

Standardizing double-blind, placebo-controlled oral food challenges: American Academy of Allergy, Asthma & Immunology–European Academy of Allergy and Clinical Immunology PRACTALL consensus report

Hugh A. Sampson, MD,^a Roy Gerth van Wijk, MD,^b Carsten Bindslev-Jensen, MD, PhD,^c Scott Sicherer, MD,^a Suzanne S. Teuber, MD,^d A. Wesley Burks, MD,^e Anthony E. J. Dubois, MD,^f Kirsten Beyer, MD,^g Philippe A. Eigenmann, MD,^h Jonathan M. Spergel, MD, PhD,ⁱ Thomas Werfel, MD,^j and Vernon M. Chinchilli, PhD^k
New York, NY, Rotterdam and Groningen, The Netherlands, Odense, Denmark, Davis, Calif, Chapel Hill, NC, Berlin and Hannover, Germany, Geneva, Switzerland, and Philadelphia and Hershey, Pa

Key words: Food allergy, oral food challenge, double-blind, placebo-controlled food challenge

Discuss this article on the JACI Journal Club blog: www.jaci-online.blogspot.com.

In an article reviewing the status of gastrointestinal allergy in the *New England Journal of Medicine* in 1949, Ingelfinger et al¹ decried the reliance on patients' "incrimination" of specific foods, outcome of trial diets, or association of abdominal complaints with symptoms believed to be allergic in making the diagnosis of food allergy. These authors did not doubt that food allergy existed but believed that the criteria used to diagnose food allergy were unacceptable. They proposed the following criteria: (1) food should be given in capsules, by stomach tube, or in such a manner that the patient is unaware of its nature; (2) reproducible symptoms should consistently follow administration of the disguised

Abbreviations used

AD:	Atopic dermatitis
APT:	Atopy patch test
DBPCFC:	Double-blind, placebo-controlled food challenge
LOAEL:	Lowest observable adverse event level
ML:	Maximum likelihood
NO:	Nitric oxide
NOAEL:	No observable adverse event level
OFC:	Oral food challenge
SPT:	Skin prick test

food at a more or less constant interval; (3) other foods given to the patient in the same manner should not produce similar changes; and (4) suspected foods given in the same manner to nonallergic healthy subjects should not cause the observed

From ^aMount Sinai School of Medicine, New York; ^bErasmus Medical Center, Rotterdam; ^cOdense University Hospital, Odense; ^dthe University California-Davis School of Medicine, Davis; ^ethe University North Carolina School of Medicine, Chapel Hill; ^fthe University of Groningen, University Medical Center Groningen, Department of Pediatric Pulmonology and Pediatric Allergy, and the GRIAC Research Institute, Groningen; ^gCharité University Hospital, Clinic for Pediatric Pneumology and Immunology, Berlin; ^hChildren's University Hospital of Geneva; ⁱChildren's Hospital of Philadelphia and Perelman School of Medicine, University of Pennsylvania, Philadelphia; ^jHannover Medical School; and ^kPennsylvania State University College of Medicine, Hershey.

Disclosure of potential conflict of interest: H. A. Sampson is on the Danone Scientific Advisory Board; has received consultancy fees from Allertein Therapeutics and the Food Allergy Initiative (FAI); is employed by Mount Sinai Medical School; has received research support from the National Institute of Allergy and Infectious Diseases (NIAID) and the National Institutes of Health (NIH); has received royalties from Elsevier-Wiley and UpToDate; has received travel expenses from the NIAID; and is 42.5% shareholder of Herbs Springs, LLC. R. Gerth van Wijk has received travel support from the European Academy of Allergy and Clinical Immunology (EAACI); is on the board for the World Allergy Organization and EAACI; has received consultancy fees from MSD, Crucell, and Haarlems Allergenen Lab; is employed by Erasmus Medical Center; has received research support from Zon-MW, STW, and Biomay; has received lecture fees from Allergopharma; and receives royalties from Handboek Allergologie, Het Allergie Formulatius, and Het Allergie Boek. C. Bindslev-Jensen has received lecture fees from MSD and Thermo-Fisher. S. Sicherer has received consultancy fees from the FAI. S. S. Teuber has consulted for Sanofi-Aventis. A. W. Burks has received grants and support for travel from the NIH/NIAID; is on boards for the NIH HAI, US Food and Drug Administration (FDA), and American Academy of Allergy, Asthma & Immunology (AAAAI); has consulted for Dannon Company, Merck & Co, Exploramed Development Intelliject, McNeil Nutritionals, Novartis

Pharma AG, and Schering-Plough; has received grants from the Food Allergy & Anaphylaxis Network, the FAI, the NIH, and Wallace Research Foundation; has received royalties from UpToDate; has received payment for educational presentations from Current Views; and owns stock or stock options in Allertein and MastCell Pharmaceuticals. A. Dubois has received travel and hotel costs from the EAACI, is on the ALK-Abelló Adrenaline Board, and has received research support from ALK-Abelló. K. Beyer has received consultancy fees from Danone, Novartis, and ALK-Abelló; has received research support from the German Research Foundation, European Union, Danone, and Foundation of the Treatment of Peanut Allergy; and has received lecture fees from Danone, Phadia, Infectopharm, CSL Behring, Novartis, UCB, Meda Pharma, and Med Update. P. A. Eigenmann has received travel support from the EAACI; has consulted for DBV Technologies, Danone, Nestlé, and ALK-Abelló; has received grants from LETI and Nestlé; has received payment for lectures from Stallergenes, Sodilac, and Phadia; has received royalties from UpToDate and Elsevier; and owns stock or stock options in DBV Technologies. J. M. Spergel is on the advisory board for and has stock/stock options in DBV; has received consultancy fees from Danone; has received research support from the Department of Defense, Ception, NIH, and Nutricia; has received lecture fees from Abbott and Nutricia; receives royalties from UpToDate; and has received payment for development of educational presentations from MEI. T. Werfel has received travel support from the EAACI. V. M. Chinchilli has received travel support from the AAAAI.

Received for publication August 22, 2012; revised October 11, 2012; accepted for publication October 11, 2012.

Corresponding author: Hugh A. Sampson, MD, Department of Pediatrics, Box 1198, Mount Sinai School of Medicine, One Gustave L. Levy Place, New York, NY 10029-6574. E-mail: hugh.sampson@mssm.edu.
0091-6749/\$36.00

© 2012 American Academy of Allergy, Asthma & Immunology
<http://dx.doi.org/10.1016/j.jaci.2012.10.017>

symptoms. One year later, Mary Loveless, a New York allergist, published 3 articles outlining the use of blind, placebo-controlled oral food challenges to demonstrate that the practice of the day for diagnosing food allergy (ie, elimination diets and open feedings) inadequately accounted for confounding factors.²⁻⁴ Dr Loveless called for “the introduction of controlled, objective methods to the study of food allergy,” but her appeal went largely unnoticed for more than 10 years until A. S. Goldman and his colleagues published a series of articles describing their evaluation of children with milk allergy. In this series of 89 children, the authors accepted a diagnosis of milk allergy when the children met the following criteria⁵⁻⁷: (1) the patient had to experience resolution of symptoms when milk was eliminated from the diet and (2) presenting symptoms had to be replicated on 3 occasions when milk was added back to the diet, preferably in a blinded fashion. However, it was not until 1976, when Charles May published his seminal article,⁸ that the allergy community began to accept the need for double-blind, placebo-controlled food challenges (DBPCFCs), which became known as the gold standard, to accurately diagnose food allergy. In his article May stressed that a negative challenge result had to be followed up with an open feeding of the challenged food prepared in the usual way and given in a normal proportion at a later time point before a patient could be defined as nonallergic. This was to ensure that the use of dehydrated powdered foods, as advocated for use in challenges by May, did not alter the allergenicity of the food and that the use of progressively increasing doses did not lead to the theoretic “rapid desensitization” of the patient, resulting in a “false-negative” challenge. The latter recently was shown to be a real concern by Niggemann et al⁹ when using lower initiating doses, semilogarithmic dose increases, and more prolonged periods between doses, as advocated by some investigators. Overall, 13% of negative incremental challenge results were subsequently found to be positive when given as a cumulative dose compared with 1% to 3% reported by American investigators.¹⁰

A number of expert panels from various organizations have published consensus documents on the appropriate diagnosis of food allergies. Recently, the National Institutes of Health–sponsored expert panel “Guidelines for the diagnosis and management of food allergy,” reaffirmed the utility of the DBPCFC for diagnosing food allergy after an extensive review of the current literature.¹¹ However, the expert panel noted that open or single-blind challenges could also be acceptable when the challenge outcome is negative or when objective symptoms are elicited that exactly recapitulate the history of the reaction. This is in agreement with an algorithm proposed by Niggemann and Beyer,¹² as shown in Fig 1. Although most authorities agree that the DBPCFC is the gold standard, there is yet to be universal standardization of the challenge procedure and its interpretation. Both the European Academy of Allergology and Clinical Immunology¹³ and the American Academy of Asthma, Allergy & Immunology¹⁴ have published opinion papers on the standardization of oral food challenges (OFCs).

In October 2008, a number of food allergy experts from the European Academy of Allergology and Clinical Immunology and the American Academy of Asthma, Allergy & Immunology met to develop a consensus document on the conduct and interpretation of the DBPCFC. Members of both academies realized the need to develop an international standard that would advance the field by facilitating our ability to compare studies involving diagnosis, natural history, and therapeutic trials in food allergy done around the world.

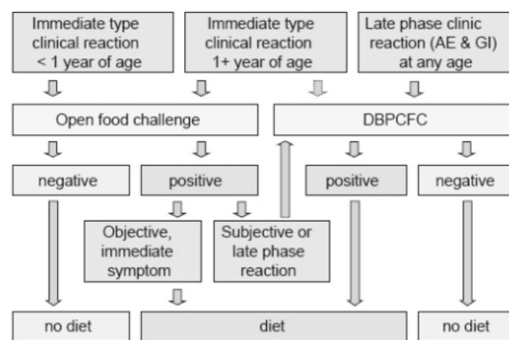


FIG 1. Toward a standardization of food challenges. Reprinted with permission from Niggemann et al.¹²

In the sections that follow, the authors have summarized various aspects of the OFC that were considered. Although this consensus document was meant to provide a standard approach for researchers in the field, most sections are equally applicable to the clinician in practice. Safety was considered paramount in these discussions. The risks and benefits of performing food challenges must be considered, including the appropriate setting for conducting the challenge and the expertise of those administering the challenge. Considerable discussion was directed at the appropriate prescreening and preparation of the patient undergoing the DBPCFC; the prechallenge and intra-challenge assessment, including how to deal with subjective symptoms, parameters to follow, and what objective symptoms constitute a positive challenge; and how outcomes should be reported.

With the increasing number of immunotherapeutic trials for food allergy being conducted around the world and the pivotal role that the DBPCFC plays in interpreting the outcomes of these trials, it is imperative that investigators use a standardized approach. Regulatory agencies, such as the European Medical Agency and the US Food and Drug Administration, are increasingly demanding more stringent standards for the conduct of these studies, and therefore it is critically important that the allergy/immunology community provides input into the best practices for generation of these standard operating procedures. It will also be important for the allergy/immunology community to stress the importance of including all age groups in these clinical trials, especially younger children, who are more frequently affected by food allergies.

PRECHALLENGE ASSESSMENT

Before undertaking a DBPCFC, the patient or research subject should be avoiding the food to be challenged and have attained a stable baseline with regard to atopic disease. The length of time required for chronic disease to improve on an elimination diet can vary by disease and individual patient, but at least 2 weeks is suggested. For safety considerations, the clinician should be aware that prolonged elimination of a food to which IgE is detectable might occasionally be associated with acute severe reactions on reintroduction.^{15,16} Challenges are preferably performed on an empty stomach, although age and practical issues might require compromises to allow liquids or light nonfatty snacks. The following additional prechallenge assessments and considerations should be reviewed in deciding on selection of patients/research participants and for decisions regarding proceeding with administering a DBPCFC to increase the safety and accuracy of the DBPCFC:

1. *Atopic disease that might interfere with assessment.* Patients should not be challenged if they are experiencing unstable or

exacerbated atopic disease, such as asthma, atopic dermatitis (AD), urticaria, or allergic rhinitis. Chronic atopic disease symptoms (eg, rhinitis, AD, and asthma) should be stable and defined as controlled/mild as possible before challenge. Challenges should be deferred for acute infections that might interfere with challenge interpretation.

2. *Diseases and conditions that might affect safety.* Patients should not be challenged if they have chronic medical conditions that would pose a significant health threat in the event of anaphylaxis/treatment of anaphylaxis. Examples of such conditions include unstable angina pectoris, cardiac disease or dysrhythmias, severe chronic lung disease, and pregnancy. Pregnancy should be ruled out by testing or based on history (eg, before menarche or after menopause), as appropriate.
3. *Medications that might interfere with assessment or affect safety.* Antihistamines and medications with antihistaminic properties can mask symptoms and should be avoided for a period of 5 half-lives of the specific agent. Inhaled and topical steroids or anti-inflammatory medications, such as calcineurin inhibitors or leukotriene antagonists and β -agonists, used at the lowest doses possible and on an established schedule by using fixed doses of medication to maintain a low and stable baseline of atopic disease can usually be continued because their use is unlikely to significantly influence challenge outcomes and their withdrawal might result in exacerbations, affecting disease management and challenge interpretation.

Patients should not be challenged proximate to treatment with systemic steroids (eg, within 7-14 days) because disease rebound might confound the interpretation of the food challenge result. Prolonged high-dose steroids, omalizumab, or possibly other new drugs to control atopic disease are likely to modify challenge outcomes and should be avoided. Aspirin/nonsteroidal anti-inflammatory drugs, angiotensin-converting enzyme inhibitors, alcohol, and antacids can act as eliciting factors that increase reactivity in susceptible patients.¹⁷ β -Blockers can pose safety concerns if epinephrine is required for treatment.¹⁸

PRECHALLENGE ASSESSMENT: SURROGATE PARAMETERS

Because food challenges are time-consuming and not always without risk to the patient, much focus has been given to other diagnostic procedures aimed at reducing the need for food challenges. Among these are case history, skin prick testing, measurement of specific IgE levels to the food or components in the food in question, and atopy patch testing.

Much effort has been invested in trying to establish standards for the size of skin prick test (SPT) responses, the level of specific IgE, or both to food proteins above which the probability of a positive outcome of a food challenge would exceed 95% (decision point). Decision points have been established for various foods (milk, egg, peanut, and hazelnut) in various patient groups, but thus far, no common values have been accepted.

Case history

Only in very clear-cut cases of an acute severe reaction to a specific single food can the patient's case history substitute for challenge.¹⁹⁻²¹ In most cases the incidence of self-reported adverse reactions to food exceeds the challenge-proved incidence

by up to a factor of 10.^{13,22-25} Furthermore, in cases of delayed reactions to foods, case history is rarely of clinical relevance.²⁶

Skin prick testing

SPTs can be performed either with commercial extracts or with fresh foods.²⁷⁻³⁰ The sensitivity and specificity varies between extracts, and direct comparisons between different reports are therefore difficult. In the most positive reports³¹ the specificity of a positive SPT response compared with challenge outcome is 100%, but this is achieved at the expense of sensitivity because in the studies by Sporik et al,³¹ positive challenge outcomes were seen in patients with negative SPT responses. Normally, the use of fresh extracts has a higher sensitivity than the use of commercial extracts, but the opposite has also been demonstrated.^{28,30} There are major differences between different foods, as demonstrated by Verstege et al,²⁹ who established decision points for milk and egg in a cohort of children but failed to do so with wheat and soy.

In addition, attempts have been made to increase the diagnostic accuracy by using single allergenic proteins, such as casein in milk³²: the sensitivity of SPTs with fresh milk was superior to the individual components, whereas the specificity of SPTs could be improved by skin testing with purified casein in a cohort of children with milk allergy. Another attempt to increase diagnostic performance has recently been published by Tripodi et al³⁰ using titrated SPTs to predict the outcome of egg challenges in children: here SPTs with serial dilutions of the allergen extract significantly increased the diagnostic accuracy up to 99%.

Allergenic cross-reactions (caused by common allergenic epitopes present in various related or unrelated foods and pollens) might play a role in the evaluation of a positive SPT response (or serum IgE measurement), as demonstrated in a cohort of young patients with grass pollen allergy in which a number were also found to have positive results to peanut without clinical relevance.³³ On the other hand, panallergens, such as profilin, were recently suggested to increase diagnostic accuracy by using SPTs with the recombinant protein.³⁴

Allergen-specific IgE

Only in very selected cases, such as the presence of IgE against ω -5 gliadin in patients with food-dependent, exercise-induced anaphylaxis, does the presence of measurable levels of specific IgE to a food establish a diagnosis of food allergy.³⁵ Quantification of ω -5 gliadin in patients with wheat allergy without food-dependent, exercise-induced anaphylaxis has not been found to be so specific.³⁶

Decision points predicting a positive outcome of a food challenge have been investigated in several centers worldwide and provide a useful tool for predicting outcome, especially if the level is high (>20 kU_A/L, as determined by using the CAP system FEIA or UniCAP [Thermo Fisher Scientific, Uppsala, Sweden]).³⁷⁻⁴⁴ Most data are available for egg and milk, followed by peanut and tree nuts, whereas decision points for soy or wheat allergy have not been successfully established. Although the level of specific IgE can sometimes correlate with the severity of a reaction, such as egg and peanut allergy,^{45,46} no correlation between the clinical threshold of the patient during challenge and the specific IgE level has been demonstrated.³⁹

One major point hampering the use of decision points as a surrogate parameter for food allergy is the diversity of the reported levels of specific IgE resulting in a greater than 95%

probability for a positive challenge outcome. Regarding egg allergy in children, published decision points vary between 1.5 kU_A/L,³⁹ 7 kU_A/L or greater,⁴⁷ and 10.2 kU_A/L,⁴⁴ and the decision points can even vary within the same clinical center.^{41,43,44,47} One reason for these discrepancies is the age of the patients studied because Celik-Biligi et al⁴¹ calculated the decision point in infants less than 1 year of age to be 10.9 kU_A/L and that in infants greater than 1 year of age to be 13.2 kU_A/L, giving rise to the overall level of 12.6 kU_A/L.

Decision points should thus be used with caution and adjusted to the actual patient population in a center. Furthermore, most data are established with the Phadia CAP System FEIA or UniCAP system, and data generated with these assays are not readily transferrable to other systems because differences between different commercial test systems have been described.⁴⁸

Another issue is the probability of serologic, clinically insignificant cross-reactivity seen between different allergens, such as in patients with grass pollen allergy with positive specific IgE levels to peanut³³ or in patients with a challenge-proved food allergy to one food and cross-reacting IgE to other foods.^{49,50} This topic might be at least partially resolved in the future, when component-resolved diagnostics measuring IgE directed against allergenic epitopes⁵¹ or single purified allergenic proteins in the food might improve the diagnostic accuracy and eliminate insignificant cross-reactions to pollen allergens. Antibodies toward numerous proteins have been identified,⁵² and data are now emerging on the relative allergenic significance of various proteins in hazelnut,^{53,54} soy,⁵⁵ egg,^{56,57} milk,^{44,57} and peanut.⁵⁸⁻⁶⁰ However, at the current time, component-resolved diagnostics cannot serve as a substitute for OFCs in determining either thresholds or severity.⁴⁴

Other tests

Atopy patch testing might improve the diagnostic accuracy in patients with isolated delayed reactions to food^{61,62} but offers limited benefit over measurement of specific IgE or SPTs in the daily routine. Measurement of specific IgG levels or subclasses thereof have not been demonstrated to predict outcomes of food challenges.^{21,63} The use of basophil leukocyte activation tests as an *in vitro* marker for clinical reactivity has not been thoroughly investigated, and no data exist on the use of these methods as surrogate markers for clinical food allergy.⁶⁴

CONCLUSION

Controlled OFCs still remain the gold standard for the diagnosis of food allergy, but certain levels of specific IgE might reduce the need for challenges. However, the patient population in question, including age and possible cross-reactions, should be evaluated. It is important to always keep in mind that the surrogate parameters measure sensitization and not clinical disease in the patient.

PRECHALLENGE ASSESSMENT: CHALLENGE SETTINGS—SAFETY ISSUES

Food challenges, whether DBPCFCs, single-blind challenges, or open challenges, pose certain safety issues that must be addressed before beginning any challenge. The purpose of the challenge, the type of expected allergic symptoms, and the

patient's previous clinical history of a reaction, among others, are all important factors to consider.^{12-14,65,66}

Purpose of challenge

The purpose of the challenge is important; it might simply be a clinical food challenge to prove or disprove a patient is allergic to a certain food. Challenges can be done for research purposes and again might confirm or refute a patient is allergic to the food or might show the minimal threshold of food needed to cause initial clinical symptoms.^{67,68} The risk of the challenge will be dependent on the patient's history of either a previous likely reaction, a history suggestive of a reaction but without likelihood of a reaction, and a history of no known exposure. In general, only approximately 50% of challenges in patients thought to have food reactions will be positive.^{13,67,69} Other important factors to consider include whether the suspected allergic reaction is likely IgE or non-IgE mediated, the age of the patient, and the food to be used for challenge.

Safety/challenge setting

There are several options for a possible location of the food challenge, including the intensive care unit, a regular hospital room, a hospital clinic room, an outpatient clinic room, a private practice office clinic, or the home of the patient.²⁰ Between 20% and 40% of patients challenged in some series had severe reactions on challenge,^{69,70} often to one of the lower doses. Thus the starting dose and schedule of dose escalation is part of overall safety. For example, Sicherer et al⁶⁹ reported on a series of 196 children who underwent DBPCFCs in whom the results of 26 milk challenges and 22 egg challenges were positive at the first dose of 250 mg and 11% of the first dose reactions were severe. Additionally, certain foods seem to cause more severe reactions on challenge than others.^{14,67} The amount of allergen-specific IgE does not correlate with the severity of reaction on food challenge.^{46,71,72} The likelihood of a late-phase or delayed reaction should be considered before the challenge to decide on the appropriate place and period of observation.^{13,14,65} Other factors to consider include the age of the subject; the presence of other concomitant diseases, such as AD, asthma, or risk factors for cardiovascular disease; whether they have reacted to trace amounts of the food based on clinical history; the criteria used for determining a positive challenge result (objective and subjective symptoms); and whether the patient was on an elimination diet before challenge.^{14,73-75}

Personnel

The personnel involved in a challenge procedure must be specifically trained in the management of acute allergic reactions, and anaphylaxis practice drills should be conducted periodically with the challenge team. Centers for performing food challenges are desirable, especially for challenges conducted for research purposes, so that personnel have adequate experience in conducting food challenges.⁷⁶

Medications

All necessary medication should be readily available if needed during the food challenge. Epinephrine, oxygen, antihistamines (both H₁ and H₂), β-agonists, corticosteroids, and intravenous fluids are all needed for specific types of allergic reactions.^{18,76} All patients with a history of a previous severe reaction who are being challenged should have intravenous access established.

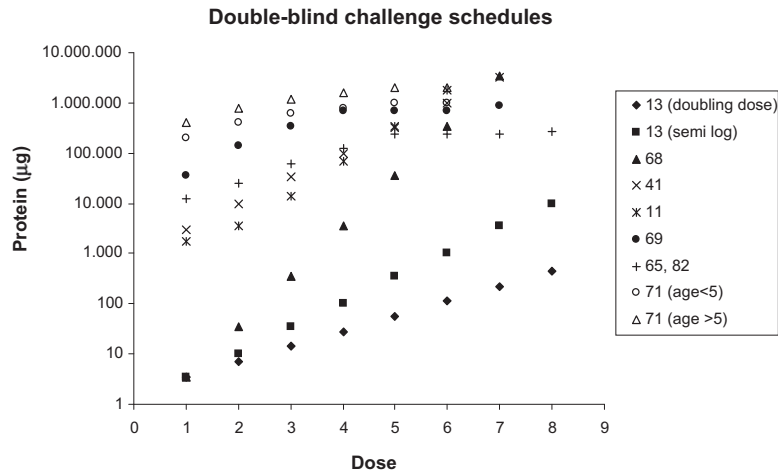


FIG 2. Incremental doses of various schedules used for DBPCFCs. Individual doses are expressed as micrograms of protein in the food administered during the active food challenge. Schedules shown are either generic or derived from those devised for challenges with cow's milk. The numbers shown with symbols refer to the source publications in the reference list.

Other instances in which intravenous access should be considered are food challenges for peanuts, tree nuts, fish, and shellfish, depending on the situation and the patient.^{14,20,69} In small children intravenous access is necessary only in selected patients; if there is any doubt as to the outcome (severity) of the challenge, intravenous access might be needed.^{13,66} Also, patients with severe asthma, even without a history of an anaphylactic reaction to food, and older children and adults with difficult intravenous access should also be considered for prior line insertion.¹⁴

Patients can generally be discharged after an observation period of at least 2 hours, provided no reaction occurred. If the clinical history indicates the allergic symptoms on the initial reaction occurred later, then a longer period of observation might be necessary. If the patient had significant clinical symptoms, then a period of up to 4 hours of observation might be needed before discharge. If a patient experiences a severe systemic allergic reaction requiring significant treatment, then he or she should be kept under observation at the hospital overnight with appropriate treatment.^{13,77}

Conclusions: future directions

Safety issues are an important consideration in deciding who to challenge and where the challenge should be done. A food challenge might be done for various reasons, generally to identify a food that the patient should be avoiding or a food that the patient can put into his or her diet. It is important that OFCs be conducted in a location that is well equipped to deal with anaphylactic reactions by physicians and staff who are well trained in dealing with such reactions. This does not exclude practicing allergists' offices because most allergists administering immunotherapy are fully equipped to deal with anaphylactic reactions. Food challenges for research purposes will continue to allow us to better understand the mechanism of food allergy, as well as to begin to develop new therapies.

CHALLENGE PARAMETERS: SCHEDULES AND MATERIALS

Schedules

Numerous dosage schedules using diverse materials are currently in use for performing DBPCFCs for research purposes.

These schedules differ in starting dose, incremental scale, time between doses, and top dose (see Fig 2).^{13,14,41,65,68,69,71,78} A distinction must be made between dosage schedules designed for studies determining lowest observable adverse event levels (LOAELs) and no observable adverse event levels (NOAELs)^{13,68,79} and those used for other purposes. Similarly, schedules will be different if late-phase reactions, such as those seen in patients with AD, are the object of the study.

Fig 2^{11,13,41,65,68,69,71,82} shows several schedules for DBPCFCs with cow's milk, in which the challenge dose is expressed as protein content of the particular dose. Although schedules for other foods are similar, calculation from protein content to weight of food in either the dried or native form might vary, especially when protein contents of such foods are high (eg, fish) or low (eg, celery).¹³ For these high and low protein-containing foods, other amounts might be necessary to deliver reasonable quantities of food to the patient. Dosages have been derived from either absolute quantities^{13,41,68,78,79} or percentages of either maximum single dose⁶⁵ or cumulative doses.^{65,70} As can be seen from the figure below, starting doses vary widely from the low microgram level for studies aimed at NOAEL/LOAEL determinations to doses from the low milligram level to several hundred milligrams in schedules used for other purposes. Incremental scales vary from 10-fold increases through 5-fold,^{19,68} semilogarithmic (ie, 1, 3, 10, 30, 100, and 300 mg and so on),^{13,41} doubling-dose,⁷⁹ or even smaller^{65,69,71} increases, with the latter being associated with schedules aiming to deliver cumulative doses. These last-mentioned schedules also generally have higher starting doses and shorter time intervals between doses (15 minutes) than the schedules aimed at delivery of discrete doses, where intervals between doses are typically 20 to 30 minutes. One of the pitfalls of using lower starting doses and longer intervals is the increased likelihood of partial desensitization and false-negative results.⁹

With schedules developed for NOAEL/LOAEL determination using starting doses at the low microgram level, it is more difficult to achieve meaningful top doses with acceptable increments in an acceptable period of time. A combination of logarithmic and semilogarithmic increases, as done in EuroPrevall, can be helpful.⁸⁰

Because of differences in reporting safety and validity outcomes resulting from studies using different schedules, robust evidence demonstrating the relative superiority of any of these schedules is limited. Comparing schedules is further hampered by the inability to assess the contribution of a given schedule parameter to a purported difference in outcome. Many parameters have been established on the basis of practical considerations, such as the length of time required to perform the challenge procedure. Despite these limitations, the following recommendations are proposed.

The preferred method for DBPCFCs is the administrations of active and placebo challenges on separate days. If the active and placebo test challenges are administered on the same day, these challenges should be separated by at least 3 hours. However, such schedules will not be capable of diagnosing reactions more than 3 hours after the challenge. Administration of active and placebo doses in an interspersed fashion (ie, administering both placebo and active doses in the same challenge) is not recommended.

Current evidence suggests that high starting doses are associated with more severe reactions during the challenge. Starting doses at the low milligram level are generally safe and seem to result in fewer severe reactions than higher doses.^{70,81} Experience with 10-fold dose increases as being relatively safe over the entire dose range is limited. Many schedules using semilogarithmic increases are associated with good safety. Smaller increments (eg, doubling) are likely to be safe but lengthen the challenge procedure. Longer time intervals between doses will also lengthen the challenge procedure but will probably decrease the chance of accumulating high doses, which might result in more severe reactions. An interval of at least 15 to 20 minutes between doses is recommended.

The top dose required to avoid false-negative DBPCFC results is uncertain but seems to be at least 2 g. Sicherer et al⁶⁹ reported approximately 5% false-negative test results with a schedule with a top dose of 875 mg (cumulative dose, 3500 mg) of protein. Schedules with top doses of approximately 1750 mg (cumulative dose, 2190 mg) report a false-negative rate of approximately 9% and 12.7%,⁸² as gauged by recurrence of symptoms during subsequent introduction of the food. The original recommendation of May⁸ of 8 to 10 g of dried food as being an adequate top dose is also in this range because this would be 2800 to 3500 mg of protein for milk. However, it is possible that further future experience will show that top doses might need to be higher for certain foods (eg, those of which the usual portion contains relatively high amounts of protein, such as fish), patients, or situations. Follow-up of negative DBPCFC results with open food challenges with the native food in amounts usually consumed thus remains recommended if clinical tolerance is to be unambiguously demonstrated.

On the basis of the above recommendations, a general challenge schedule consisting of 3, 10, 30, 100, 300, 1000, and 3000 mg of food protein at intervals of at least 20 minutes (Fig 2) would be suitable. Such a schedule is likely to be appropriate for most foods, patients, clinical situations, and settings. Appropriate challenge doses for common foods are shown in Table I and can also be found in a recent working group report.¹⁴ On the basis of current knowledge, it is likely to be relatively safe while minimizing false-negative test outcomes.

Materials

Materials used in DBPCFCs should be well documented. The form and source of the food to be used should preserve maximum allergenicity. Matrices should be capable of blinding the sensory

qualities of the food in a volume that is as small as possible and compatible with use in the target patient group (eg, children). Blinding should be verified by using appropriate testing in a sensory laboratory.⁷⁸ Current evidence suggests that a high fat content of matrices slows absorption and masks early warning symptoms, leading to more severe reactions.⁸³ A large variety of matrices and matrix food combinations is important for palatability, especially in young children.

PARAMETERS TO FOLLOW DURING ORAL CHALLENGE

The assessment of a food challenge is mostly clinical and will always largely depend on the investigator. However, decisions on the positivity of a food challenge result need to be standardized as much as possible and should, if possible, include tools for objective measurement.

Clinical assessment

Before each dose, the patient should have a clinical examination, including inspection of the skin, lung auscultation, abdominal auscultation, and blood pressure, heart rate, and oxygen saturation measurement. In addition, the patient should be questioned for pruritus (mouth mucosa, skin, or both), laryngeal symptoms, abdominal pain, dizziness, and any other new complaints. Although the physical examination will provide objective symptoms, most complaints arising from the patient without observable changes need to be classified as subjective and, if isolated, might not account for a positive challenge result (see the section on interpretation).

Assessment of respiratory parameters

Respiratory reactions are often difficult to assess early on, and an objective tool that could identify pulmonary changes early is desirable because it might predict the potential severity of the reaction. Respiratory reactions were reported in 59% of patients with a positive food challenge result, mostly as nasal symptoms (63%), laryngeal reactions (43%), or involvement of the lower airways (64%).⁸⁴ Interestingly, significant spirometric changes lagged behind clinical symptoms and were not helpful in the early recognition of pulmonary changes secondary to the food challenge. The same authors hypothesized that patients undergoing food challenges might also present with more subtle reactions, implicating only a change in bronchial reactivity.⁸⁵ They investigated a group of 26 asthmatic children undergoing DBPCFCs by performing methacholine challenges both before and 4 to 6 hours after a DBPCFC. Chest symptoms (cough, wheeze, or both) developed during 12 of 38 DBPCFCs. Of these, 7 subjects had a significant increase in bronchial hyperreactivity in addition to 1 patient who had no apparent chest symptoms. Overall, methacholine challenges are probably too time-consuming for a general assessment of clinical reactivity in research protocols or especially clinical practice involving DBPCFCs, although they might provide an objective measurement of a respiratory reaction, even without a significant change in pulmonary function testing. However, use of methacholine inhalation challenges will not provide a useful early indicator of ongoing pulmonary changes.

The measurement of nitric oxide (NO) has been shown to provide an objective and reliable tool to quantify airway

TABLE I. Equivalent amount of food in milligrams (and milliliter) for the major foods in various forms often used in challenges

Food protein	Pasteurized cow's milk with 3.3% protein content	Skim milk powder with 36% protein content	Pasteurized whisked hen's egg with 12.8% protein content	Hen's egg powder with 47% protein content
3 mg	91 mg \approx 0.1 mL	8.3 mg	23.4 mg	6.4 mg
10 mg	303 mg \approx 0.3 mL	27.8 mg	78.1 mg	21.3 mg
30 mg	909 mg \approx 0.9 mL	83.3 mg	234.4 mg	63.8 mg
100 mg	3,030 mg \approx 3.0 mL	277.8 mg	781.3 mg	212.8 mg
300 mg	9,091 mg \approx 9.1 mL	833.3 mg	2,343.8 mg	638.3 mg
1000 mg	30,303 mg \approx 30.3 mL	2,777.8 mg	7,812.5 mg	2,127.7 mg
3000 mg	90,909 mg \approx 90.9 mL	8,333.3 mg	23,437.5 mg	6,383 mg

inflammation caused by an allergic trigger. Offline NO levels have been measured in 44 infants undergoing open milk challenges before and immediately after the termination of the challenges.⁸⁶ A wide range of exhaled NO values were observed, and no significant changes were observed in children with positive or negative milk challenge results. A more recent study measured NO levels up to 2 hours after the last challenge dose and reported a significant NO level decrease, peaking at 90 minutes after positive challenge results.⁸⁷ Although these procedures might provide an objective measurement for respiratory reactions, the delay in the change of exhaled NO levels will not be helpful for immediate assessment of the positivity of a DBPCFC result.

Assessment of blood parameters

Early studies have linked positive DBPCFC results with increased histamine levels in peripheral blood samples,⁸⁸ leading several investigators to explore changes in various parameters measurable in the blood for the assessment of a food challenge. Niggemman et al⁸⁹ have shown that peripheral blood eosinophil numbers decreased immediately after positive DBPCFC results and that eosinophil cationic protein levels were increasing.⁸⁹ Similarly, other mediators suggestive of mast cell and basophil degranulation have been investigated. Some investigators have reported significant increases in urinary 1-methylhistamine levels 1 hour after a positive OFC result, as well as a significant increase in serum tryptase levels, but the sensitivity of these parameters is low.⁹⁰ Similar results were observed by Ohtsuka et al⁹¹ when measuring plasma tryptase and histamine levels up to 4 hours after a positive challenge result but were not confirmed by others for eosinophil cationic protein level in the blood and tryptase levels in the blood or saliva.⁹²

Other objective measures of clinical reactivity

For immediate availability of objective measurements, Clark et al⁹³ measured the temperatures of 3 areas of the face using facial thermography and showed a significant increase in facial temperature measured on the nose during a positive food challenge result. The same method has been recently used with success to evaluate peanut challenges on the nasal mucosa.⁹⁴

The clinical judgment of an experienced investigator will always be the most important factor in deciding when to call a DBPCFC result positive or negative. However, an objective, easy to perform, and immediately available objective assessment, such as facial thermography or exhaled NO measurement, would

provide a decisive tool that could help to further standardize DBPCFC results.

SCORING AND STOPPING OFCs

To allow comparison of outcomes of DBPCFCs, standards must be followed to report results, including which symptoms are classified as subjective or objective and what outcomes constitute a positive challenge result. There are currently no agreed upon published parameters, likely because clinical judgment is needed, and circumstances might vary by patient or study characteristics. Decisions to discontinue dosing can be made for reasons that vary according to requirements of specific study protocols, the symptoms observed, safety issues, and patients' characteristics. In response to safety concerns, dosing can be ceased before eliciting clear objective reactions, which in turn might reduce the diagnostic accuracy of the procedure. It is beyond the scope of this report to develop a comprehensive approach to conducting OFCs, but guidelines toward improved reporting are presented.

Challenge results are typically considered positive, and dosing is stopped when objective symptoms occur. However, in some situations mild objective symptoms might be considered insufficient to discontinue dosing or to consider a challenge result positive (eg, 1 or 2 transient perioral urticarial lesions from contact with the food or 1 episode of vomiting in a child with anxiety and a distaste for the challenge substance).¹² In some circumstances subjective symptoms can indicate a positive challenge response and present a good reason to cease dosing, such as by having repetitive symptoms or multiple subjective symptoms in several organ systems. However, stopping a challenge for subjective symptoms increases the risk of a false-positive test result compared with only allowing objective symptoms to indicate a positive test result. Subjective or initially mild objective symptoms can be subtle in a young child, who can become suddenly quiet or begin to show behaviors such as food refusal or scratching in the ears or at the tongue and neck.

At a minimum, the parameters for stopping and declaring a challenge result positive or negative should be prespecified in challenge protocols and the details should be reported in publications. Symptoms typically considered subjective are listed in Table II. Regarding the options for discontinuing dosing and considering a challenge result positive, options include "worsening" subjective symptoms, repeated elicitation (eg, on 3 doses), or persistence (eg, 40 minutes).¹²⁻¹⁴ These parameters for stopping a challenge and declaring a challenge result positive have not been evaluated with regard to the effect on the sensitivity and

TABLE I. (Continued)

Peanut butter with 24% protein content	Peanut flour with 50% protein content	Gluten powder with 80% protein content	Soy drink with 3.3% protein content	Soy powder with 50% protein content
12.5 mg	6.0 mg	3.8 mg	91 mg ≈ 0.1 mL	6.0 mg
41.7 mg	20 mg	12.5 mg	303 mg ≈ 0.3 mL	20.0 mg
125.0 mg	60 mg	37.5 mg	909 mg ≈ 0.9 mL	60 mg
416.7 mg	200 mg	125 mg	3,030 mg ≈ 3.0 mL	200 mg
1,250 mg	600 mg	375 mg	9,091 mg ≈ 9.1 mL	600 mg
4,166.7 mg	2,000 mg	1,250 mg	30,303 mg ≈ 30.3 mL	2,000 mg
12,500 mg	6,000 mg	3,750 mg	90,909 mg ≈ 90.9 mL	6,000 mg

specificity of the test. When mild objective or subjective symptoms occur, decisions include stopping the challenge, waiting longer for the next dose, or repeating a dose.^{12,75} Judgments about proceeding must balance safety against the certainty of the challenge outcome. Although more time-consuming, undertaking additional challenges (eg, 5 challenges per food with 3 challenges containing placebo and 2 challenges containing the food allergen or *vice versa* randomly) can increase the accuracy of conclusions when symptoms are subjective.¹³

A scoring system appropriate for acute allergic responses is shown in Fig 3. This scoring system has not been validated but represents the authors' clinical experience. The scoring system indicates symptoms and signs that might warrant caution (repeating a dose, delaying a dose, and consideration for stopping) or are clear enough to warrant stopping a challenge and declaring the result positive. When DBPCFCs are used to evaluate disease outcomes in which symptoms are delayed (eg, isolated gastrointestinal reactions and AD) or controversial (migraine, behavior, and arthritis), different approaches, such as symptom diaries and SCORAD scores, might be needed. Similarly, evaluation of isolated oral symptoms from pollen-food-related syndrome requires modification of outcome measures. Challenge protocols might require modifications relevant to specific ages, diseases, and challenge circumstances; it is suggested to report in detail how symptoms were assessed with regard to stopping dosing and determination of positive, negative, or inconclusive challenge results to allow comparison of outcomes.

LATE REACTIONS TO FOOD IN PATIENTS WITH AD Clinical reaction patterns to food in patients with AD

Although food as a trigger of immediate-type allergic reactions can be readily suspected in many patients with a detailed history, skin test responses, and demonstration of specific IgE toward the suspected food, the identification of triggers for delayed reactions, such as eczematous reactions in patients with AD, is often more difficult.²⁶ Provocation tests can lead to 3 different reaction patterns in patients with AD⁹⁵:

1. noneczematous reactions, which are usually IgE mediated and present on the skin as pruritus, urticaria, or flush reactions, as well as with other immediate-type reactions of the gastrointestinal or respiratory tracts, or as anaphylaxis;
2. isolated eczematous delayed reactions (ie, usually flare of eczema after hours or 1-2 days); or

3. a combination of a noneczematous early reaction and an eczematous delayed reaction.

Analyses of food challenges in children show that isolated late reactions (ie, those after 2 hours) occur in about 12% of children with AD.^{41,96} Adolescents and adults can also react to foods, but reactions to "classical" food allergens, such as hen's eggs and cow's milk, are not as common as in childhood. Some patients with AD react to pollen-associated foods with late reactions.^{97,98} Triggering AD with pollen-associated foods is especially relevant in adolescent and adult patients.

Practical aspects of the diagnostic algorithm in suspected eczematous reactions to food allergens

Diagnosing food allergy in patients with AD does not differ in principle from the diagnostic workup in other allergic diseases.⁶⁷ However, serum food-specific IgE levels, SPT responses, and clinical history often do not correlate well with clinical findings. In particular, the value of history has been shown to be less useful in suspected late reactions compared with immediate reactions.^{96,99,100} When food allergy is suspected, appropriate *in vivo* tests (eg, skin tests), *in vitro* tests (eg, serum specific IgE), or both should be performed. For routine skin testing, only SPTs are recommended for the diagnosis of food allergy in patients with AD.

The atopy patch test (APT) with cow's milk, hen's egg, cereals, and peanut might be of some value, identifying specific food allergies in patients with AD in the following cases⁶¹: (1) suspicion of food allergy without predictive specific IgE levels or positive SPT responses; (2) severe AD, persistent AD, or both with unknown trigger factors; and (3) multiple IgE sensitizations without proved clinical relevance in patients with AD. The APT can therefore be considered an additional diagnostic tool that can be used in specialized institutions. However, standardized reagents and methods for performing APTs and interpreting their results have not been established, and therefore APTs are not recommended for routine diagnoses of food-induced eczema. Of note, it was concluded from an evaluation of a large number of children with AD that the APT does not lead to a significant reduction in the need for OFCs when food-induced eczema is suspected.¹⁰¹

In cases of suspected food allergy (by history, specific sensitization, or both), a diagnostic elimination diet of suspected food items over a period of up to 4 to 6 weeks is recommended. However, OFCs should be performed, even if the skin improved

TABLE II. Subjective symptoms

General nonspecific pruritus
Scratching
Nasal pruritus
Ocular pruritus
Dyspnea (without objective signs)
Throat "tightness"
Nausea
Abdominal pain
Oral/throat pruritus
Complaints of weakness, dizziness, not feeling well, etc

during a diagnostic elimination diet, to confirm the diagnosis. It is generally not sufficient to accept the result of an elimination diet as diagnostic because a variety of factors can influence the outcome during the elimination period.

Oral provocation tests in suspected eczematous reactions

In patients with AD, oral provocation should be performed after a period of suspected food allergen elimination when symptoms are minimal. If a stable situation cannot be attained by avoidance alone, topical therapy must be intensified before the provocation and continued in the same fashion throughout the OFCs. Accompanying therapy and conditions should be changed as little as possible. If a corticosteroid is absolutely necessary, then the application of a weak preparation (eg, 1% hydrocortisone) once daily is possible. Clinically relevant reactions will not be prevented with this minimal treatment, although on the other hand, natural fluctuations of the eczema are minimized. Steroid therapy should be continued over the entire provocation; no antihistamine or UV therapy should be given.

Even when no immediate-type reactions are reported, provocation should only be performed by clinicians with experience in the treatment of immediate allergic reactions. In children with AD who have been maintaining elimination diets for a prolonged period of time, very severe to life-threatening reactions have been reported after reintroduction of a food allergen.^{15,16}

Oral provocation should be initiated in a form similar to that for expected immediate reactions, as discussed above. The dose should be increased every 20 minutes up to the highest dose or until a clinical reaction occurs. The total dose should approximately correspond to the average serving at a meal (eg, 1 egg or 150 mL of milk). Because of the expected delayed reaction, the skin must be inspected by a medical doctor who is familiar in scoring eczema before the OFC and on the day after the challenge. If possible, repetitive provocation with the same food for another 1 to 2 days is recommended in those patients who did not react after the first challenge day. Clinical evaluation with a standardized eczema scoring system (eg, SCORAD or Eczema Area and Severity Index score) must be uniform throughout the period. Fig 4 summarizes the diagnostic algorithm for the identification of late reactions in patients with AD.

DEALING WITH SUBJECTIVE SYMPTOMS IN DBPCFCs

The DBPCFC is a rigorous tool that has become popular for evaluating adverse reactions to foods. The standard use of the DBPCFC has been to document food allergies for individual

patients, but it recently has been gaining acceptance as a procedure for investigating the effectiveness of therapies to prevent/minimize food-induced anaphylaxis. The DBPCFC is administered in increasing (fixed) doses to each participant, and the challenge is discontinued when the participant exhibits objective symptoms (eg, vomiting, diarrhea, and urticaria) at a specific dose or when the top dose is consumed without evidence of reactivity. This presentation addresses statistical design and analysis issues for the DBPCFC when subjects exhibit subjective symptoms (eg, abdominal pain and throat tightness).

The DBPCFC

Assume that the DBPCFC consists of K increasing doses of the food item, denoted as follows:

$$d_1 < d_2 < \dots < d_k.$$

A subject undergoing the active component of the DBPCFC progresses through the increasing doses of the food item. The DBPCFC result can be discontinued for a subject because of any of the following events:

- objective symptoms occurring at one of the doses;
- no objective symptoms occurring through dose d_k ; and
- withdrawal before objective symptoms occur.

If B or C occurs, then that subject is said to be "right censored" at the final dose that was administered. Censoring is said to occur when results of an observation are only partially known. The observation is considered right censored when the final (true) result is greater than the observed result, but it is not known by how much. In the case of B and C, the investigator would know that the subject has tolerated a certain dose without symptoms but how much more the subject could tolerate and remain symptom free is unknown.

For the i th subject in the study ($i = 1, 2, \dots, n$), 2 items, D_i and λ_i , are rerecorded, where D_i is defined as the final dose administered, λ_i is defined as 1 if objective symptoms were observed at D_i , and λ_i is defined as 0 if objective symptoms were not observed at D_i .

As an example, suppose there are 5 dose levels of the food item designated as follows:

$$d_1 = 5 \text{ mg}, d_2 = 10 \text{ mg}, d_3 = 20 \text{ mg}, d_4 = 40 \text{ mg}, d_5 = 80 \text{ mg},$$

and 4 subjects undergo the active component of the DBPCFC (Table III).

Subjective symptoms

What if a subject undergoing the DBPCFC describes subjective symptoms? What should be done? Hourihane et al¹⁰² continued the challenge after the subjective symptoms were completely resolved. Wensing et al¹⁰³ discontinued the challenge if the subjective symptoms endured for at least 60 minutes. Flinterman et al¹⁶ discontinued the challenge if the subjective symptoms occurred on 3 consecutive doses. Peeters et al¹⁰⁴ discontinued the challenge if the subjective symptoms lasted for at least 45 minutes.

An approach that minimizes the potential bias caused by the occurrence of subjective symptoms is as follows:

I. SKIN

A. Erythematous Rash- % area involved_____

B. Pruritus

0 = Absent

1 = Mild, occasional scratching

2 = Moderate -scratching continuously for > 2 minutes at a time

3 = Severe – hard continuous scratching – excoriations

C. Urticaria/Angioedema

0 = Absent

1 = Mild – < 3 hives, or mild lip edema

2 = Moderate - < 10 hives but >3, or significant lip or face edema

3 = Severe – generalized involvement

D. Rash

0 = Absent

1 = Mild – few areas of faint erythema

2 = Moderate – areas of erythema

3 = Severe – generalized marked erythema (>50%)

II. UPPER RESPIRATORY

A. Sneezing/Itching

0 = Absent

1 = Mild – rare bursts, occasional sniffing

2 = Moderate – bursts < 10, intermittent rubbing of nose, and/or eyes or frequent sniffing

3 = Severe – continuous rubbing of nose and/or eyes, periorcular swelling and/or long bursts of sneezing, persistent rhinorrhea

III. LOWER RESPIRATORY

A. Wheezing

0 = Absent

1 = Mild – expiratory wheezing to auscultation

2 = Moderate – inspiratory and expiratory wheezing

3 = Severe – use of accessory muscles, audible wheezing

B. Laryngeal

0 = Absent

1 = Mild – >3 discrete episodes of throat clearing or cough, or persistent throat tightness/pain

2 = Moderate – hoarseness, frequent dry cough

3 = Severe – stridor

IV. GASTROINTESTINAL

A. Subjective Complaints

0 = Absent

1 = Mild–complaints of nausea or abdominal pain, itchy mouth/throat

2 = Moderate – frequent c/o nausea or pain with normal activity

3 = Severe - notably distressed due to GI symptoms with decreased activity

B. Objective Complaints

0 = Absent

1 = Mild – 1 episode of emesis or diarrhea

2 = Moderate – 2-3 episodes of emesis or diarrhea or 1 of each

3 = Severe – >3 episodes of emesis or diarrhea or 2 of each

V. CARDIOVASCULAR/NEUROLOGIC

0 = normal heart rate or BP for age/baseline

1 = mild-subjective response (weak, dizzy), or tachycardia

2 = moderate-drop in blood pressure and/or >20% from baseline, or significant change in mental status.

3 = severe-cardiovascular collapse, signs of impaired circulation (unconscious)

TABLE LEGEND:

GREEN:

- Not usually an indication to alter dosing.
- Not generally sufficient to consider a challenge positive.

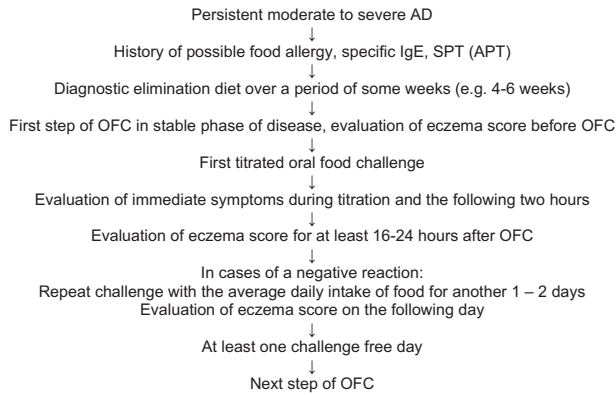
Orange (scores increasing to orange):

- Caution, dosing could proceed, be delayed, have a dose repeated rather than escalated.
- If clinically indicated, dosing is stopped.
- Symptoms that recur on 3 doses, or persist (e.g., 40 minutes) are more likely indicative of a reaction than when such symptoms are transient and not reproducible.
- 3 or more scoring areas in orange more likely represent a true response.

RED:

- Objective symptoms likely to indicate a true reaction
- Usually an indication to stop dosing.

FIG 3. Scoring the challenge outcome (modified from Bock et al⁶⁵ and Nowak-Wegrzyn et al¹⁴). The scoring system proposed here can be used for IgE-mediated reactions to determine the degree of response in various target organs and changes from baseline. Challenges should usually not commence if there are baseline symptoms exceeding descriptions in *green* (an exception might be AD that remains moderate despite maximal therapy). See the text for additional comments. (Please note that I,C,1 was made *orange* because it is similar to mild objective symptoms in other areas, is not a stopping indication, and, depending on clinical judgment, might or might not represent contact urticaria).



OFC: oral food challenge (ideally DBPCFC should be performed); APT: Atopy Patch Test, SPT: Skin prick test, AD: atopic dermatitis

FIG 4. Diagnostic algorithm for the identification of food allergy in patients with AD.

1. continue the challenge, in spite of the subjective symptoms, until the subject exhibits objective symptoms, or
2. if the subject is unable or unwilling to continue the challenge because of the subjective symptoms, then discontinue the challenge and consider the final administered dose to be right censored.

Although this approach is more accurate (less biased), it will be less precise (more variable) because of the increased amount of right censoring. Less precision results in a larger sample size requirement.

Statistical analysis

The most appealing approach for the between-group comparison is a discrete-time survival analysis. This analysis is based on a model for the conditional probability of a subject responding to dose d_k ($k = 1, 2, \dots, K$), given that the subject has not responded to previously administered doses.

This conditional probability is called the hazard function, and an appropriate model to describe this process is the discrete-time hazard function (in this situation the discrete-dose hazard function) as follows:

$$\lambda(k; x_1, x_2, \dots, x_r) = 1 - \exp\{-\exp(a_k + x_1\beta_1 + x_2\beta_2 + \dots + x_r\beta_r)\},$$

where x_1, x_2, \dots, x_r denote a set of r regressors, such as treatment group assignment, baseline score, sex, and age.

This particular form for the discrete-time hazard function, as proposed by Prentice and Gloeckler¹⁰⁵ and Allison,¹⁰⁶ is called the extreme value hazard function, and a statistical analysis based on this model requires a logistic regression analysis with the complementary log-log link function. The extreme value hazard function is the discrete-time analog of the Cox proportional hazards function. Researchers might not be familiar with the extreme value hazard function because they are more familiar with the exponential hazard function that is used for continuous time survival analysis.

The parameters $\alpha_1, \alpha_2, \dots, \alpha_K$ and the parameters $\beta_1, \beta_2, \dots, \beta_r$ are estimated through maximum likelihood (ML) estimation. Statistical inference, such as hypothesis testing for comparing

TABLE III. Subject responses to doses in DBPCFC

Individual i	D_i	λ_i	Explanation
1	$d_1 = 5$ mg	1	Objective symptoms at d_1
2	$d_4 = 40$ mg	1	Objective symptoms at d_4
3	$d_5 = 80$ mg	0	Right censored at d_5
4	$d_3 = 20$ mg	0	Right censored at d_3

the hazard rates of the experimental therapy and control groups, is based on the ML estimates of $\beta_1, \beta_2, \dots, \beta_r$. SAS PROC LOGISTIC (SAS, Institute, Cary, NC) is available for the logistic regression analysis with the complementary log-log link function.

An example is taken from Chinchilli et al.¹⁰⁷ A dataset for a clinical trial comparing an experimental therapy with a control therapy in patients with peanut allergy was simulated as follows:

Example: Six doses of peanut flour were administered during a baseline DBPCFC and study's end DBPCFC:

$d_1 = 1$ mg	$\log_4(d_1) = 0$
$d_2 = 4$ mg	$\log_4(d_2) = 1$
$d_3 = 16$ mg	$\log_4(d_3) = 2$
$d_{4tt} = 64$ mg	$\log_4(d_4) = 3$
$d_5 = 256$ mg	$\log_4(d_5) = 4$
$d_6 = 1024$ mg	$\log_4(d_6) = 5$

Fifty patients were randomized to each of the experimental therapy and control groups. The final dose administered was coded as 0, 1, 2, 3, 4, or 5 (\log_4 of the dose). The results of the analysis indicate that the baseline score and group assignment are statistically significant (Table IV).

The negative sign for the ML estimate of the experimental therapy group (-1.157) effect indicates that the hazard (conditional probability of objective symptoms) for the experimental therapy group is less than the hazard for the control group; that is, the experimental therapy is significantly more effective than the control group.

Estimating sample size

Because the discrete-time survival analysis is the discrete analog of the Cox proportional hazards regression model, sample size for the former can be approximated by using a sample size formula based on the latter. One such formula is provided by Schoenfeld.¹⁰⁸ This sample size formula requires that the effect size be expressed in terms of the hazard ratio, which is the ratio of the hazard functions for the experimental and control groups. Typical values that researchers use for effect sizes in sample size calculations range from a hazard ratio of 1.5 (small effect size) to 3.0 (moderate effect size).

In addition to the effect size, the researcher needs to account for the anticipated censoring rate (averaged across the 2 treatment arms) at the study's end food challenge. For example, if a researcher anticipates that the experimental therapy group will experience approximately 20% censoring and the control group will experience approximately 0% censoring, then this is an average of 10% censoring in the study. If subjective symptoms are anticipated, then there will be a higher level of right censoring, and hence an increased sample size will be necessary. Sample sizes (per treatment arm) for various hazard ratios (1.5, 2.0, 2.5,

TABLE IV. Outcome of peanut challenges in 100 patients randomized to experimental therapy or control group

	β (ML estimate)	SE	P value
Baseline score	-1.109	0.124	<.0001
Therapy group	-1.157	0.263	<.0001

TABLE V. Sample size varies dependent upon hazard ratio and anticipated right-censoring

Hazard ratio	0% Censoring	10% Censoring	20% Censoring	30% Censoring
1.5	128	142	160	182
2.0	44	49	55	63
2.5	25	28	31	36
3.0	18	20	23	26

and 3.0) and censoring rates (0%, 10%, 20%, and 30%) for a 2-sided, .05 significance level test with 90% statistical power are presented in the Table V. Notice that the sample size increases as (1) the hazard ratio decreases and (2) the censoring rate increases.

In summary, an appropriate statistical analysis of a DBPCFC in the presence of objective symptoms is a discrete-time survival analysis. Statistical software packages, such as SAS, have available routines for applying such an analysis. When planning a DBPCFC, it is important to calculate sample size based on the assumed hazard ratio and the level of right censoring (the percentage of participants who will not experience objective symptoms). The formula presented by Schoenfeld¹⁰⁸ can be invoked to approximate the sample size. The statistical approach described above is a very conservative approach with respect to subjective symptoms in that a participant who discontinues because of subjective symptoms is regarded as right censored. A more sensitive statistical approach needs to be developed that accounts for subjective symptoms.

OFC: REPORTING RESULTS

Not only should the reasons for stopping a challenge be reported (ie, objective symptoms vs repeated subjective, worsening subjective, or persistent subjective symptoms) but also studies should tabulate the numbers of patients who fall into the different categories. The number of patients with severe systemic reactions and their characteristics should be reported. The use of epinephrine to treat a reaction is not an adequate measure of severity because multiple patient-based factors, site issues (eg, the threshold to administer epinephrine might be lower in an outpatient setting), and the physician's judgment will affect the treatment of a reaction. Investigators are encouraged to make individual data available when possible in online repositories. In summary, Tables VI and VII list the recommended data that should optimally be included in research reports.

Placebo responses

Outcomes of DBPCFCs for research purposes should be published, with the results of the placebo challenges clearly described. The interpretation of equivocal challenge results should be determined in advance, as shown in the example in

TABLE VI. Data to report in studies using OFCs

1. Demographics (can be summarized in larger studies but helpful to have in an online repository for individual data) Age, sex, age at diagnosis, reaction symptoms, SPT responses, <i>in vitro</i> specific IgE levels, other atopic disorders
2. Were patients with a history of anaphylaxis included? How defined? Subanalysis might be indicated
3. Food matrix used (recipe), including percentage fat content
4. Method to validate blinding between placebo and verum
5. Exact description of challenge food's physical state (eg, raw, cooked, dehydrated, or defatted)
6. Dose escalation schedule Should be expressed as mg of protein
7. Setting for challenges
8. Scoring system with predetermined stopping/delaying/repeating/progressing decision points
9. Percentage of patients stopped for objective responses and subjective responses
10. Eliciting dose in mg of protein
11. Cumulative dose in mg of protein
12. Percentage positive, negative, or inconclusive, with criteria defined in advance

TABLE VII. Example of a decision matrix

Active food challenge	Placebo challenge	Interpretation
Positive	Negative	Positive
Positive (clearly more positive than placebo)	Positive	Positive
Negative	Negative	Negative
Negative (or positive, but clearly less positive than placebo)	Positive	Negative

Reprinted with permission from Vlieg-Boerstra et al.¹⁹

Fig 3.¹⁹ Placebo responses are important to publish because these rates should be used in the statistical handling of the results to accommodate and estimate false-positive rates; such information is also important in study design.¹⁰⁹

Eliciting versus cumulative dose

Both the eliciting and cumulative dose in milligrams of protein should be reported because this provides the maximum information for later interpretation of the challenge outcome.⁶⁸ For example, it might seem that the eliciting dose is the sole dose responsible for a reaction that begins in the mouth with pruritus and angioedema, but to date, it cannot be excluded that previous subthreshold doses served to prime the response or were subject to delayed absorption, and therefore both should be reported. Studies are currently lacking that compare a possible discrete dose effect versus the cumulative dose contribution.

Reporting of randomized, controlled therapeutic trials

As food allergy research continues to move beyond characterization to randomized therapeutic trials, the importance of adopting consistent parameters for reporting results is evident because this will allow a critical comparison of parameters

between studies and facilitate the application of meta-analyses. A framework developed in the 1990s (as the use of meta-analysis expanded) has been widely adopted by journal editorial boards on an international level. The Consolidated Standards of Reporting Trials (CONSORT) has outlined 22 aspects of clinical trial design and execution that are essential components of a high-quality randomized trial.^{110,111} The main principles of transparent reporting of trials include full disclosure of each step of enrollment (eg, eligibility and recruitment), allocation to intervention, follow-up, and analysis. A flow chart facilitates visualization of the numbers of participants at each phase of the study and whether analysis was on an intent-to-treat basis. Because the power of an individual study is extended when combined with other trials, it is important for food allergy intervention trials to adhere to CONSORT to ensure high quality, even if the proposed trial is small or of inadequate power for generalization of results.

REFERENCES

- Ingelfinger FJ, Lowell FC, Franklin W. Gastrointestinal allergy. *N Engl J Med* 1949;241:303.
- Loveless MH. Milk allergy: a survey of its incidence; experiments with a masked ingestion test. *J Allergy* 1950;21:489-99.
- Loveless MH. Allergy for corn and its derivatives: experiments with a masked ingestion test for its diagnosis. *J Allergy* 1950;21:500-9.
- Loveless MH. Cottonseed protein vs. cottonseed oil sensitivity; an objective approach to the diagnosis of food-allergy as applied to cottonseed atopy. *Ann Allergy* 1950;8:15-22.
- Goldman AS, Anderson DW, Sellers WA, Saperstein A, Kniker WT, Halpern SR. Milk allergy. I. Oral challenge with milk and isolated milk proteins in allergic children. *Pediatrics* 1963;32:425-43.
- Goldman AS, Sellers WA, Halpern SR, Anderson DW, Furlow TE, Johnson CH. Milk allergy. II. Skin testing of allergic and normal children with purified milk proteins. *Pediatrics* 1963;32:572-9.
- Saperstein S, Anderson DW Jr, Goldman AS, Kniker WT. Milk allergy. III. Immunological studies with sera from allergic and normal children. *Pediatrics* 1963;32:580-7.
- May CD. Objective clinical and laboratory studies of immediate hypersensitivity reactions to food in asthmatic children. *J Allergy Clin Immunol* 1976;58:500-15.
- Niggemann B, Lange L, Finger A, Ziegert M, Muller V, Beyer K. Accurate oral food challenge requires a cumulative dose on a subsequent day. *J Allergy Clin Immunol* 2012;130:261-3.
- Sicherer SH, Sampson HA. Food allergy. *J Allergy Clin Immunol* 2010;125(suppl 2):S116-25.
- 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: report of the NIAID-sponsored expert panel. *J Allergy Clin Immunol* 2010;126(suppl):S1-58.
- Niggemann B, Beyer K. Diagnosis of food allergy in children: toward a standardization of food challenge. *J Pediatr Gastroenterol Nutr* 2007;45:399-404.
- Bindslev-Jensen C, Ballmer-Weber BK, Bengtsson U, Blanco C, Ebner C, Hourihane J, et al. Standardization of food challenges in patients with immediate reactions to foods—position paper from the European Academy of Allergology and Clinical Immunology. *Allergy* 2004;59:690-7.
- Nowak-Wegrzyn A, Assa'ad AH, Bahna SL, Bock SA, Sicherer SH, Teuber SS. Work Group report: oral food challenge testing. *J Allergy Clin Immunol* 2009;123(suppl):S365-83.
- David TJ. Anaphylactic shock during elimination diets for severe atopic dermatitis. *Arch Dis Child* 1984;59:983-6.
- Flinterman AE, Knulst AC, Meijer Y, Bruijnzeel-Koomen CA, Pasmans SG. Acute allergic reactions in children with AEDS after prolonged cow's milk elimination diets. *Allergy* 2006;61:370-4.
- Untersmayr E, Jensen-Jarolim E. The role of protein digestibility and antacids on food allergy outcomes. *J Allergy Clin Immunol* 2008;121:1301-8.
- Lieberman P, Kemp SF, Oppenheimer JJ, Land I, Bernstein IL, Nicklas RA, et al. The diagnosis and management of anaphylaxis: an updated practice parameter. *J Allergy Clin Immunol* 2005;115(suppl):S485-523.
- Vlieg-Boerstra BJ, van der Heide S, Bijleveld CM, Kukler J, Duiverman EJ, Dubois AE. Placebo reactions in double-blind, placebo-controlled food challenges in children. *Allergy* 2007;62:905-12.
- Sampson HA. Food allergy—accurately identifying clinical reactivity. *Allergy* 2005;60(suppl 79):19-24.
- Burks AW, Jones SM, Boyce JA, Sicherer SH, Wood RA, Assa'ad A, et al. NIAID-sponsored 2010 guidelines for managing food allergy: applications in the pediatric population. *Pediatrics* 2011;128:955-65.
- Sampson HA. Food Allergy. Part 2: diagnosis and management. *J Allergy Clin Immunol* 1999;103:981-99.
- Host A, Husby S, Osterballe O. A prospective study of cow's milk allergy in exclusively breast-fed infants. Incidence, pathogenetic role of early inadvertent exposure to cow's milk formula, and characterization of bovine milk protein in human milk. *Acta Paediatr Scand* 1988;77:663-70.
- Osterballe M, Hansen TK, Mortz CG, Bindslev-Jensen C. The clinical relevance of sensitization to pollen-related fruits and vegetables in unselected pollen-sensitized adults. *Allergy* 2005;60:218-25.
- Osterballe M, Mortz CG, Hansen TK, Andersen KE, Bindslev-Jensen C. The prevalence of food hypersensitivity in young adults. *Pediatr Allergy Immunol* 2009;20:686-92.
- Werfel T, Ballmer-Weber B, Eigenmann PA, Niggemann B, Rance F, Turjanmaa K, et al. Eczematous reactions to food in atopic eczema: position paper of the EAACI and GA2LEN. *Allergy* 2007;62:723-8.
- Bock S, Buckley J, Holst A, May C. Proper use of skin tests with food extracts in diagnosis of food hypersensitivity. *Clin Allergy* 1978;8:559-64.
- Rosen J, Selcow J, Mendelson L, Grodofsky M, Factor J, Sampson H. Skin testing with natural foods in patients suspected of having food allergies...is it necessary? *J Allergy Clin Immunol* 1994;93:1068-70.
- Verstege A, Mehl A, Rolinck-Werninghaus C, Staden U, Nocon M, Beyer K, et al. The predictive value of the skin prick test wheal size for the outcome of oral food challenges. *Clin Exp Allergy* 2005;35:1220-6.
- Tripodi S, Businco AD, Alessandri C, Panetta V, Restani P, Matricardi PM. Predicting the outcome of oral food challenges with hen's egg through skin test endpoint titration. *Clin Exp Allergy* 2009;39:1225-33.
- Sporik R, Hill DJ, Hosking CS. Specificity of allergen skin testing in predicting positive open food challenges to milk, egg and peanut in children. *Clin Exp Allergy* 2000;30:1540-6.
- Calvani M, Alessandri C, Frediani T, Lucarelli S, Miceli SS, Panetta V, et al. Correlation between skin prick test using commercial extract of cow's milk protein and fresh milk and food challenges. *Pediatr Allergy Immunol* 2007;18:583-8.
- Mortz CG, Andersen KE, Bindslev-Jensen C. The prevalence of peanut sensitization and the association to pollen sensitization in a cohort of unselected adolescents—the Odense Adolescence Cohort Study on Atopic Diseases and Dermatitis (TOACS). *Pediatr Allergy Immunol* 2005;16:501-6.
- Asero R, Monsalve R, Barber D. Profilin sensitization detected in the office by skin prick test: a study of prevalence and clinical relevance of profilin as a plant food allergen. *Clin Exp Allergy* 2008;38:1033-7.
- Inomata N. Wheat allergy. *Curr Opin Allergy Clin Immunol* 2009;9:238-43.
- Beyer K, Chung D, Schulz G, Mishoe M, Niggemann B, Wahn U, et al. The role of wheat omega-5 gliadin IgE antibodies as a diagnostic tool for wheat allergy in childhood. *J Allergy Clin Immunol* 2008;122:419-21.
- Sampson H, Ho D. Relationship between food-specific IgE concentration and the risk of positive food challenges in children and adolescents. *J Allergy Clin Immunol* 1997;100:444-51.
- Sampson HA. Utility of food-specific IgE concentrations in predicting symptomatic food allergy. *J Allergy Clin Immunol* 2001;107:891-6.
- Osterballe M, Bindslev-Jensen C. Threshold levels in food challenge and specific IgE in patients with egg allergy: is there a relationship? *J Allergy Clin Immunol* 2003;112:196-201.
- Mehl A, Verstege A, Staden U, Kulig M, Nocon M, Beyer K, et al. Utility of the ratio of food-specific IgE/total IgE in predicting symptomatic food allergy in children. *Allergy* 2005;60:1034-9.
- Celik-Bilgili S, Mehl A, Verstege A, Staden U, Nocon M, Beyer K, et al. The predictive value of specific immunoglobulin E levels in serum for the outcome of oral food challenges. *Clin Exp Allergy* 2005;35:268-73.
- Ridout S, Matthews S, Gant C, Twiselton R, Dean T, Arshad SH. The diagnosis of Brazil nut allergy using history, skin prick tests, serum-specific immunoglobulin E and food challenges. *Clin Exp Allergy* 2006;36:226-32.
- Maloney JM, Rudengren M, Ahlstedt S, Bock SA, Sampson HA. The use of serum-specific IgE measurements for the diagnosis of peanut, tree nut, and seed allergy. *J Allergy Clin Immunol* 2008;122:145-51.
- Ott H, Baron JM, Heise R, Ocklenburg C, Stanzel S, Merk HF, et al. Clinical usefulness of microarray-based IgE detection in children with suspected food allergy. *Allergy* 2008;63:1521-8.

45. Benhamou AH, Zamora SA, Eigenmann PA. Correlation between specific immunoglobulin E levels and the severity of reactions in egg allergic patients. *Pediatr Allergy Immunol* 2008;19:173-9.
46. Hourihane JO, Grimshaw KEC, Lewis SA, Briggs RA, Trewin JB, King RM, et al. Does severity of low-dose, double-blind, placebo-controlled food challenges reflect severity of allergic reactions to peanut in the community? *Clin Exp Allergy* 2005;35:1227-33.
47. Sampson HA. Food Allergy. *J Allergy Clin Immunol* 2003;111:S540-7.
48. Wang J, Godbold JH, Sampson HA. Correlation of serum allergy (IgE) tests performed by different assay systems. *J Allergy Clin Immunol* 2008;121:1219-24.
49. Eller E, Kjaer HF, Host A, Andersen KE, Bindslev-Jensen C. Food allergy and food sensitization in early childhood: results from the DARC cohort. *Allergy* 2009;64:1023-9.
50. Sicherer SH. Clinical implications of cross-reactive food allergens. *J Allergy Clin Immunol* 2001;108:881-90.
51. Lin J, Bardina L, Shreffler WG, Andreae DA, Ge Y, Wang J, et al. Development of a novel peptide microarray for large-scale epitope mapping of food allergens. *J Allergy Clin Immunol* 2009;124:315-22.
52. Sastre J. Molecular diagnosis in allergy. *Clin Exp Allergy* 2010;40:1442-60.
53. Hansen KS, Ballmer-Weber BK, Sastre J, Lidholm J, Andersson K, Oberhofer H, et al. Component-resolved in vitro diagnosis of hazelnut allergy in Europe. *J Allergy Clin Immunol* 2009;123:1134-41.
54. Flinterman AE, Akkerdaas JH, Knulst AC, van Ree R, Pasmans SG. Hazelnut allergy: from pollen-associated mild allergy to severe anaphylactic reactions. *Curr Opin Allergy Clin Immunol* 2008;8:261-5.
55. Holzhauser T, Wackermann O, Ballmer-Weber BK, Bindslev-Jensen C, Scibilia J, Perono-Garoffo L, et al. Soybean (Glycine max) allergy in Europe: Gly m 5 (beta-conglycinin) and Gly m 6 (glycinin) are potential diagnostic markers for severe allergic reactions to soy. *J Allergy Clin Immunol* 2009;123:452-8.
56. Ando H, Moverare R, Kondo Y, Tsuge I, Tanaka A, Borres MP, et al. Utility of ovomucoid-specific IgE concentrations in predicting symptomatic egg allergy. *J Allergy Clin Immunol* 2008;122:583-8.
57. D'Urbano LE, Pellegrino K, Artesani MC, Donnanno S, Luciano R, Riccardi C, et al. Performance of a component-based allergen-microarray in the diagnosis of cow's milk and hen's egg allergy. *Clin Exp Allergy* 2010;40:1561-70.
58. Flinterman AE, Pasmans SG, Hoekstra MO, Meijer Y, Van Hoffen E, Knol EF, et al. Determination of no-observed-adverse-effect levels and eliciting doses in a representative group of peanut-sensitized children. *J Allergy Clin Immunol* 2006;117:448-54.
59. Nicolaou N, Poorafshar M, Murray C, Simpson A, Winell H, Kerry G, et al. Allergy or tolerance in children sensitized to peanut: prevalence and differentiation using component-resolved diagnostics. *J Allergy Clin Immunol* 2010;125:191-7.
60. Codreanu F, Collignon O, Roitel O, Thouvenot B, Sauvage C, Vilain AC, et al. A novel immunoassay using recombinant allergens simplifies peanut allergy diagnosis. *Int Arch Allergy Immunol* 2011;154:216-26.
61. Turjanmaa K, Darsow U, Niggemann B, Rance F, Vanto T, Werfel T. EAACI/GA2LEN position paper: present status of the atopy patch test. *Allergy* 2006;61:1377-84.
62. Spergel JM, Brown-Whitehorn T. The use of patch testing in the diagnosis of food allergy. *Curr Allergy Asthma Rep* 2005;5:86-90.
63. Stapel SO, Asero R, Ballmer-Weber BK, Knol EF, Strobel S, Vieths S, et al. Testing for IgG4 against foods is not recommended as a diagnostic tool: EAACI Task Force Report. *Allergy* 2008;63:793-6.
64. Kleine-Tebbe J, Erdmann S, Knol EF, MacGlashan DW Jr, Poulsen LK, Gibbs BF. Diagnostic tests based on human basophils: potentials, pitfalls and perspectives. *Int Arch Allergy Immunol* 2006;141:79-90.
65. Bock SA, Sampson HA, Atkins FM, Zeiger RS, Lehrer S, Sachs M, et al. Double-blind, placebo-controlled food challenge (DBPCFC) as an office procedure: a manual. *J Allergy Clin Immunol* 1988;82:986-97.
66. Rance F, Deschildre A, Villard-Truc F, Gomez SA, Paty E, Santos C, et al. Oral food challenge in children: an expert review. *Eur Ann Allergy Clin Immunol* 2009;41:35-49.
67. Sicherer SH, Sampson HA. Food allergy. *J Allergy Clin Immunol* 2006;117(suppl):S470-5.
68. Taylor SL, Hefle SL, Bindslev-Jensen C, Atkins FM, Andre C, Brujinzeel-Koomen C, et al. A consensus protocol for the determination of the threshold doses for allergenic foods: how much is too much? *Clin Exp Allergy* 2004;34:689-95.
69. Sicherer SH, Morrow EH, Sampson HA. Dose-response in double-blind, placebo-controlled oral food challenges in children with atopic dermatitis. *J Allergy Clin Immunol* 2000;105:582-6.
70. Perry TT, Matsui EC, Connover-Walker MK, Wood RA. Risk of oral food challenges. *J Allergy Clin Immunol* 2004;114:1164-8.
71. Perry TT, Matsui EC, Connover-Walker MK, Wood RA. The relationship of allergen-specific IgE levels and oral food challenge outcome. *J Allergy Clin Immunol* 2004;114:144-9.
72. Mankad VS, Williams LW, Lee LA, LaBelle GS, Anstrom KJ, Burks AW. Safety of open food challenges in the office setting. *Ann Allergy Asthma Immunol* 2008;100:469-74.
73. Gellerstedt M, Bengtsson U, Niggemann B. Methodological issues in the diagnostic work-up of food allergy: a real challenge. *J Investig Allergol Clin Immunol* 2007;17:350-6.
74. Niggemann B, Wahn U, Sampson HA. Proposals for standardization of oral food challenge tests in infants and children. *Pediatr Allergy Immunol* 1994;5:11-3.
75. Niggemann B, Beyer K. Pitfalls in double-blind, placebo-controlled oral food challenges. *Allergy* 2007;62:729-32.
76. Cox L, Nelson H, Lockey R, Calabria C, Chacko T, Finegold I, et al. Allergen immunotherapy: a practice parameter third update. *J Allergy Clin Immunol* 2011;127(suppl):S1-55.
77. Sampson HA, Munoz-Furlong A, Campbell RL, Adkinson NF Jr, Bock SA, Branum A, et al. Second symposium on the definition and management of anaphylaxis: summary report—Second National Institute of Allergy and Infectious Disease/Food Allergy and Anaphylaxis Network symposium. *J Allergy Clin Immunol* 2006;117:391-7.
78. Vlieg-Boerstra BJ, Bijleveld MA, van der Heide S, Beusekamp BJ, Wolt-Plompen SAA, Kukler J, et al. Development and validation of challenge materials for double-blind, placebo-controlled food challenges in children. *J Allergy Clin Immunol* 2004;113:341-6.
79. Crevel RW, Ballmer-Weber BK, Holzhauser T, Hourihane JO, Knulst AC, Mackie AR, et al. Thresholds for food allergens and their value to different stakeholders. *Allergy* 2008;63:597-609.
80. Cochrane SA, Salt LJ, Wantling E, Rogers A, Coutts J, Ballmer-Weber BK, et al. Development of a standardized low-dose double-blind placebo-controlled challenge vehicle for the EuroPrevall project. *Allergy* 2012;67:107-13.
81. Vlieg-Boerstra BJ, Duiverman EJ, van der Heide S, Bijleveld CM, Kukler J, Dubois AE. Should children with a history of anaphylaxis to foods undergo challenge testing? *Clin Exp Allergy* 2008;38:1935-42.
82. Eigenmann PA, Caubet JC, Zamora SA. Continuing food-avoidance diets after negative food challenges. *Pediatr Allergy Immunol* 2006;17:601-5.
83. Grimshaw KEC, King RM, Nordlee JA, Hefle SL, Warner JO, Hourihane JOB. Presentation of allergen in different food preparations affects the nature of the allergic reaction—a case series. *Clin Exp Allergy* 2003;33:1581-5.
84. James JM, Bernhisel-Broadbent J, Sampson HA. Respiratory reactions provoked by double-blind food challenges in children. *Am J Respir Crit Care Med* 1994;149:59-64.
85. James JM, Eigenmann PA, Eggleston PA, Sampson HA. Airway reactivity changes in food-allergic, asthmatic children undergoing double-blind placebo-controlled food challenges. *Am J Respir Crit Care Med* 1996;153:597-603.
86. Gabriele C, Hol J, Kerkhof E, Elink Schuurman BE, Samsom JN, Hop W, et al. Fractional exhaled nitric oxide in infants during cow's milk food challenge. *Pediatr Allergy Immunol* 2008;19:420-5.
87. Benhamou AH, Koehli A, Rochat I, Inci D, Moeller A, Tamarcaz P, et al. Exhaled nitric oxide decreases after positive food-allergen challenge. *Clin Transl Allergy* 2011;1:14.
88. Sampson HA, Jolie PL. Increased plasma histamine concentrations after food challenges in children with atopic dermatitis. *N Engl J Med* 1984;311:372-6.
89. Niggemann B, Beyer K, Wahn U. The role of eosinophils and eosinophil cationic protein in monitoring oral challenge tests in children with food-sensitive atopic dermatitis. *J Allergy Clin Immunol* 1994;94:963-71.
90. Beyer K, Niggemann B, Schulze S, Wahn U. Serum tryptase and urinary 1-methylhistamine as parameters for monitoring oral food challenges in children. *Int Arch Allergy Immunol* 1994;104:348-51.
91. Ohtsuka T, Matsumaru S, Uchida K, Onobori M, Matsumoto T, Kuwahata K, et al. Time course of plasma histamine and tryptase following food challenges in children with suspected food allergy. *Ann Allergy* 1993;71:139-46.
92. Vila L, Sanz ML, Sanchez-Lopez G, Garcia-Aviles C, Dieguez I. Variations of serum eosinophil cationic protein and tryptase, measured in serum and saliva, during the course of immediate allergic reactions to foods. *Allergy* 2001;56:568-72.
93. Clark AT, Mangat JS, Tay SS, King Y, Monk CJ, White PA, et al. Facial thermography is a sensitive and specific method for assessing food challenge outcome. *Allergy* 2007;62:744-9.
94. Clark A, Mangat J, King Y, Islam S, Anagnostou K, Foley L, et al. Thermographic imaging during nasal peanut challenge may be useful in the diagnosis of peanut allergy. *Allergy* 2012;67:574-6.
95. Werfel T, Breuer K. Role of food allergy in atopic dermatitis. *Curr Opin Allergy Clin Immunol* 2004;4:379-85.

96. Breuer K, Heratizadeh A, Wulf A, Baumann U, Constien A, Tetau D, et al. Late eczematous reactions to food in children with atopic dermatitis. *Clin Exp Allergy* 2004;34:817-24.
97. Reekers R, Busch M, Whittmann M, Kapp A, Werfel T. Birch pollen-related foods trigger atopic dermatitis in patients with specific cutaneous T-cell responses to birch pollen antigens. *J Allergy Clin Immunol* 1999;104:466-72.
98. Breuer K, Wulf A, Constien A, Tetau D, Kapp A, Werfel T. Birch pollen-related food as a provocation factor of allergic symptoms in children with atopic eczema/dermatitis syndrome. *Allergy* 2004;59:988-94.
99. Atherton DJ, Soothill JF, Sewell M, Wells RS, Chilvers CED. A double-blind controlled crossover trial of an antigen-avoidance diet in atopic eczema. *Lancet* 1978;1:401-3.
100. Niggemann B, Sielaff B, Beyer K, Binder C, Wahn U. Outcome of double-blind, placebo-controlled food challenge tests in 107 children with atopic dermatitis. *Clin Exp Allergy* 1999;29:91-6.
101. Mehl A, Rolinck-Werninghaus C, Staden U, Verstege A, Wahn U, Beyer K, et al. The atopy patch test in the diagnostic workup of suspected food-related symptoms in children. *J Allergy Clin Immunol* 2006;118:923-9.
102. Hourihane JO'B, Kilburn SA, Nordlee JA, Hefle SL, Taylor SL, Warner JO. An evaluation of the sensitivity of subjects with peanut allergy to very low doses of peanut protein: a randomized, double-blind, placebo-controlled food challenge study. *J Allergy Clin Immunol* 1997;100:596-600.
103. 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. *J Allergy Clin Immunol* 2002;110:915-20.
104. Peeters KA, Nordlee JA, Penninks AH, Chen L, Goodman RE, Bruijnzeel-Koomen CA, et al. Lupine allergy: not simply cross-reactivity with peanut or soy. *J Allergy Clin Immunol* 2007;120:647-53.
105. Prentice RL, Gloeckler LA. Regression analysis of grouped survival data with application to breast cancer data. *Biometrics* 1978;34:57-67.
106. Allison PD. Discrete-time methods for the analysis of event histories. *Sociological Methodology* Vol 13; 1982; American Sociological Assoc, Washington, DC; p. 61-98.
107. Chinchilli VM, Fisher L, Craig TJ. Statistical issues in clinical trials that involve the double-blind, placebo-controlled food challenge. *J Allergy Clin Immunol* 2005;115:592-7.
108. Schoenfeld DA. Sample-size formula for the proportional-hazards regression model. *Biometrics* 1983;39:499-503.
109. Briggs D, Aspinall L, Dickens A, Bindslev-Jensen C. Statistical model for assessing the proportion of subjects with subjective sensitisations in adverse reactions to foods. *Allergy* 2001;56(suppl 67):83-5.
110. Moher D, Schulz KF, Altman DG. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials. *J Am Podiatr Med Assoc* 2001;91:437-42.
111. Altman DG, Schulz KF, Moher D, Egger M, Davidoff F, Elbourne D, et al. The revised CONSORT statement for reporting randomized trials: explanation and elaboration. *Ann Intern Med* 2001;134:663-94.