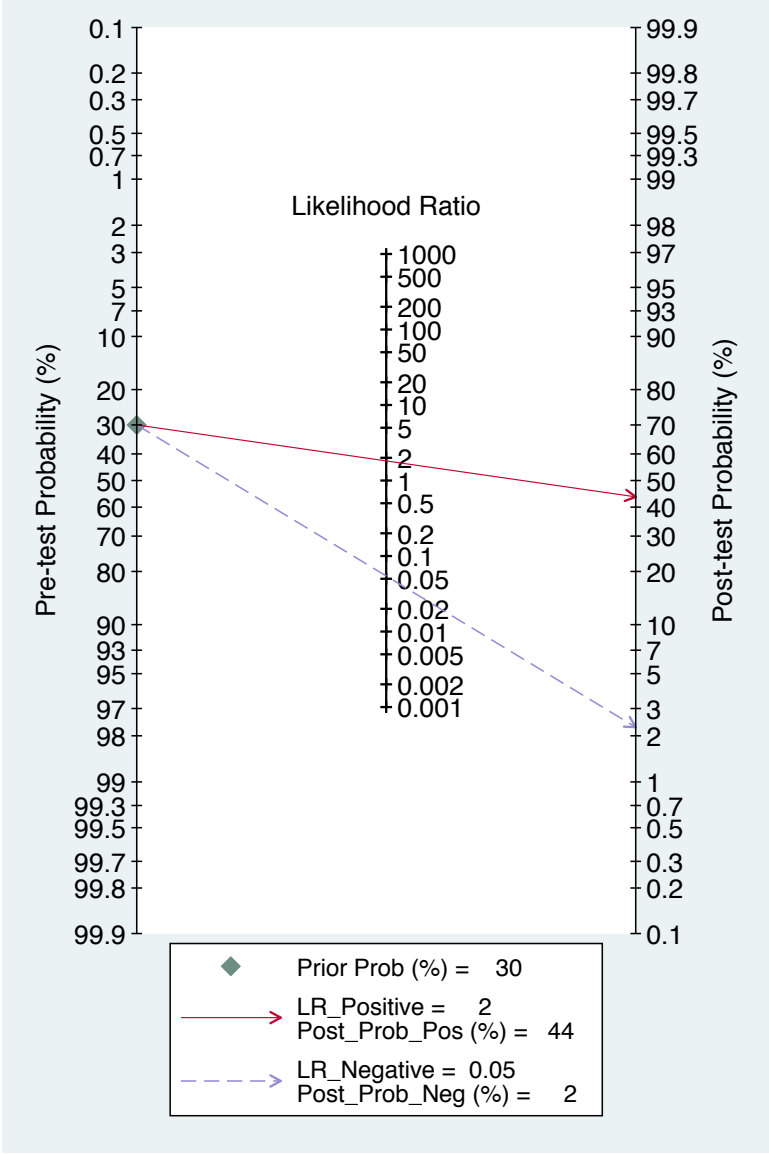


Figure 3: Fagan Nomograms for SPT 3mm Performance at Low, Moderate, and High Pre-Test Probability

a



b



c

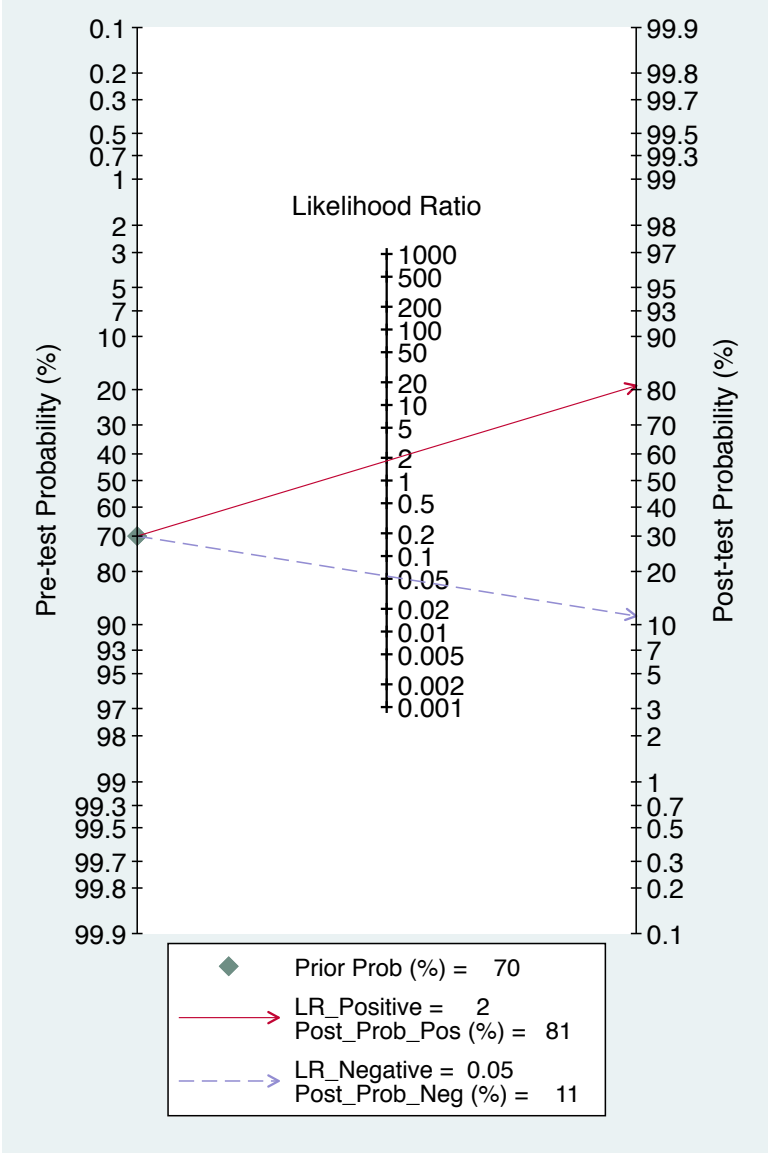


Figure 4: Fagan Nomograms for sIgE 0.35KU$_A$/L Performance at Low, Moderate, and High Pre-Test Probability

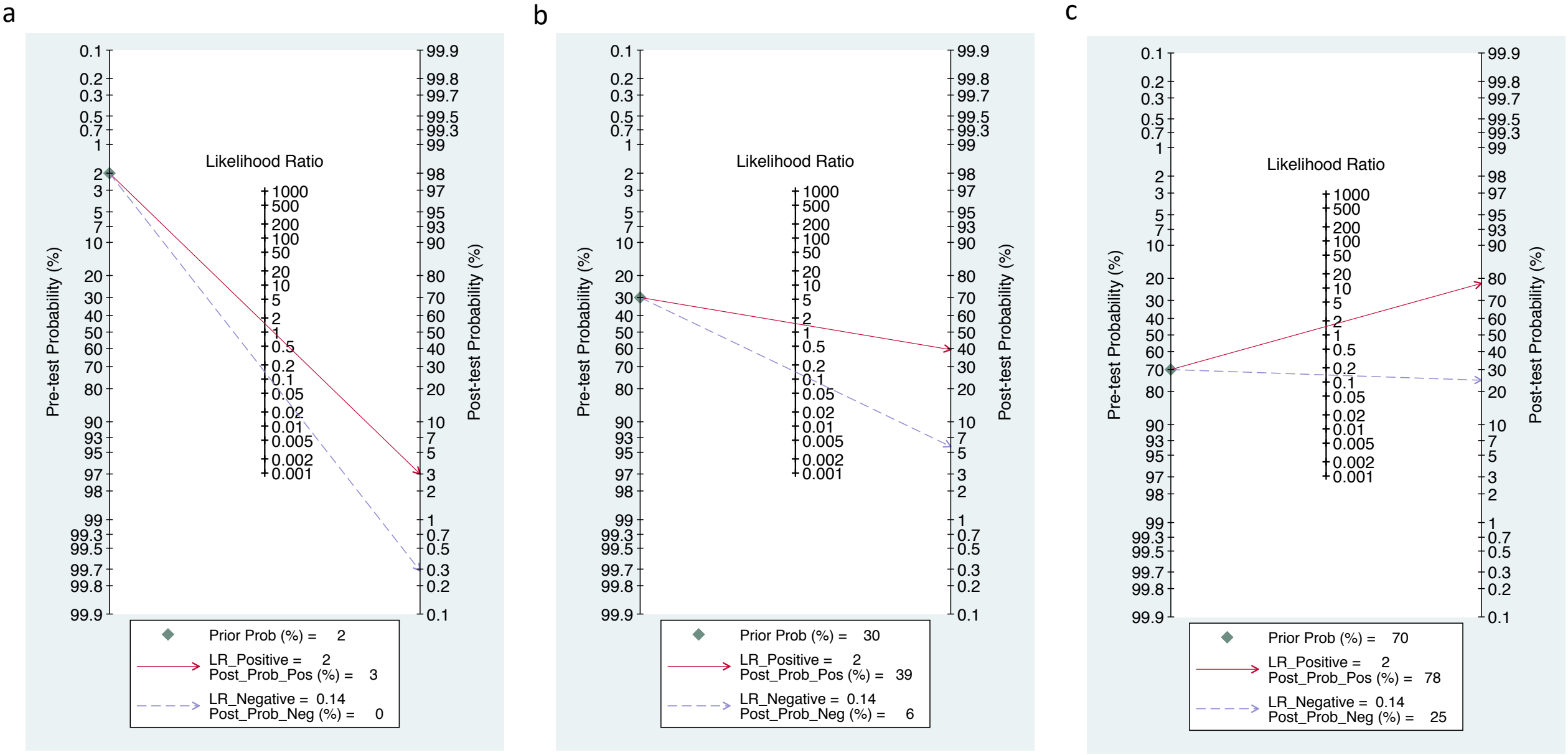


Figure 5: Summary Forest Plots for Sensitivity and Specificity of Ara h 2 sIgE testing at 0.35KU$_A$/L

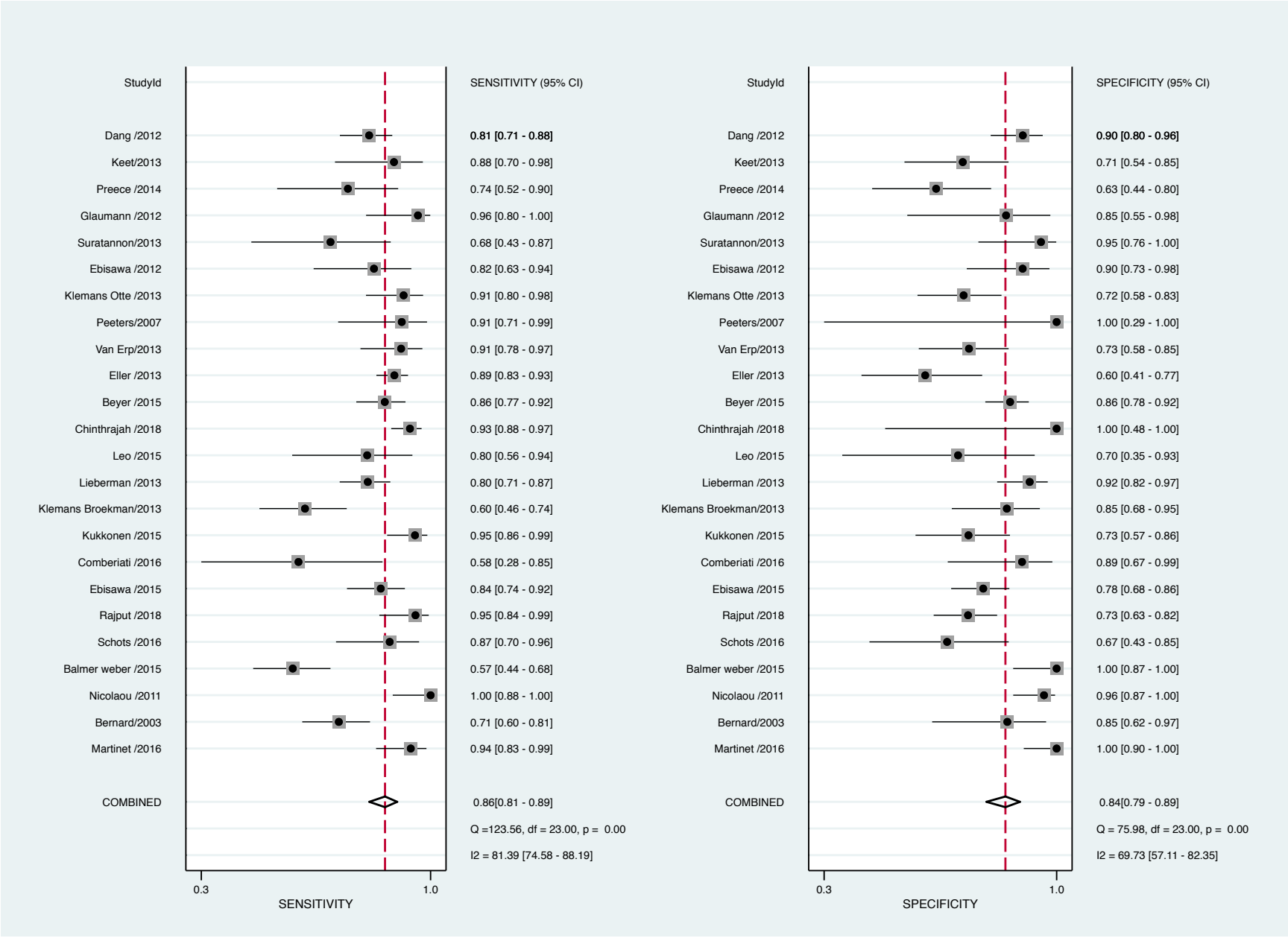
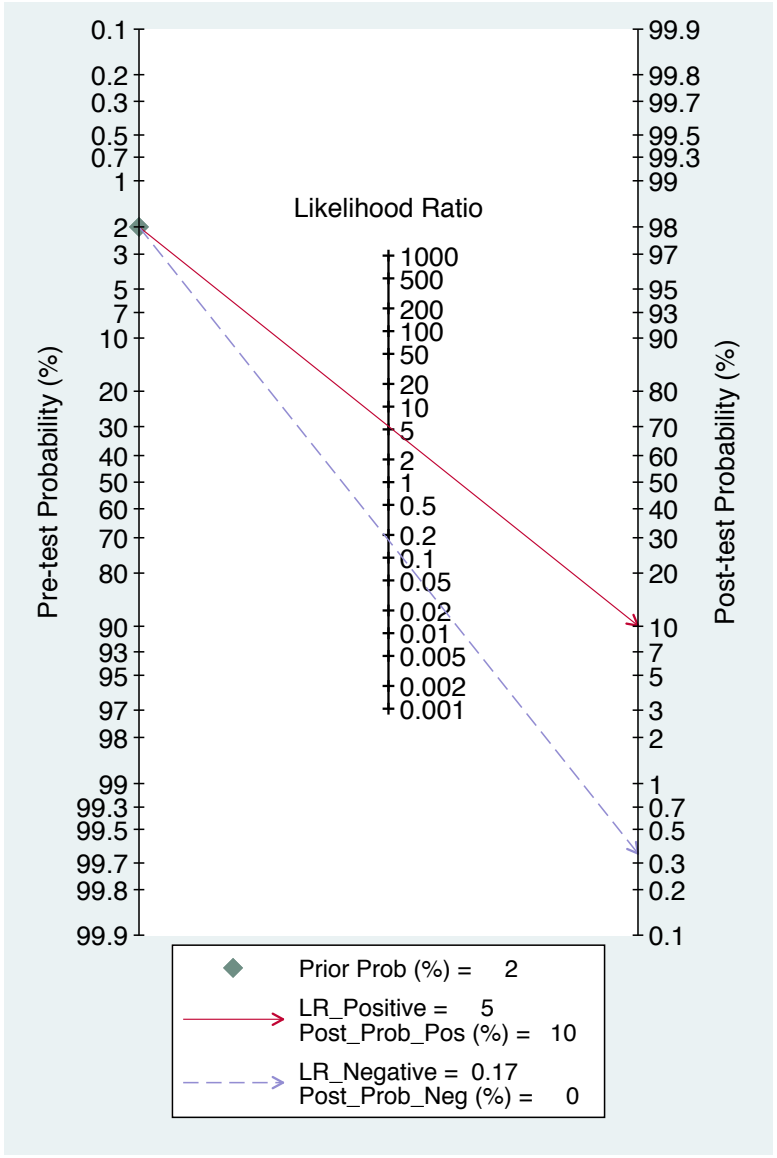
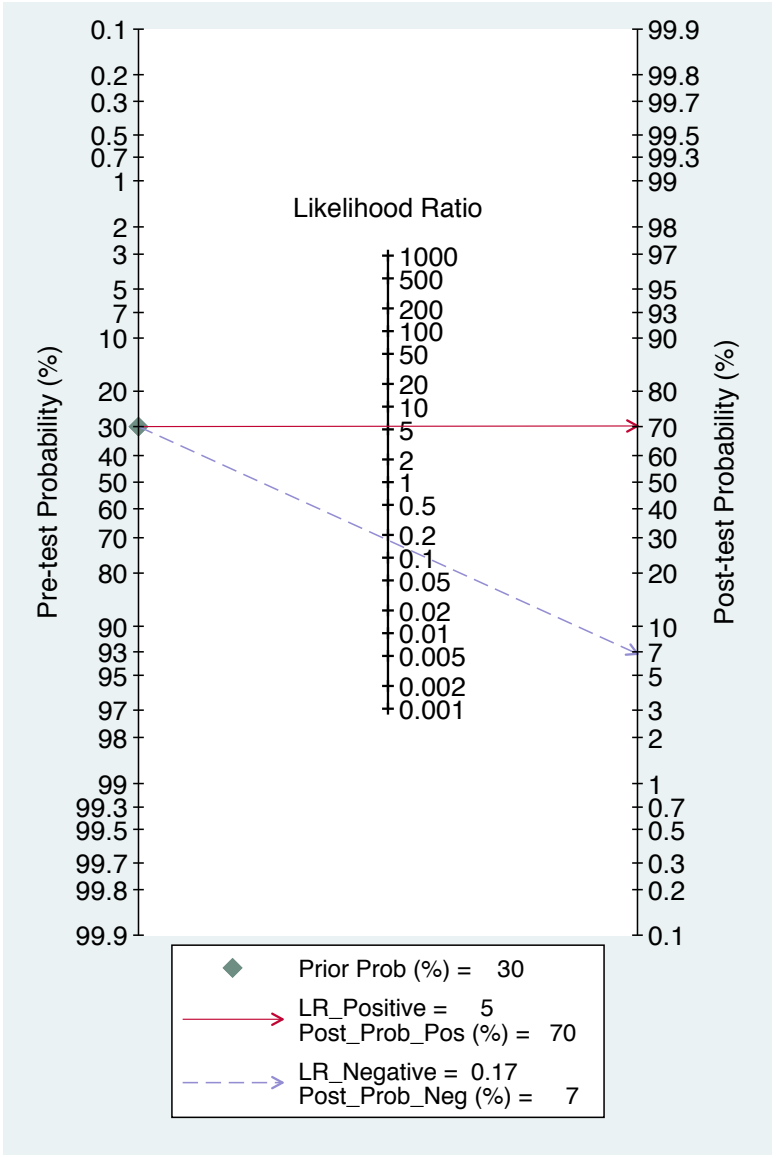


Figure 6 Fagan Nomograms for Ara h 2 sIgE 0.35KU$_A$/L Performance at Low, Moderate, and High Pre-Test Probability

a



b



c

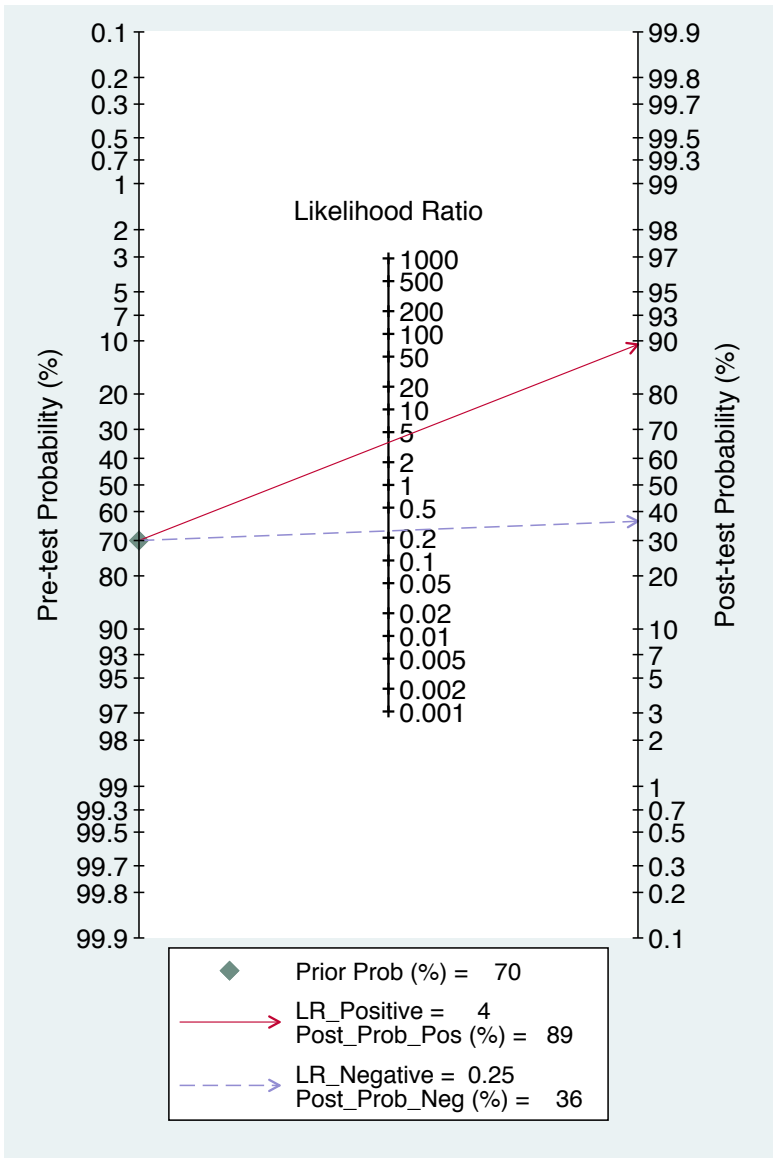


Figure 7: Summary Forest Plots for Sensitivity and Specificity of Ara h 2 sIgE testing at 2 KU_A/L Indicating a Severe Reaction

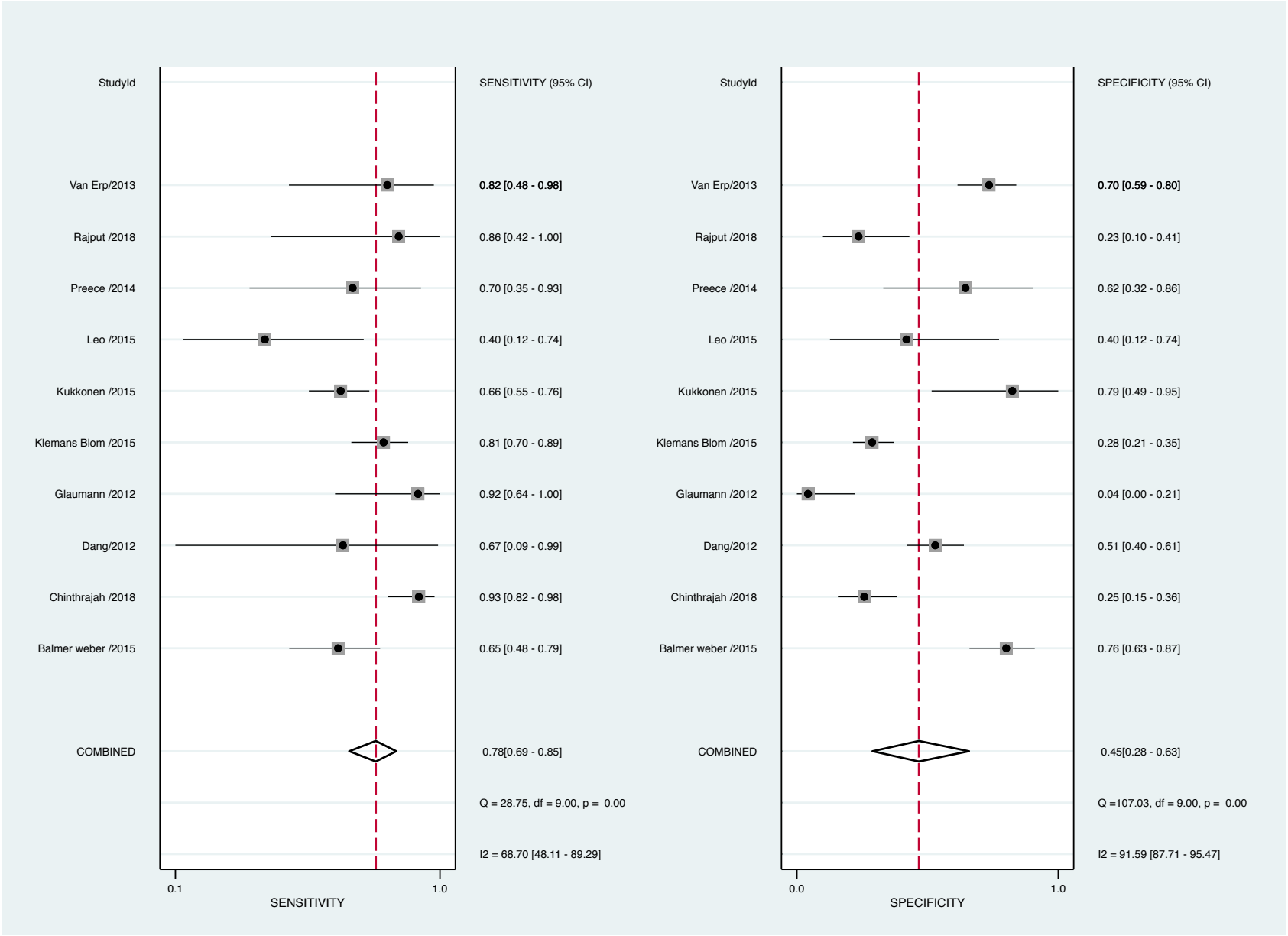


Figure 8: Summary Forest Plots for Sensitivity and Specificity of sIgE testing at 50 KU_A/L Indicating a Severe Reaction

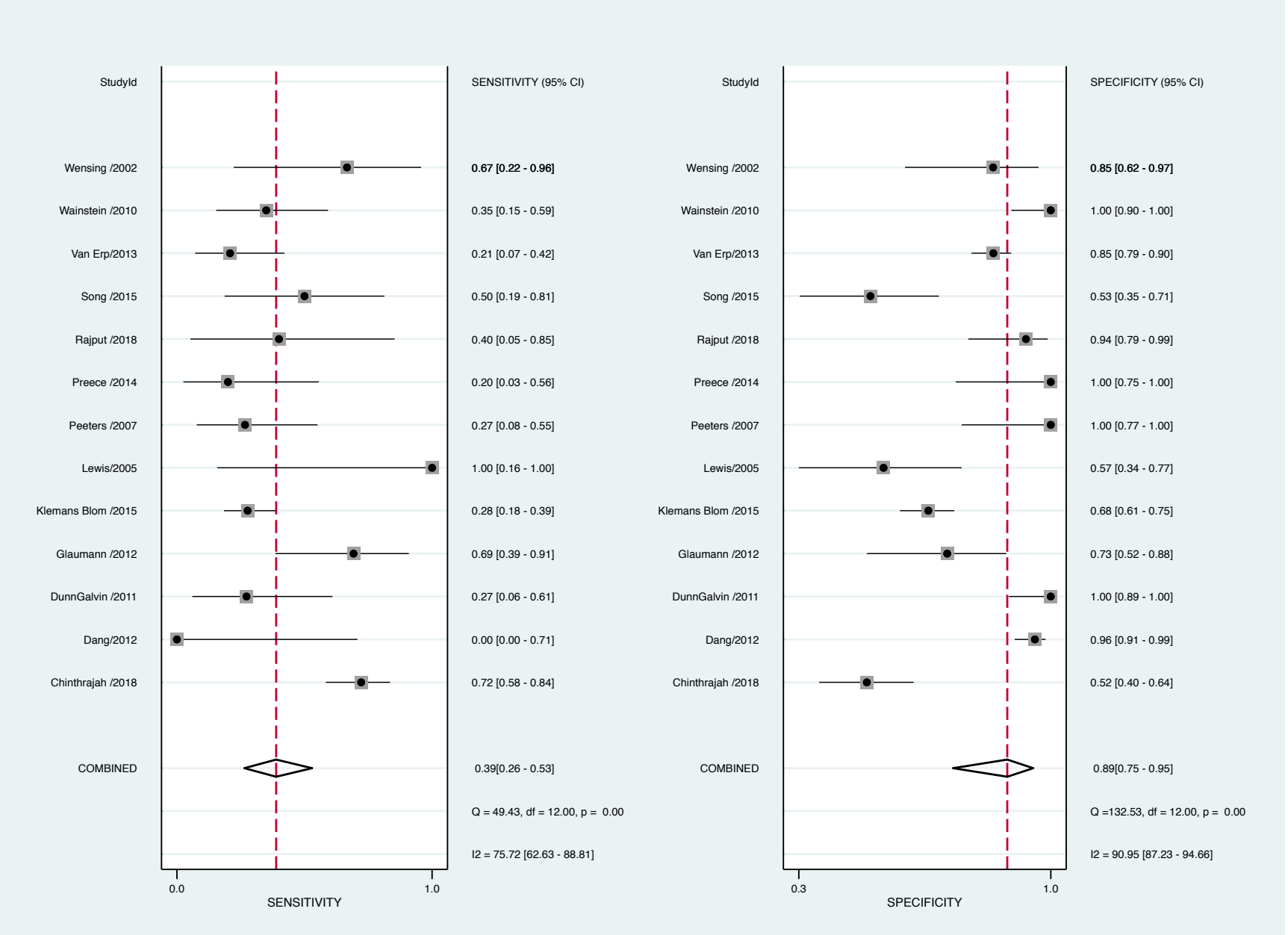


Figure 9: Summary Forest Plots for Sensitivity and Specificity of Skin Prick Testing at 10mm Indicting a Severe Reaction

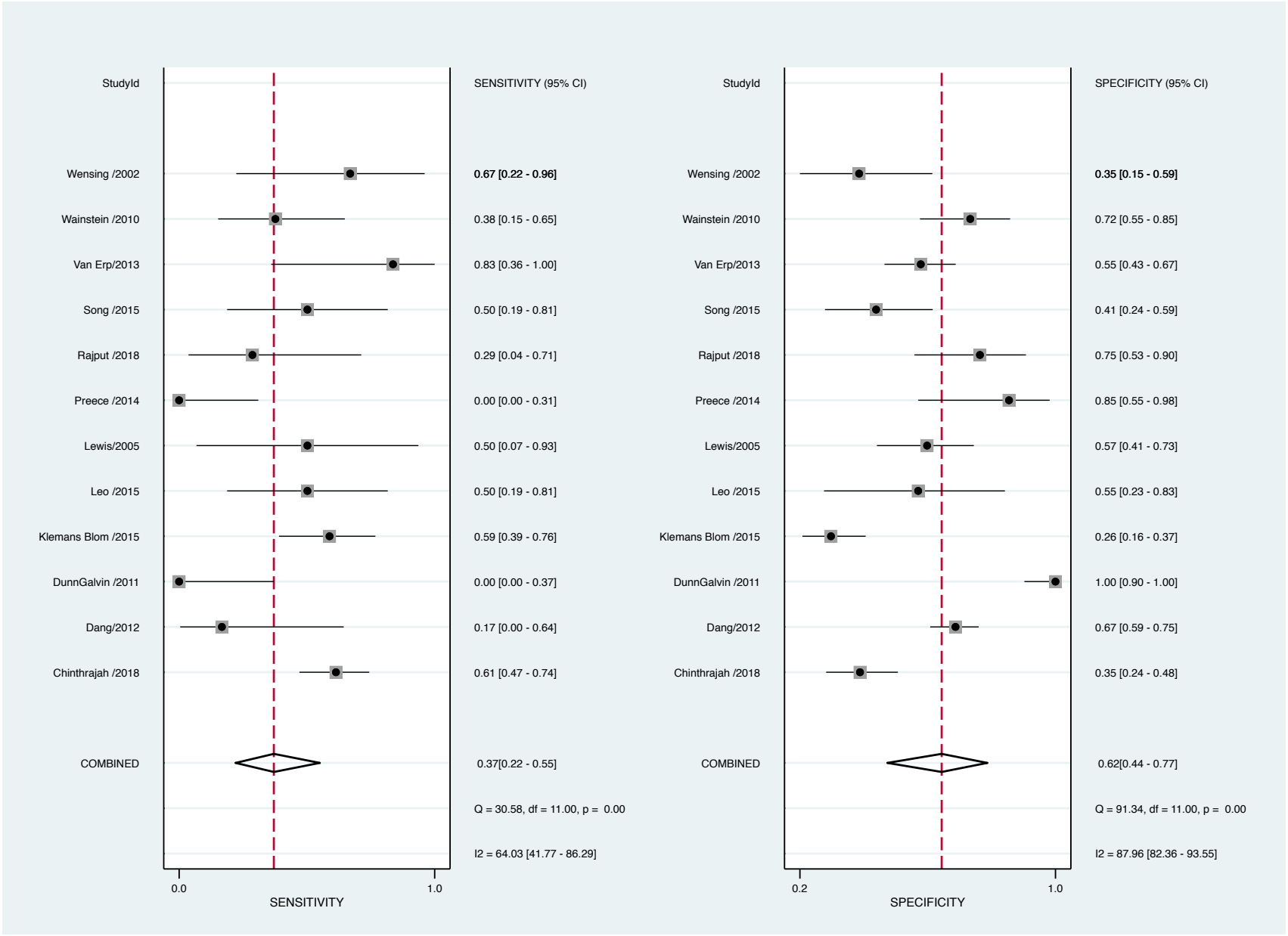
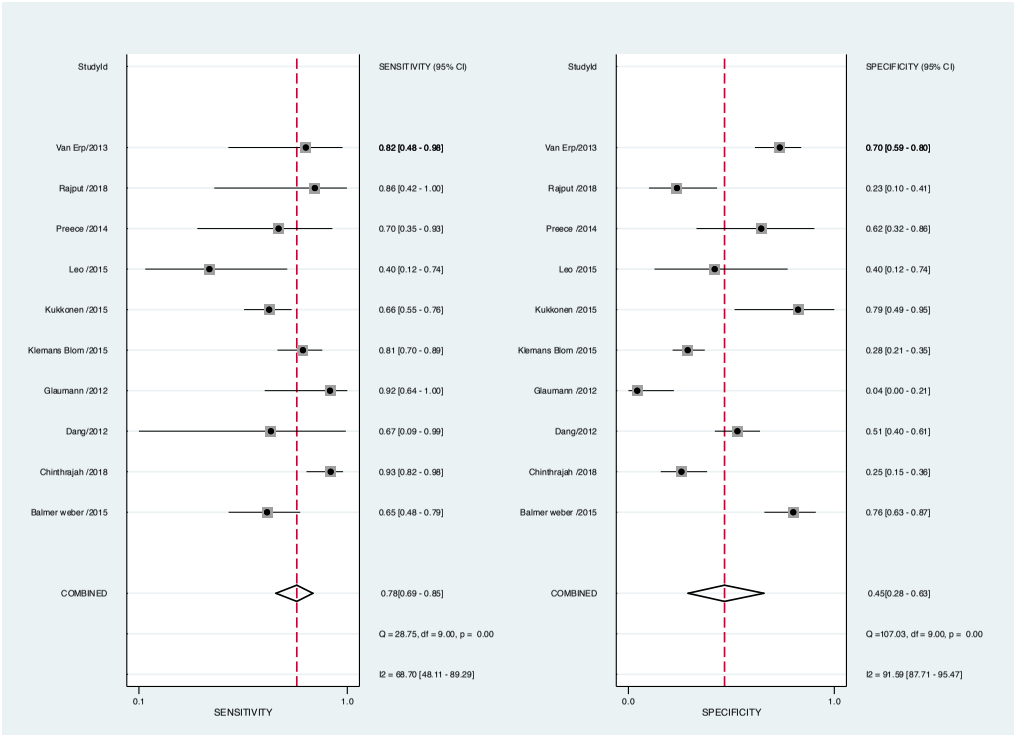
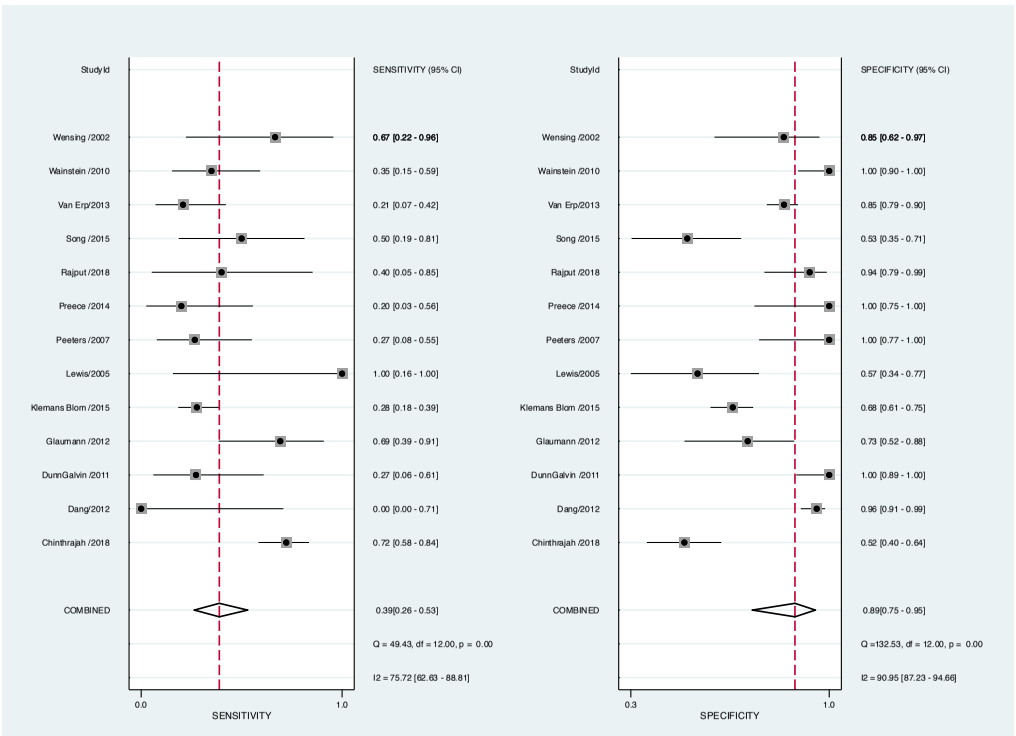


Figure 7: Summary Forest Plots for Sensitivity and Specificity of Ara h 2 sIgE 2KU$_A$/L, sIgE testing at 50 KU$_A$/L and Skin Prick Testing at 10mm In Indicating a Severe Reaction

A



B



C

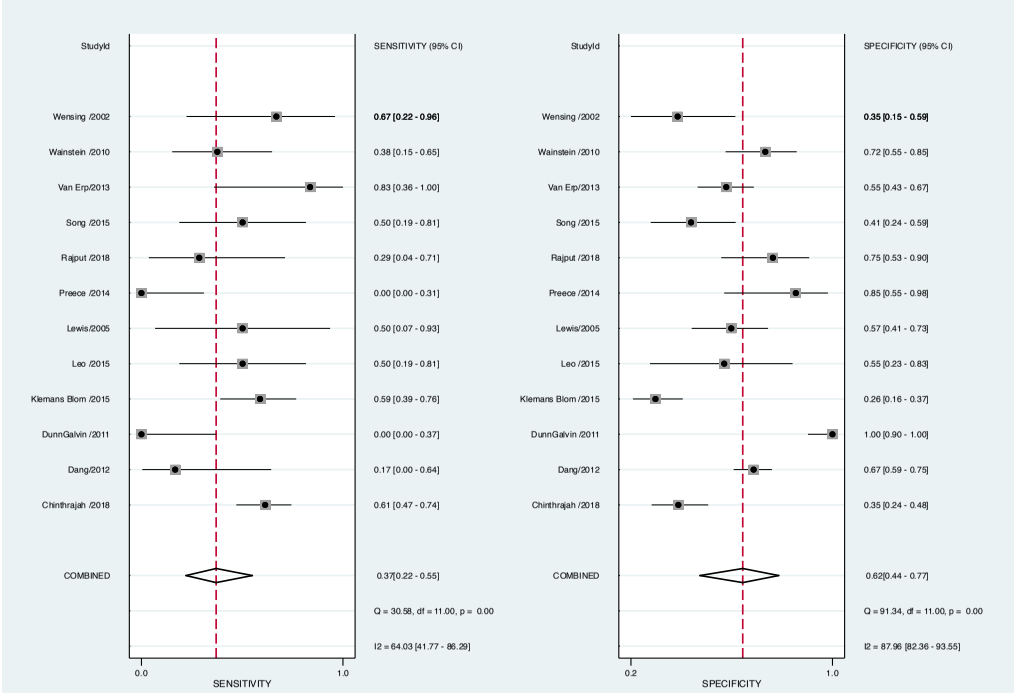


Figure 10: Outcomes of Using Diagnostic Testing for Peanut Allergy

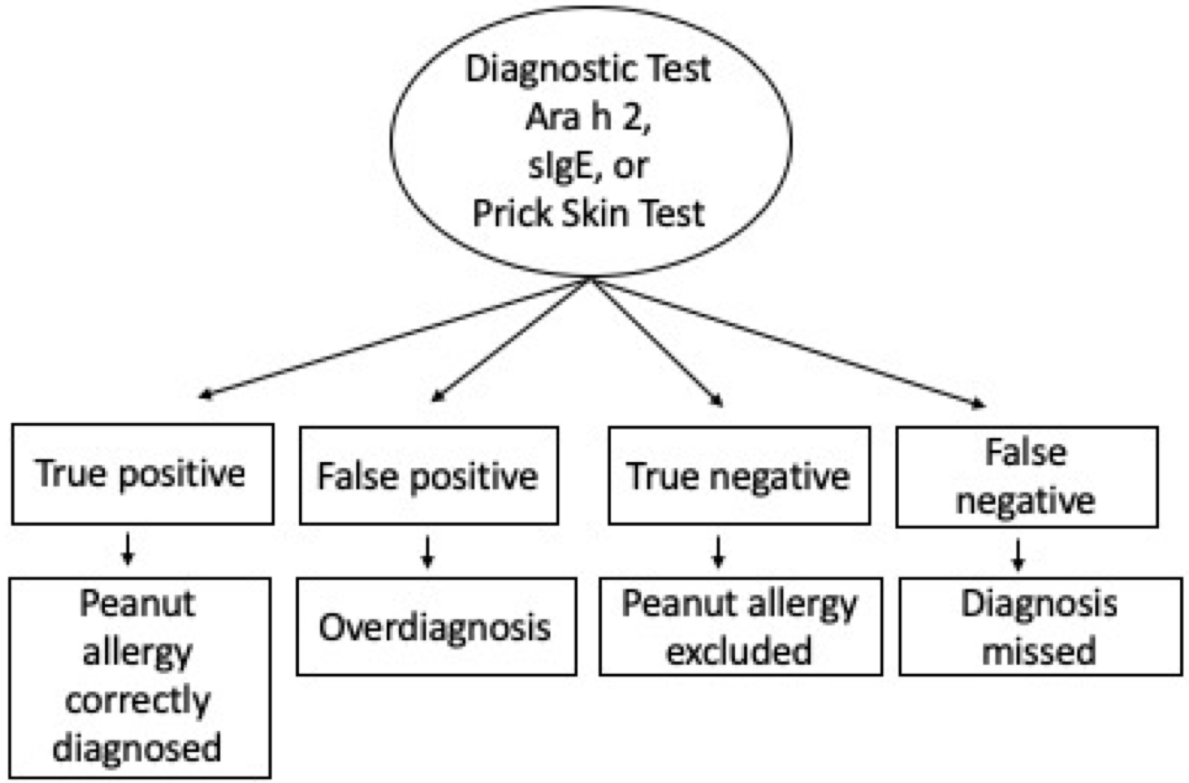


Figure 11: Decision Model for Assessing the Cost Effectiveness of the Use of Diagnostic Testing

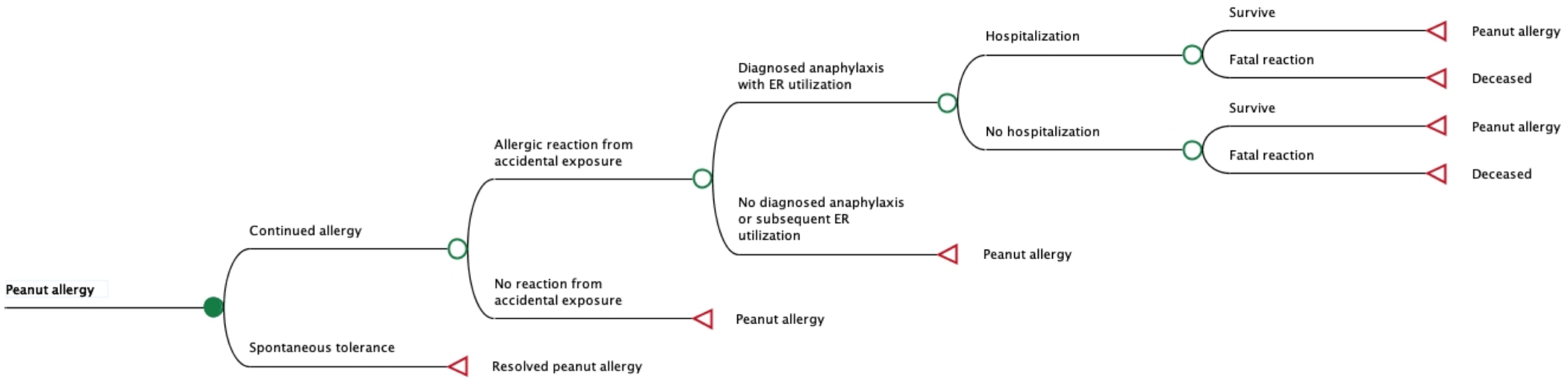


Figure 12: Deterministic Sensitivity Analysis of the Threshold of Ara h 2 Specificity Where Stand-Alone Use Is Cost-Effective

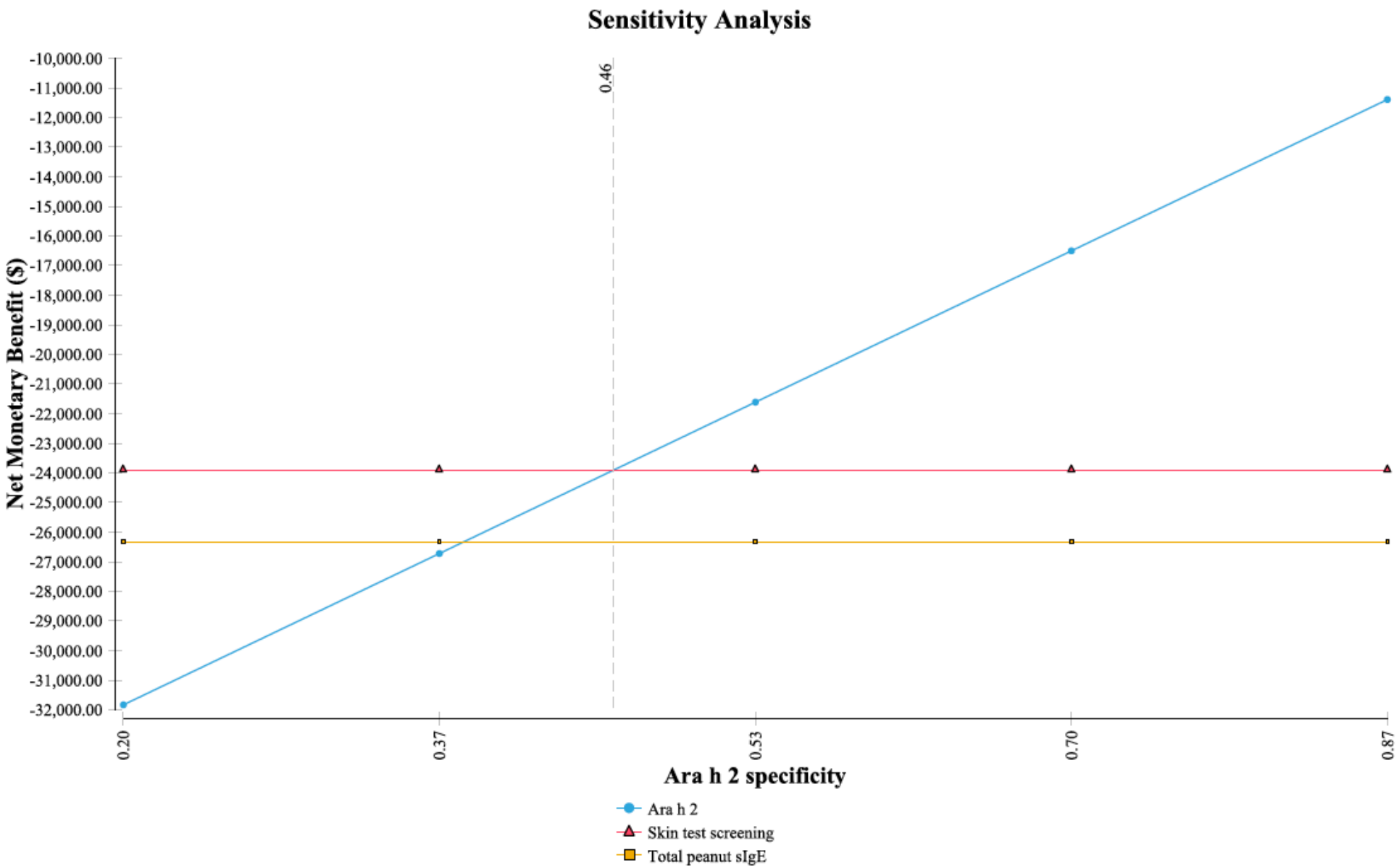
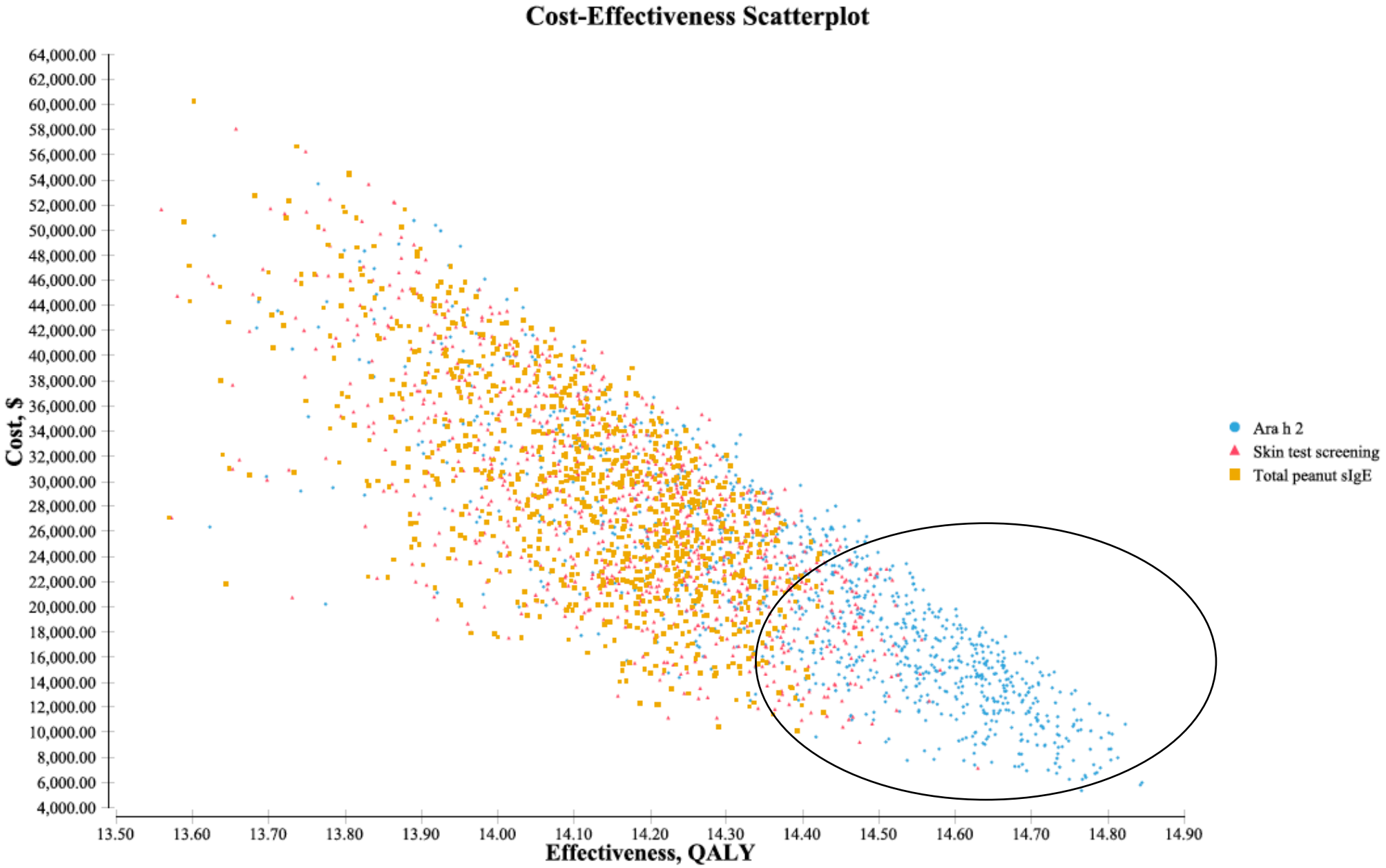
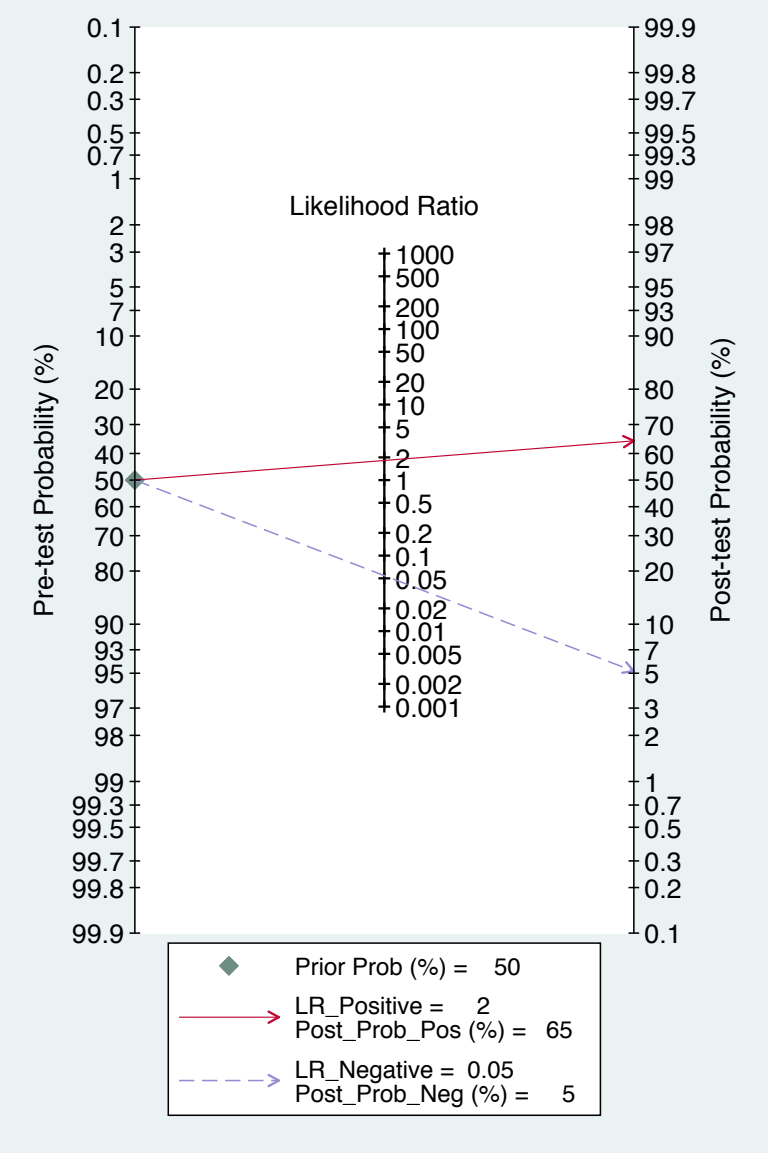


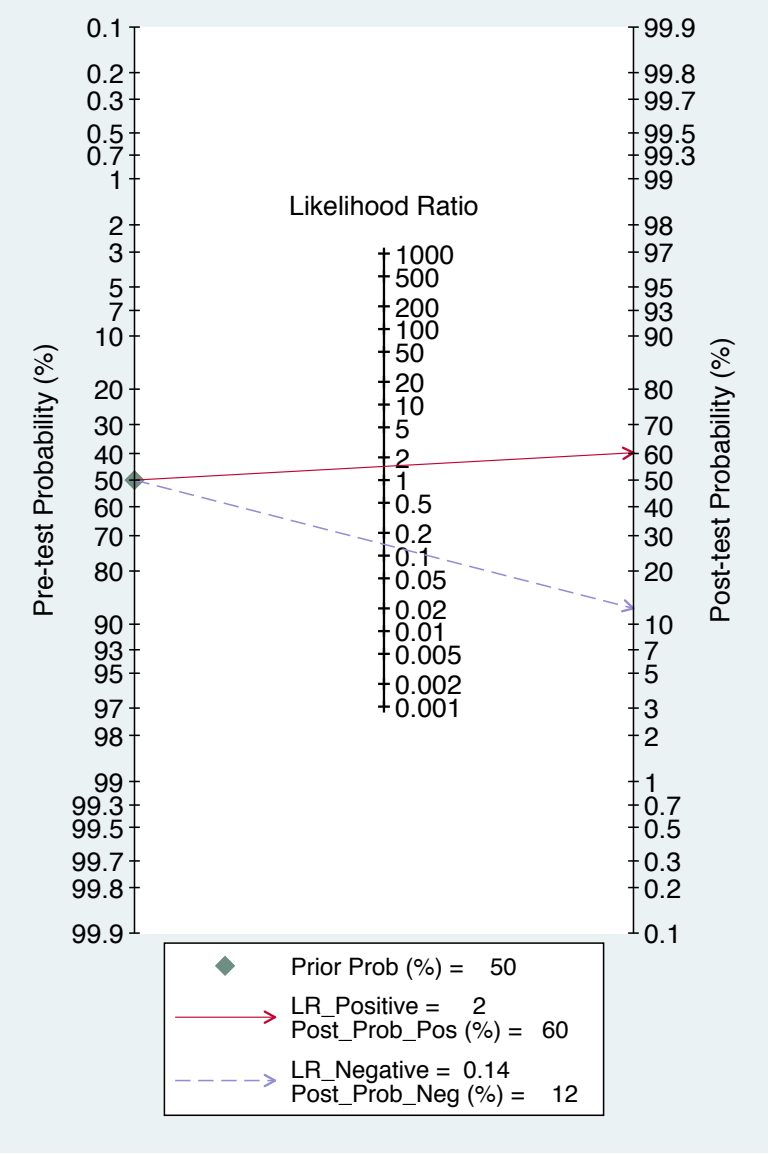
Figure 13: Probabilistic Sensitivity Analysis of Stand Alone Ara h 2 Use



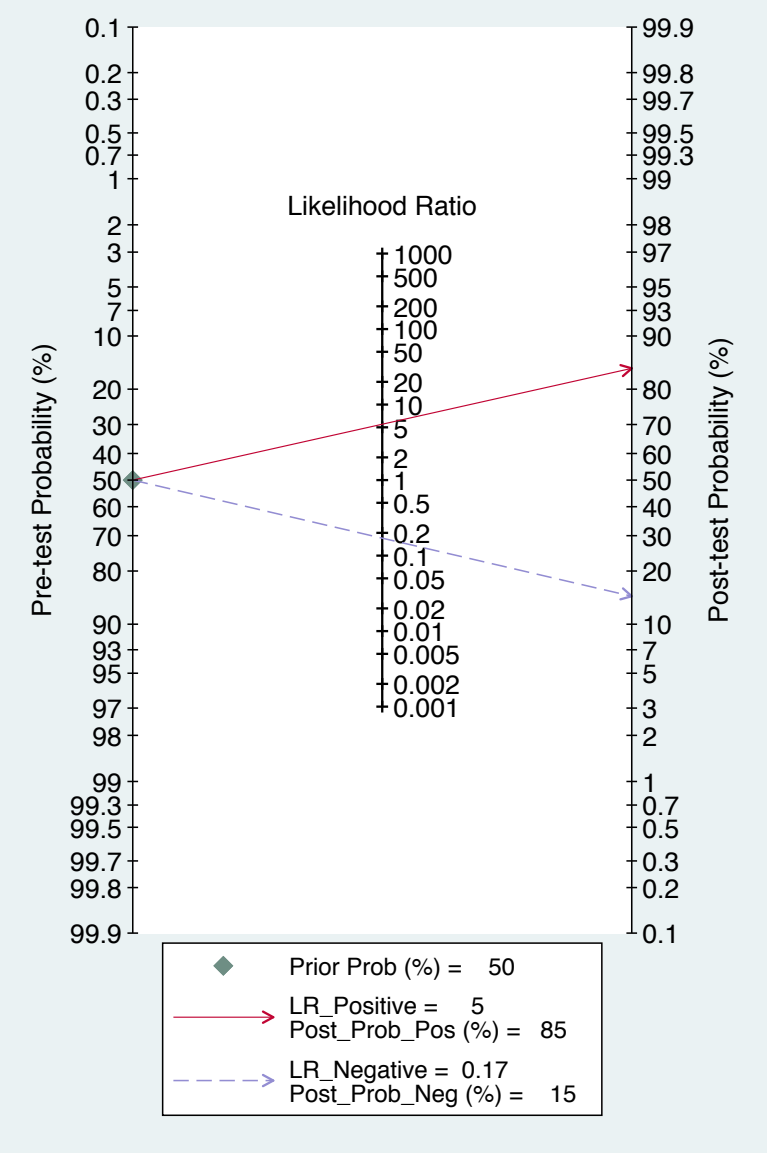
Supplemental Figure 1: SPT 3mm, sIgE 0.35 KU_A/L, and Ara h 2 sIgE 0.35KU_A/L Performance at a 50% Pre-Test Probability



SPT



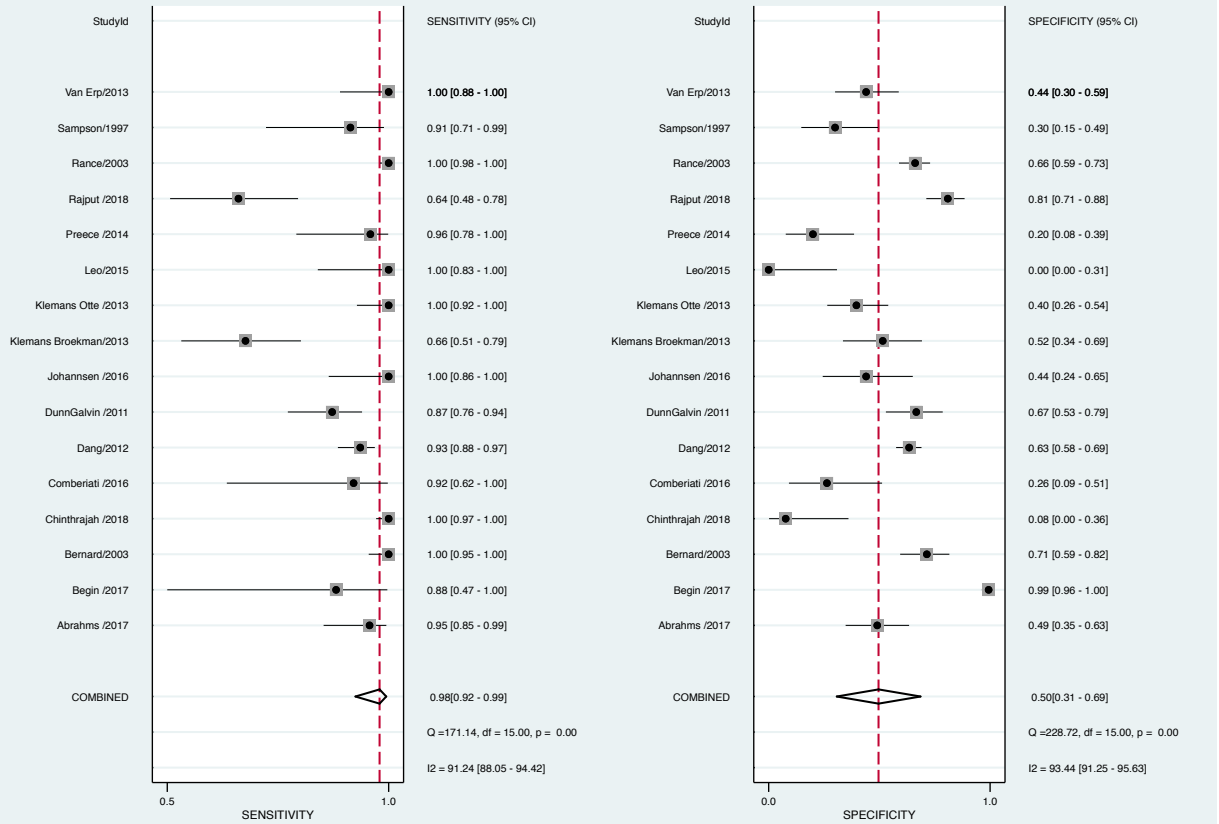
sIgE



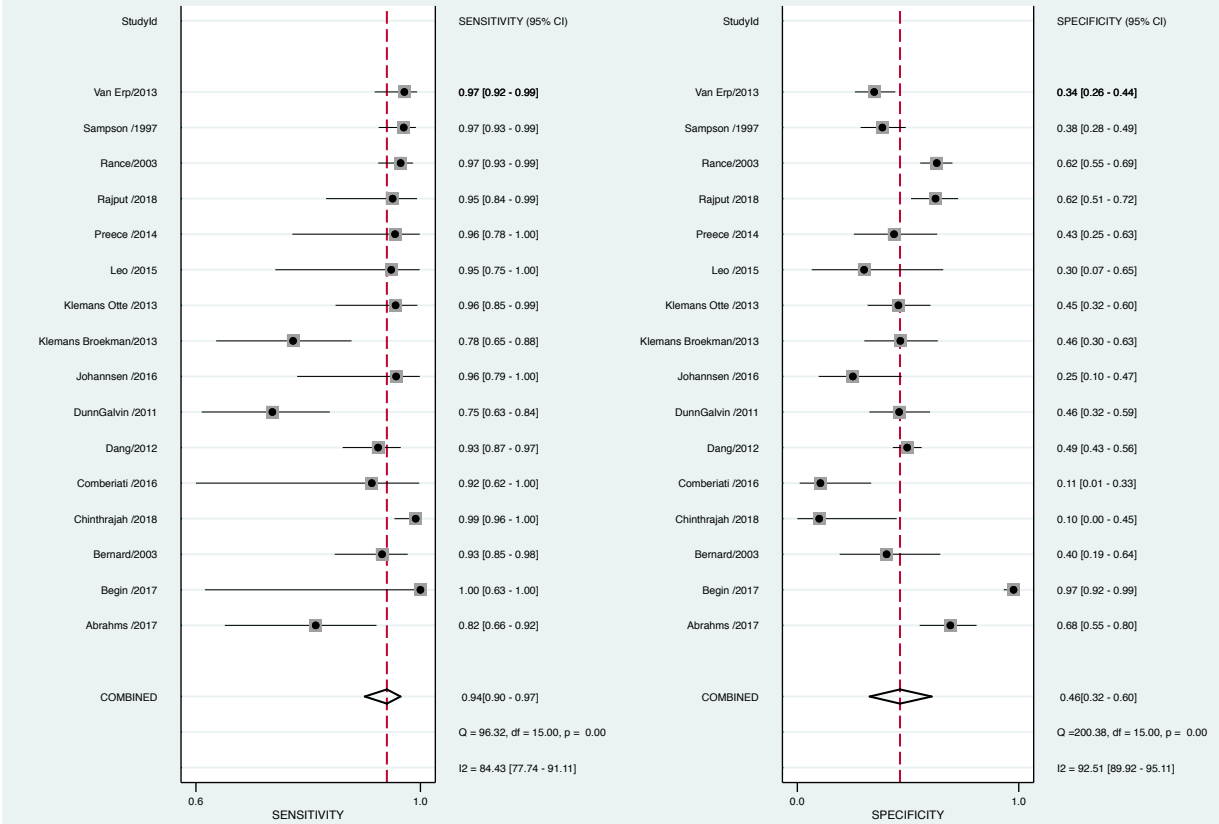
Ara h 2

Supplemental Figure 2: Summary Forest Plots for Sensitivity and Specificity of Skin Prick Testing at 3mm and sIgE testing at 0.35KU_A/L When Both Tests Run Simultaneously

A

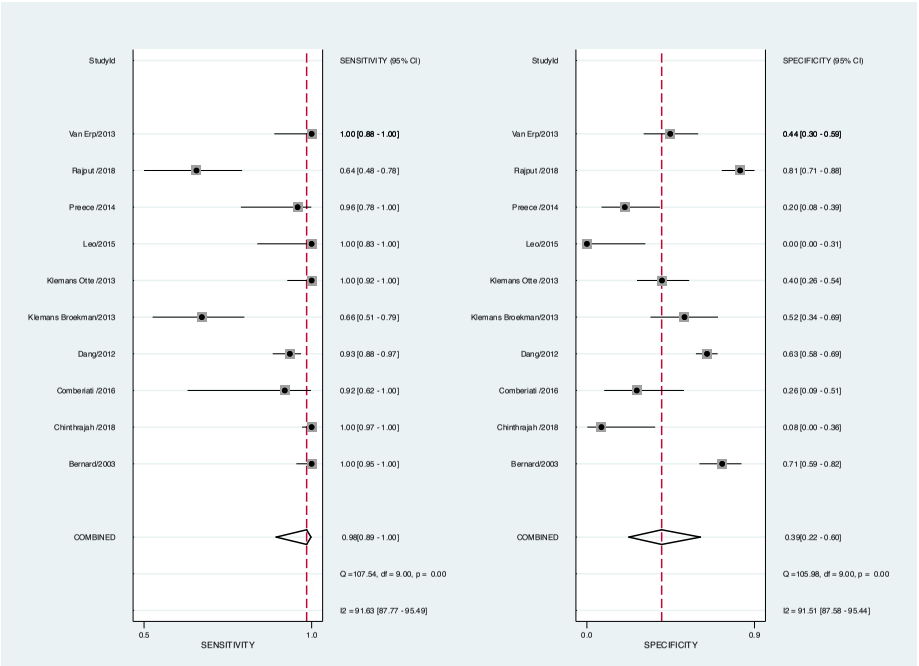


B

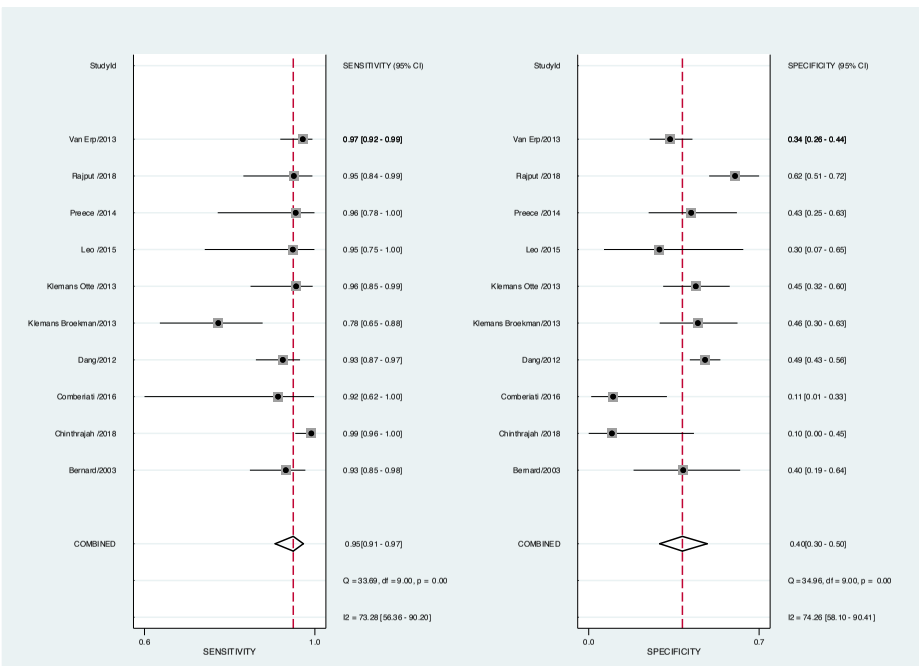


Supplemental Figure 3: Summary Forest Plots for Sensitivity and Specificity of Skin Prick Testing at 3mm, sIgE testing at 0.35KU_A/L and Ara h 2 sIgE .35KU_A/L When All Tests Run Simultaneously

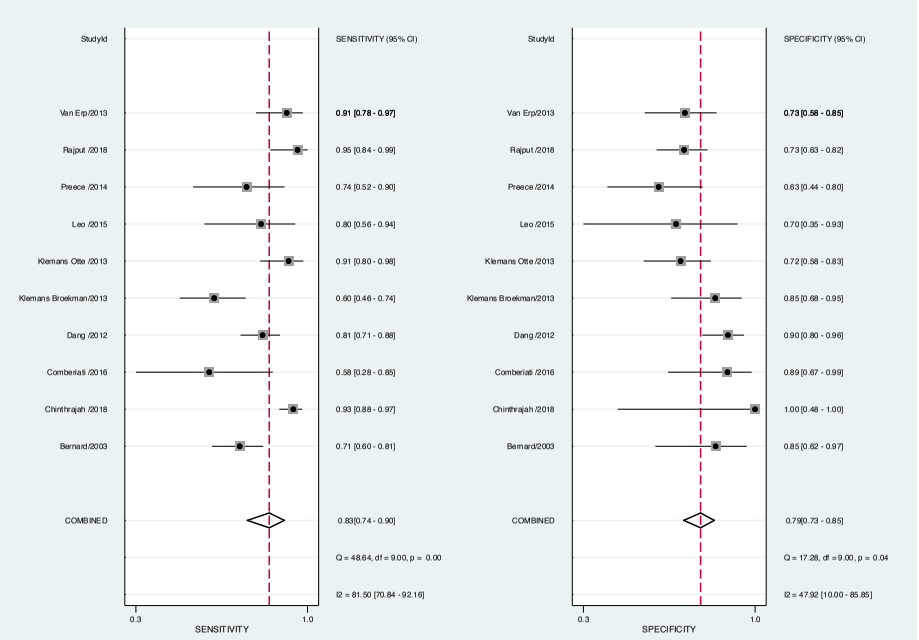
A



B

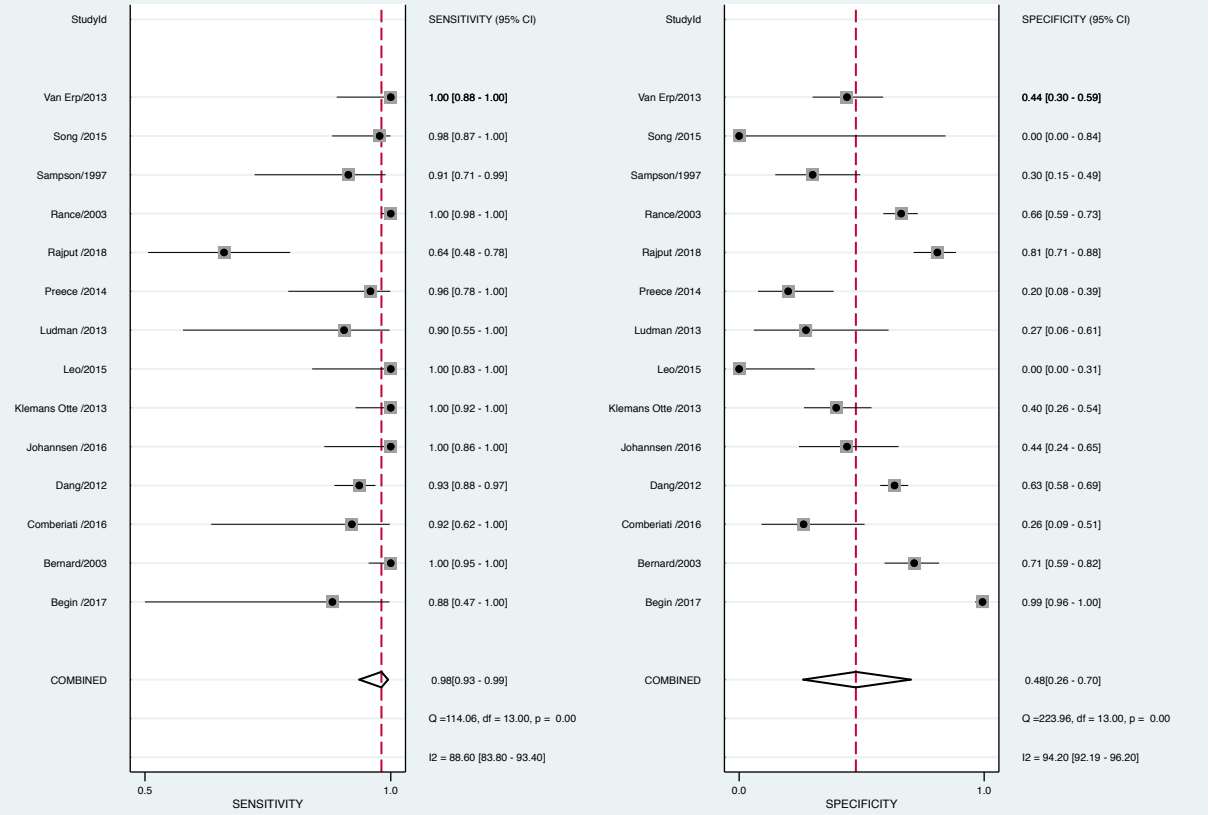


C

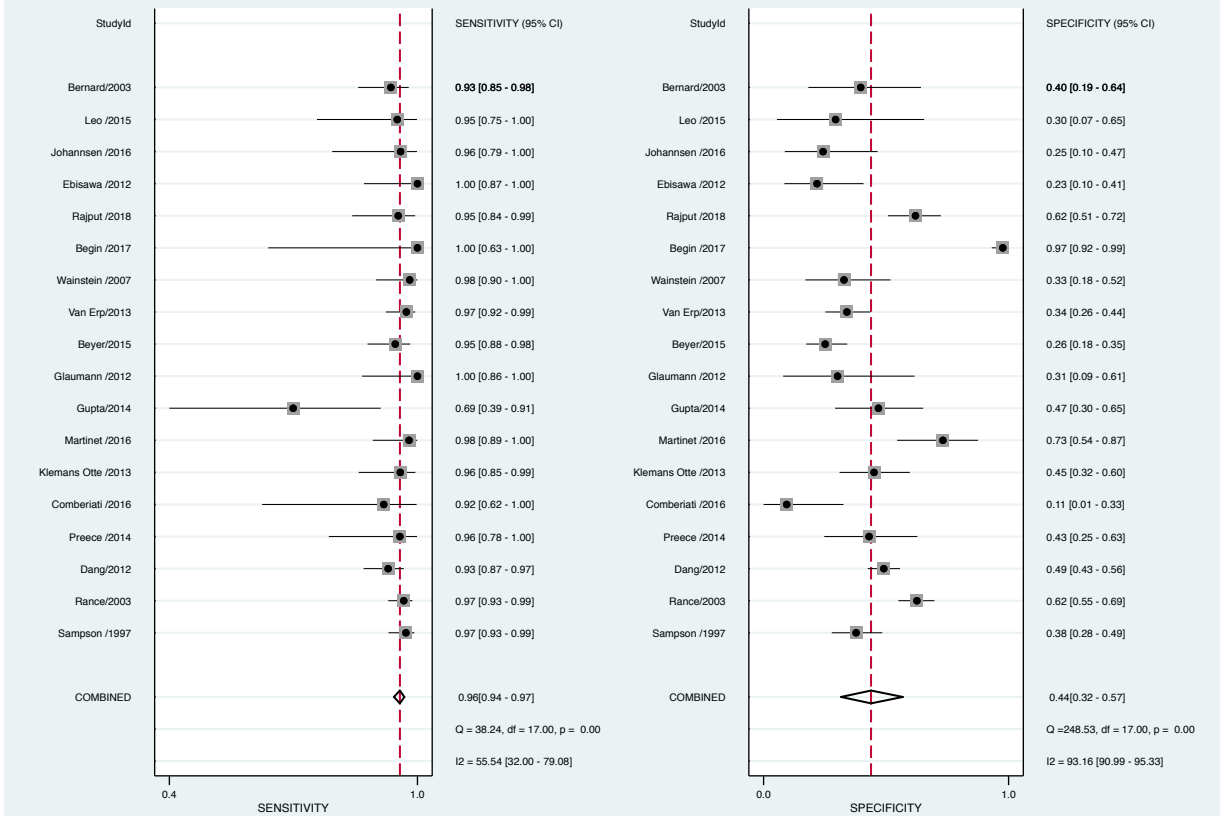


Supplementary Figure 4: Summary Forest Plots for Risk of Bias Removed Sensitivity and Specificity of Skin Prick Testing at 3mm and sIgE testing at 0.35KU_A/L

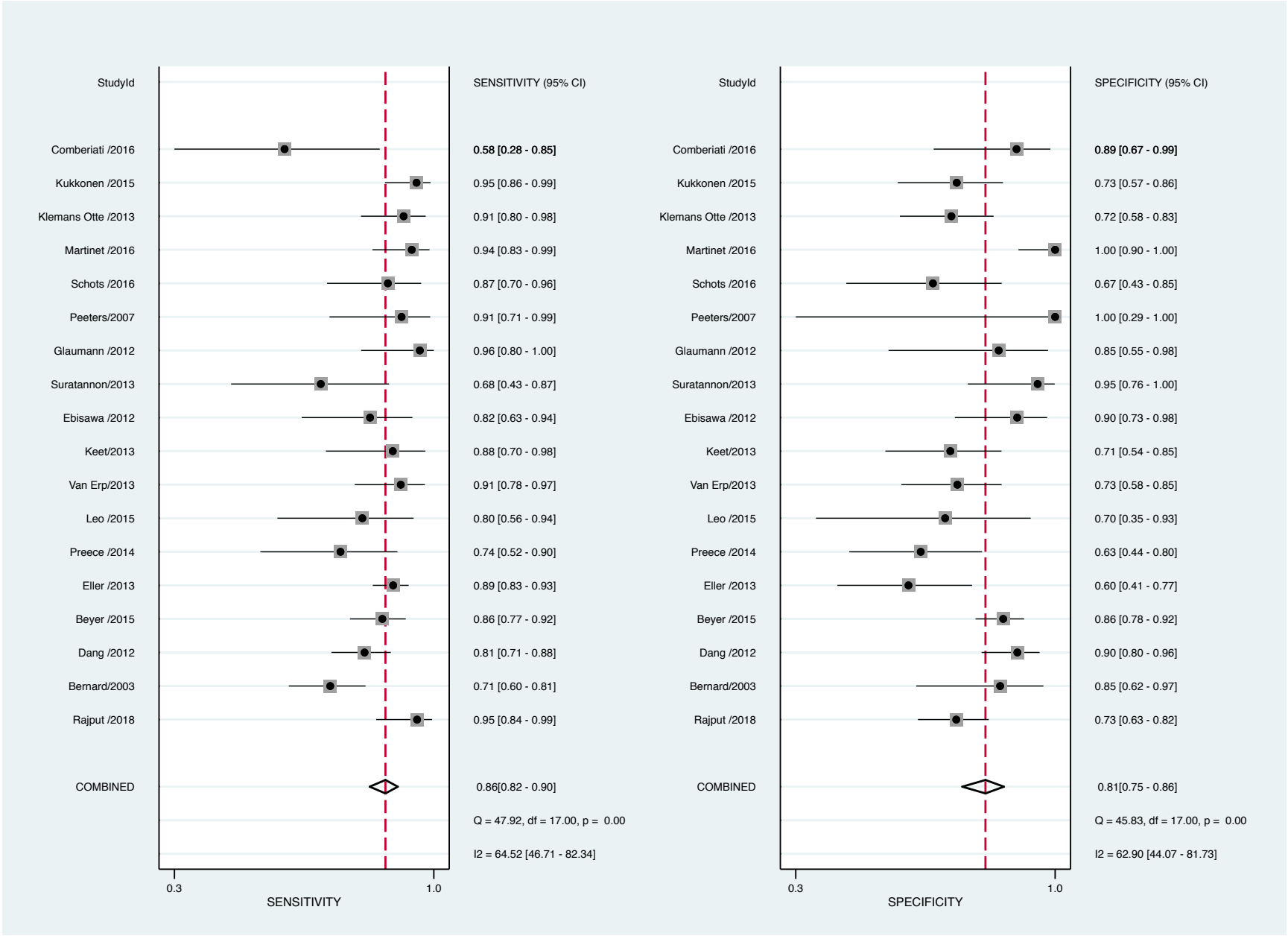
a



b



Supplementary Figure 5: Summary Forest Plots for Risk of Bias Removed Sensitivity and Specificity of Ara h 2 sIgE testing at 0.35KU_A/L



Author	Year	Design	Methods	Population	Findings	Region	n	tp	fp	tn	fn	Population	OFC type	Eczema	Asthma	Sensitized
Abrams	2017	Retrospective chart review of patients with food challenges at a tertiary care pediatric allergy clinic from 2008-2010	Open challenges performed on clinical discretion of attending physician	Children at a Canadian referral center	SPT and whole peanut sIgE were higher in those failing changes than those passing challenges	North America	96	32	18	39	7	Pediatric	Open	65%	55%	52%
Balmer Weber	2015	EuroPrevall cross-sectional study of PA patients	DBPCFC	Adults and children recruited from allergy clinics in Bulgaria, Czech Republic, France, Greece, Iceland, Italy, the Netherlands, Poland, Lithuania, Spain, Switzerland, and UK	Ara h 2 sensitization >=1.0 was associated with a 97% probability of systemic reaction	Australia	95	54	14	12	15	Pediatric	DBPCFC	NR	NR	72%
Begin	2017	Prospective cohort evaluating siblings of peanut-allergic children	DBPCFC	Children at a Canadian referral center	Negative predictive value of skin testing and sIgE was high (99%-100%) with lower positive predictive values (62-88%)	North America	155	7	1	146	1	Pediatric	DBPCFC	38%	11.70%	5%
Beigelman	2012	Retrospective cohort	Open OFC's	Children at a US referral center	85% of patients with negative OFC's had a positive skin test (vs 95% of those with positive OFCs)	North America	198	49	47	25	77	Pediatric	Open	43%	43%	48%
Bernard	2003	Retrospective cohort	OFC's performed to diagnose peanut allergy in children without regular peanut ingestion; patient with severe anaphylaxis excluded	Chidren <10 at a French referral center	Trends emerged with higher levels of sensitization occurring in subjects with more severe reactions but absolute cut-offs were not reported	Europe	91	50	3	17	21	Pediatric	Open	90%	65%	58%
Beyer	2015	Prospective multicenter cohort	Open OFC's performed to diagnose children referred for evaluation of peanut allergy	Chidren <10 at a French referral center	An Ara h 2 of 14.4 ku/L was associated with a 90% probability of positive peanut challenge	Europe	210	77	17	103	13	Pediatric	Open	71%	32%	45%
Chinthrajah	2018	Prospective cohort	DBPCFC for subjects with a convincing history of peanut allergy	Children and adults 7 to 55 years of age at a US referral center	Higher whole peanut and component levels were associated with more severe reactions but absolute cut-offs below 10mm or 50ku/L were not identified	North America	135	12	122	1	0	Mixed	DBPCFC	73%	67%	99%
Comberiati	2016	Prospective cohort	Open OFC in children with suspected peanut allergy	Italian children (median age 10 years)	25% of children with whole peanut sIgE > 15 ku/L tolerated oral challenge	Europe	31	7	2	17	5	Pediatric	Open	NR	NR	29%
Dang	2012	Prospective population-based cohort (HealthNuts)	Oral food challenges performed in coordination with SPT, whole peanut sIgE, and Ara h 2 sIgE	Australian infants recruited at 11 to 15 months of age	Ara h 2 sIgE provided higher diagnostic accuracy than whole peanut sIgE	Australia	200	81	7	93	19	Pediatric	Open	42%	NR	44%
DunnGalvin	2011	Prospective cohort	Opn OFC	Children (mean age 8 years) receiving food challenges to evaluate tolerance to a previously reactive food or to test for reactivity to foods not previously consumed	Predictive accuracy of whole peanut sIgE, SPT, and sequential testing with both was 61%, 75%, and 81%	Europe	124	58	19	38	9	Pediatric	Open	NR	NR	62%
Ebisawa	2015	retrospective cohort of consecutive patients	open OFC for 121 (78%) using objective symptoms as stopping criteria; included 44 (27%) where allergy dx'd by history	Children referred to a Japanee referral center	sp/sn, PPV/NPV for Ara h 2 sIgE cutoff 1.2 and 4.0	Asia	165	87	71	7	0	Pediatric	Open	NR	NR	97.60%
Ebisawa	2012	retrospective cohort of consecutive patients	all had open OFC (10gram)	Children referred to a Japanee referral center	ara h 2 sIgE provided higher diagnostic accuracy than whole peanut sIgE; combo Ara h1,2,3 has spec 94%	Asia	59	23	3	28	5	Pediatric	Open	49%	70%	87.70%
Eller	2013	retrospective cohort	165 open OFC and 40 DBPCFC	Children and adults (to age 26) seen at a Dutch referral cener	Ara h 2 sIgE is superior predictor of challenge outcome than any single component or whole peanut sIgE	Europe	205	155	12	18	20	Mixed	Mixed	NR	NR	99.50%
Glaumann	2012	prospective cohort	DBPCFC; blood drawn at time of challenge	Known swnsitized children without history of anaphylaxis at a Sweedish referal center	detectable Ara h 2 is linked to PN allergy	Europe	38	25	9	4	0	Pediatric	DBPCFC	NR	NR	89.50%
Guilloux	2009	cross-sectional study	DBPCFC or labial challenge, control patients not challenged	Patients seen at a French referral center	compared ImmunoCAP and Immulite with OFC outcome	Europe	99	58	22	19	0	Adult	DBPCFC	NR	NR	100%
Gupta	2014	retrospective chart review	open OFC	Pediatric patients seen at both a US referral center and private clinic	sIgE/total IgE ratio is more accurate than sIgE alone for predicting outcomes of OFC performed to confirm development of tolerance	North America	47	18	9	16	4	Pediatric	Open	68%	34%	NR
Johannsen	2016	cross-sectional study	open OFC, SPT day of OFC, sIgE within 6 mo before OFC	Known sensitized childrenwithout history of reaction < 5yrs seen at an Australian referral center	in children <5 yrs with no ingestion hx, SPT <7mm and IgE <2 kU/L identify children most likely tolerant to PN (5% likelihood of failing OFC)	Australia	49	24	14	11	0	Pediatric	Open	NR	NR	100%
Keet	2013	retrospective cohort	open OFC and had serum banked within 2 yrs of OFC	Children with banked serum saeen at a US referral center 61	Ara h2 diagnostic sensitivity 96% and diagnostic specificity of 54% (sensitization defined as >=0.1)	North America	61	23	10	3	25	Pediatric	Open	NR	NR	100%
Klemans Blom	2015	retrospective	DBPCFC	Children and adults seen at a Dutch referral center	higher PN and Ara h2 IgE associated with increased likelihood of reaction at OFC	Europe	221	58	108	14	41	Mixed	DBPCFC	82%	58%	NR
Klemans Broekman	2013	cross-sectional study	DBPCFC	Children and adults seen at a Dutch referral center	Ara h2 has best diagnostic value of all components; Ara h2 correlated with clinical severity	Europe	94	33	6	33	22	Adult	DBPCFC	57%	53%	79.80%
Klemans Liu	2013	retrospective chart review	DBPCFC	Childen, adults and atopic controls seen at a large Dutch referral center	single- and multi-plexed assay, SPTand immunoblot perform equally in both peanut allergic adults and children, with Ara h2 being most often recognized with all techniques. Specific IgE to Ara h 1, 2, and 3 inadults was correlated with severity	Europe	37	0	35	0	2	Mixed	DBPCFC	NR	NR	95.50%
Klemans Otte	2013	retrospective chart review	Open OFC	Children at a Dutch referral center	With use of ara h2, the need for peanut challenges could be reduced by approx 50%	Europe	100	43	15	38	4	Pediatric	Open	79%	NR	100%
Kukkonen	2015	Prospective cohort	DBPCFC	Children 6-18 at a Dutch referral center	SIgE	Europe	102	58	11	30	3	Pediatric	DBPCFC	52%	55%	NR
Leo	2015	Prospective cohort	Open OFC in those without convincing peanut allergy history	Children 2-17 at a Canadian referral center	using Ara h 2 sIgE at a cutoff of 0.75 kU/L to predict the outcome of a challenge was more than twice as predictive as using the combination of SPT ≥ 3 mm and peanut sIgE at a cutoff of 2 kU/L.	North America	21	5	5	5	6	Pediatric	Open	68%	NR	32%

Lewis	2005	prospective	DBPCFC	Children at a referral center	Cluster analysis failed to reveal any association between a particular protein or pattern of proteins (based on presence/absence) andchallenge score	Europe	25	2	10	0	13	Pediatric	DBPCFC	52%	57%	100%
Lieberman	2013	Prospective and retrospective cohort	open and dbpc	Chilren from several large US referral centers	(100%) severe reactions at low doses were successfully diagnosable.	North America	167	85	5	56	21	Pediatric	Open	NR	NR	100%
Ludman	2013	Retrospective chart review cohort	Open OFC	3 to 16	A positive maternal history of allergy and aspecific IgE>5 kU/l were strongly associated with a significantly increased risk of apositive food challenge (OR 3.73; 95% CI 1.31–10.59; p=0.013 and OR 3.35; 95% CI1.23–9.11; p=0.007, respectively	Europe	21	9	8	3	1	Pediatric	Open	63%	46%	53%
Martinet	2016	retrospective	Open OFC	7.7 ± 4.4	The Ara h 2 sIgE assay has the best negative predictive value (0.93) and positive predictive value (1) at a cutoff of 0.1 kU/l. Ara h 2 sIgE titers can predict the risk of anaphylaxis (14 kU/l, high risk)	Europe	83	45	0	35	3	Pediatric	Open	NR	13.60%	NR
Nicolaou	2011	Population based cohort	Open OFC	Children ages 7 to 14 in a Manchester, UK population cohort	Among school-aged children in the United Kingdom, a cutoff of 0.35 kUA/L Ara h 2 IgE confers 100% sensitivity and 96.1% specificity	Europe	81	28	38	14	1	Pediatric	Open	41%	47%	NR
Peeters	2007	Prospective cohort	DBPCFC	Teens and dults seen at a referral center	Demonstrated the relevance of SPT with diluted purified peanut allergens, showing that the reactivity to all four allergens tested is correlated to the severity of peanut allergy by history.	Europe	29	4	0	11	14	Adult	DBPCFC	80%	50%	NR
Perry	2004	Retrospective chart review cohort	Review of patients in database who underwent open OFC were performed to confirm less of allergy when peanut-specific IgE level < 0.35 kUA/L or approached one fourth of the previously established 95% PPV. Included patients given diagnoses solely on the basis of positive skin test response or food-specific IgE levels, and others had a less clear history of reaction, such as a worsening of atopic dermatitis with exposure to that food.	Children and 8 adults seen at a large US referral center.	n=173 challenges performed, 59% passed, median pass level 0.5 KU/L, median fail 1.9 KU/L. 76% of patients passed OFC at ps-IgE<= 0.35 kUA/L, 44% passed between 0.36 and 2 kUA/L, 40% passed between 2 and 4.9 kUA/L, and none passed >5 kUA/ L. For those patients without a clear reaction history, 88% of patients passed with a negative peanut-specific IgE level of less than 0.35 kUA/L, 71% passed with a level of 0.36 to 2 kUA/L, 33% passed with a level of 2 to 4.9 kUA/ L, and 77% passed with a level greater than 5 kUA/L	North America	159	60	54	44	1	Pediatric	Open	58%	48%	68.2
Preece	2014	Prospective cohort,consecutive patients	53 patients consecutively recruited for open OFC at an Australian referral center, inclusive of 32 patients with prior anaphylaxis. Patients excluded with PST>10mm. SPT, sIgE, and Ara h 2 measured.Health Nuts OFC stopping criteria used, blinded assesor used.	Children seen at a large Australian referral center	Ara h 2 had higher sensitivity and specificity than SPT or sIgE but did not discriminate patients with or withour anaphylaxis	Australia	53	22	24	6	1	Pediatric	Open	21%	32%	87%
Rajput	2018	Retrospective cohort	Open peanut OFC in sensitized patients that had not reacted to peanut since early childhood, with non-anxious families.	Pediatric patients at a large UK referral center	Ara h 2 was a better predictor of OFC outcome than SPT or sIgE in a norhtern England population	Europe	31	2	6	5	18	Pediatric	Open	NR	NR	100%
Schots	2016	Retrospective cohort	Peanut sensitized chldren undergoing open OFC	Pediatric patients at a Dutch referral center	Ara h 2 had best discriminatory value in predicting challenge outcome	Europe	52	27	7	14	4	Pediatric	Open	82%	69%	100%
Song	2015	Nested prospective cohort	DBPCFC	Adult and adolescent cohort seen at a large US referral center, undergoing DBPCFCs as part of screening for enrollment in a clinical trial for Chinese Hearbal Medicine	A low positive correlation was seen between DBPCFC severity score and Ara h 2 IgE,whereas a low negative correlation with Ara h 8 IgE was observed.	North America	44	41	2	0	1	Mixed	DBPCFC	46%	59%	100%
Suratannon	2013	Cross-sectional cohort	Peanut sensitized subjects at a referral center underwent SPT, sIgE, and component testing as well as open OFC	Children and adults at a Singapore referral center	rAra h 2, rAra h 9, and CCD are important components in the diagnosis of peanut allergy in an Asian country with low peanut allergy prevalence. The ratio between rArah h 2 sIgE to peanut sIgE can be used for predicting patients who will develop anaphylaxis.	Asia	40	16	7	14	3	Pediatric	Open	NR	NR	100%
Van Erp	2013	Retrospctive cohort	n=225 childre	Dutch referral center, pediatric population with known/suspected peanut allergy, referred for DBPCFC in a 3 year period, all having SPT, sIgE, Ara h 3 measurements.DBPCFC severity rated using Sampson criteria.	No marker of sensitization was linked to severity of reaction	Europe	80	28	30	22	0	Pediatric	DBPCFC	82%	45%	100%
Wainstein	2007	Prospective cohort	Known peanut sensitized children invited for open OFC to validate previously established predictive cut-off levels	Pediatric patients at a large Australian referral center	Using challenge outcomes as the standard, available in vitro and in vivo diagnostic tests for peanut allergy have poor sensitivity and specificity and combining them does not significantly improve their clinical usefulness. Previously described diagnostic cut-off levels do not have general applicability.	Australia	85	51	22	11	1	Pediatric	Open	71%	32%	100%
Wainstein	2010	Prospective cohort	Known peanut sensitized pediatric patients undergoingopen OFC to determine predictive markers of reaction severity	Pediatric patients at a large Australian referral center	Mean peanut SPT wheal size and specific IgE level were associated with the severity of reactions on challenge. History of anaphylaxis prior to the challenge was not predictive.	Australia	54	7	0	13	34	Pediatric	Open	65%	33%	100%
Wensing	2002	Retrospctive cohort	Known sensized patients undergoing DBPCFC to establish peanut allergy threshold dose	Adult patients at a Dutch referral center	No Observed Adverse Effect Level of 30mcg protein established	Europe	26	4	13	2	7	Pediatric	DBPCFC	NR	NR	100%