Highly accurate prediction of food challenge outcome using routinely available clinical data

Audrey DunnGalvin, PhD,[a] Deirdre Daly, RN,[a] Claire Cullinane, RN,[a] Emily Stenke, MB,[a] Diane Keeton, RSCN,[b] Mich Erlewyn-Lajeunesse, MRCPCH DM,[b] Graham C. Roberts, MRCPCH DM,[b,c] Jane Lucas, MRCPCH PhD,[a,b,c] and Jonathan O'B. Hourihane, MB, DM, MRCPI, FRCPCH[a] *Cork, Ireland, and Southampton, United Kingdom*

Background: Serum specific IgE or skin prick tests are less useful at levels below accepted decision points.

Objectives: We sought to develop and validate a model to predict food challenge outcome by using routinely collected data in a diverse sample of children considered suitable for food challenge.

Methods: The proto-algorithm was generated by using a limited data set from 1 service (phase 1). We retrospectively applied, evaluated, and modified the initial model by using an extended data set in another center (phase 2). Finally, we prospectively validated the model in a blind study in a further group of children undergoing food challenge for peanut, milk, or egg in the second center (phase 3). Allergen-specific models were developed for peanut, egg, and milk.

Results: Phase 1 (N = 429) identified 5 clinical factors associated with diagnosis of food allergy by food challenge. In phase 2 (N = 289), we examined the predictive ability of 6 clinical factors: skin prick test, serum specific IgE, total IgE minus serum specific IgE, symptoms, sex, and age. In phase 3 (N = 70), 97% of cases were accurately predicted as positive and 94% as negative. Our model showed an advantage in clinical prediction compared with serum specific IgE only, skin prick test only, and serum specific IgE and skin prick test (92% accuracy vs 57%, and 81%, respectively).

Conclusion: Our findings have implications for the improved delivery of food allergy–related health care, enhanced food allergy–related quality of life, and economized use of health service resources by decreasing the number of food challenges performed. (J Allergy Clin Immunol 2011;127:633-9.)

Key words: *Food challenge, predictive model, validation, calculator, outcomes*

Abbreviations used

AUC: Area under the curve
DBPCFC: Double-blind, placebo-controlled food challenge
FA: Food allergy
OFC: Open food challenge
ROC: Receiver operator characteristics
sIgE: Serum-specific immunoglobulin
SPT: Skin prick test

The last decade has shown both an increasing prevalence of food allergy (FA) and increasing numbers of patients and parents seeking diagnosis.[1-5] Although double-blind, placebo-controlled food challenges (DBPCFC) still represent the gold standard for the diagnosis of FA, they are time-consuming and costly and can create some concerns for patients and parents—for example, fear of risk of severe systemic reactions for patients and parents.[6,7] Many pediatric units routinely use open food challenges (OFCs) outside research protocols. In addition, not all clinical locations have the facilities to carry out high-quality food challenges. Given these factors, efforts have been made to find diagnostic tests to predict the outcome of oral food challenges. These prognostic models have focused on decision points for sIgE or skin prick test (SPT), either independently or in combination.[8-12] Referral pattern (community, secondary care, or tertiary allergy service) did not alter this relationship. However, age of children affected IgE and SPT cutoff levels in some studies,[13-15] particularly for food challenges with egg and milk, with lower cutoff levels in infants under 2 years of age.[16,17] Furthermore, a history of anaphylaxis is often used to decide not to offer a food challenge, although resolution of FA is still possible with such a history. Therefore, it appears other clinical factors are inconsistently considered, and their impact on challenge outcome has previously never been evaluated systematically, although they are likely to be influential.

We hypothesized that a more complex model incorporating many proven or suspected clinical predictors would result in a better predictive accuracy of challenge outcome (DBPCFC and OFC) because each variable in any prognostic model represents only a part of the information. Thus, the more information considered, the more accurate the equation should be. Logistic regression has been applied to the diagnosis of several common diseases in recent literature.[18-22] The advantage of using probability values is that they are expressed on a continuous scale that gives a quantified probability of likely diagnostic outcome. In addition, when a patient falls within an immunologic gray area[23] (ie, below the 95% cutoff point for certain reactivity but above a level that would predict a negative outcome, usually a completely negative SPT or sIgE), a probability of clinical allergy can still

From [a]Paediatrics and Child Health, University College Cork; [b]the Southampton University Hospitals NHS Trust; and [c]Child Health, School of Medicine, University of Southampton.

Disclosure of potential conflict of interest: J. O'B. Hourihane receives speaker fees from King Pharmaceuticals, ALK-Abelló, Mead Johnson, and Phadia; receives research support from DANONE, the Children's Research Center, and the Food Standards Agency; has provided legal consultation services or expert witness testimony in cases related to food allergy death; is honorary secretary of the Irish Association of Allergy and Immunology; is a medical advisor for the Anaphylaxis Campaign UK; and has a patent application relating to electronic use of the algorithm developed in this project. D. Daly receives research support from DANONE and Mead Johnson. The rest of the authors have declared that they have no conflict of interest.

Received for publication March 9, 2010; revised November 26, 2010; accepted for publication December 2, 2010.

Reprint requests: Audrey DunnGalvin, PhD, Department of Paediatrics and Child Health, Clinical Investigations Unit, Cork University Hospital, Wilton, Cork, Ireland. E-mail: a.dunngalvin@ucc.ie.

0091-6749/$36.00

doi:10.1016/j.jaci.2010.12.004

be calculated. We sought to test our hypothesis that combining clinical variables, routinely available prechallenge, with allergy test results would improve our ability to predict the outcome of a food challenge.

METHODS

Recruitment for peanut, milk, and egg food challenges

Patient data are routinely collected in each participating center. A standard protocol and *pro forma* sheet for food challenges was developed in Southampton by J.H. in 1994,[24] adapted from Bock et al.[25] It was updated by J.L. in 1997 and has since remained in current use in identical form in both centers. There have been no major changes to challenge protocols during the study period.

Children were selected for challenge on the basis of clinical need to confirm a suspected diagnosis, to identify children who may have become tolerant to a previously reactive food, or to test for reactivity to foods not previously consumed. The majority of children (90%) underwent OFCs. A challenge was considered positive if objective or sustained subjective symptoms were observed.[25,26] Children with inconclusive results (<5% of all challenges) were excluded from analyses. Baked egg or baked milk challenges were not included. Consent for the study was obtained from the Clinical Research Ethics Committee of the Cork Teaching Hospitals. The data were prospectively collected in Southampton as part of an approved clinical audit project.

SPT testing protocol

Skin prick testing was performed by using positive control (histamine dihydrochloride 10 mg/mL), negative control (50% glycerol and 50% buffered saline), the potential food allergen (ALK-Abelló, Reading, United Kingdom), and single-head lancet. Skin reactions were recorded after 15 minutes. Positive test was a wheal of ≥3 mm. The size of the wheal was determined as the mean of 2 perpendicular longest diameters. Serum specific IgE was measured by using the UniCAP system (Phadia, Uppsala, Sweden), and results were reported in kUa/L (kilounits of allergen-specific IgE per liter).

Design

Our design consisted of 3 distinct phases (exploration/feasibility; development and evaluation; and validation) using data from 3 separate patient groups:

1. Exploratory retrospective generation of the proto-model by using anonymous unlinked data prospectively collected from challenged patients in Southampton (1998-2005). We used listwise deletion of variables to cater for missing data. We developed dichotomized predictors defined as high (above the decision point) or low (below the decision point for IgE or SPT levels) on the basis of published decision points for allergen types[10,12,14,23] as follows: SPT and sIgE for peanut (≥8 mm; ≥15 kUa/L), egg (≥7 mm; ≥7 kUa/L), and milk (≥8 mm; ≥15 kUa/L); age (≥7 years); and previous reactions (yes/no; group A, n = 429). History of reaction was dichotomized (yes/no).
2. Retrospective development, evaluation, and modification of the initial model using a similar but complete set of data (no missing data) already collected from patients challenged in Cork including sex, age, symptom history, serum specific IgE, total IgE, and SPT results. To create the symptom history variable, parents were presented with a list of 32 symptoms before challenge by using the clinical severity system devised by Ewan and Clark.[27,28] Symptoms were categorized and recategorized to create a graded variable with 4 categories in order of predictive ability: 1, skin or oral or gastrointestinal or upper respiratory only; 2, upper respiratory and gastrointestinal or 2 systems; 3, lower respiratory or 3 systems; 4, cardiovascular or 4 systems. This symptom grading corresponds closely with those put forward by other authors.[27-29] Three separate allergen-specific algorithms were developed for 3 main food types (peanut, egg, milk; group B, n = 289).
3. Prospective blind application and validation of the accuracy (including specificity and sensitivity) of the modified model using data routinely collected from other more recent patients undergoing a food challenge in Cork (group C, n = 70).

The main outcome was an allergen-specific calculator (that may act as a complementary decision tool for clinicians) predicting the probability of a positive challenge result with a high degree of accuracy (at least 80%, positive and negative) from a routinely available prechallenge set of variables.

Statistical analyses

The descriptive statistics were presented as means and SDs for measurement data and percentages for categoric data. Differences between groups were assessed by ANOVA for normally distributed measurement data, Wilcoxon test for nonnormally distributed measurement data, and the χ^2 test for categoric data. Continuous variables with a nonnormal distribution were log-transformed before inclusion in the regression models.[28]

In phase 1, we performed a feasibility study by using univariate logistic regression analysis to screen factors relating to the diagnosis of FA.[30] After a series of multivariate logistic regression analyses (stepwise forward, probability of entering a variable, .05; probability of removing a variable, .06), the final significant factors were included in the prognostic model.[30] The model itself was based on the logistic regression formula

$$\pi = ez/(1 + ez)$$

(see this article's Fig E1 in the Online Repository at www.jacionline.org). We present relative risk ratios instead of odds ratios to evaluate the contribution of each variable because the probability of a positive outcome was high.[30]

In phase 2, we used logistic regression analyses (as described) to develop a model for a calculator and to increase accuracy of diagnoses. We examined the independent and combined contribution of each reported symptom. By using stepwise forward multivariate regression analysis,[29-31] we identified groups of independent correlating categoric clinical symptoms from previous allergic reactions. We graded them into 4 categories; each of the 4 levels had its own estimator with no symptoms (0) as the indicator. In addition, we examined the independent and combined contribution of SPT, total IgE, sIgE, and total IgE minus sIgE to the overall predictive ability of the model.[30-32] On the basis of these analyses, we developed 3 separate allergen-specific models for the 3 foods studied (peanut, egg, milk). We subtracted sIgE from total IgE so that sIgE would be accounted for only once.

In phase 3, we conducted a prospective validation test of the model by using receiver operator characteristics (ROC) curve to evaluate sensitivity and specificity.[33] ROC analyses specify that the closer the curve is to the diagonal —that is, the closer the area under the curve (AUC) is to 0.5—the worse the model. In contrast, the closer the AUC is to 1.0, the better the model. The optimum balance of sensitivity and specificity indicated by the ROC output was used to determine the positive/negative cutoff points.

All tests were performed with SPSS (version 16.0 for Windows; SPSS Inc, Chicago, Ill).

RESULTS

Tables I, II, and III show that the clinical and demographic profiles of groups A, B, and C.

The mean age was 7 years, and there was approximately a 2:1 male:female ratio for each group. More boys than girls had positive challenges ($F = 4.054$; $P < .05$), although there were no significant differences across foods. Almost 80% of our total sample had reported symptoms after ingestion of peanut, milk or egg. Two percent of the children reporting symptoms after ingesting peanut, 4% of milk, and 5% of egg were suspected to have developed tolerance. All reactions were immediate. No significant differences were found between complete and missing profiles for SPT, sIgE, or reported symptoms. Fifteen percent of our sample

TABLE I. Demographic and clinical characteristics of group A, phase 1 (exploratory)

	Peanuts		Milk		Egg	
Group A	**n = 239***	**56%**	**n = 110***	**26%**	**n = 80***	**18%**
Male	153	64%	69	63%	52	65%
Mean age (y)	8.0	4.0	6.5	3.6	6.0	3.7
Mean SPT wheal (mm)	4.0	2.2	3.7	2.8	3.8	3.1
Mean sIgE (kUa/L)	5.8	4.0	6.0	4.1	5.7	3.6
Total IgE						
Positive sIgE/not previously eaten	56	23%	4	4%	24	30%
Skin or oral or gastrointestinal or upper respiratory only	54	23%	57	52%	26	32%
Upper respiratory and gastrointestinal or 2 systems	55	23%	26	24%	16	20%
Lower respiratory or 3 systems	46	19%	19	17%	11	14%
Cardiovascular or 4 systems	28	12%	4	4%	3	4%
Positive diagnosis on challenge	163	38%	172	40%	167	39%

Data represent means (SDs) or numbers (percentages) for each group. Groups A and B formed the retrospective development groups, whereas group C was the prospective validation group.
Entire data set, N = 680.
*Missing data: sex (35 cases), age (81 cases), SPT (174 cases), sIgE (250 cases), and symptoms reported (250 cases). No significant differences were found between complete and missing cases for SPT, sIgE, or reported symptoms.

had an SPT level of 7 mm or above, and 30% had an sIgE level of 5 kUA/L or above. SPT was highest for peanut in all groups (mean = 4.7 mm), and sIgE was highest for milk (mean = 8.1 kUA/L). In terms of positive and negative outcomes, there were no significant associations with age. As expected, positive and negative profiles differed for all foods and populations in terms of total IgE (t = 1.961; $P = .04$), sIgE (F = 9.625; $P = .008$), SPT (F = 9.625; $P = .001$), grade of symptoms (F = 7.768; $P = .001$), and history (F = 6.127; $P = .002$). There were no significant differences in type of challenge, whether open (n = 708) or double-blind, placebo-controlled (n = 80).

Development of the diagnostic model

In phase 1, univariate regression analysis identified factors that were associated with a positive food challenge (including history, SPT level, sIgE level, sex, and age). Total IgE results were not available. This model had an 80% accuracy in correctly predicting positive outcomes and 75% accuracy in negative outcomes for all foods. In this article's Table E1 in the Online Repository at www.jacionline.org, estimators for SPT and sIgE are shown as high (peanut, ≥8 mm, ≥15 kUa/L; egg, ≥7 mm, ≥7 kUa/L; milk, ≥8 mm, ≥15 kUa/L) or low. High SPT and high sIgE had the strongest relative risk for positive challenge (2.7 and 1.7, respectively).

In phase 2, we used logistic regression analyses to develop the model for the calculator and increase the accuracy of diagnoses. Log-transformation of continuous variables resulted in an improved performance of 0.4% for peanut, 0.2% for milk, and 0.1% for egg. The estimators for peanut, milk, and egg along with relative risks and CIs are shown in Table IV. The symbol π is the probability of a diagnosis of FA by a positive food challenge, x represents the 6 predictors, α is the constant and β is the coefficient estimator for each predictive variable. On the basis of this analysis, the diagnostic model was developed as shown in Fig E1.

The best combination of symptoms (in ascending order of predictive ability) corresponded to graded systems outlined by previous authors.[27-29] Upper respiratory and gastrointestinal symptoms, if present alone, had a weaker predictive value than upper respiratory and gastrointestinal symptoms presenting together. The presence of any 2 or more symptoms from different organ systems had more weight in prediction than multiple symptoms from 1 system. Although total IgE + sIgE and Total IgE/sIgE were both significant when evaluated, total IgE minus sIgE had the best predictive ability, adding almost 10% to the number of cases correctly classified as positive or negative.

Retrospective evaluation of the diagnostic model

Performance in group B was very satisfactory. Analysis showed that sensitivity and specificity were 94.0 (95% CI, 82.3-97.6) and 91.0 (95% CI, 76.8-97.5) for peanut, 92.0 (95% CI, 85.0-96.7) and 91.0 (95% CI, 81.2-96.4) for milk, and 92.0 (95% CI, 78.1-98.3) and 89.0 (95% CI, 75.9-70.8-97.6) for egg. The AUC was 0.95 (95% CI, 0.90-0.99; $P < .001$).

Prospective evaluation of the diagnostic model

We used this model to predict outcome of FA in 70 further patients (group C). The predictive value of this model was excellent (see this article's Table E2 in the Online Repository at www.jacionline.org). The area under the ROC curve was 0.97 for peanut, 0.95 for egg, and 0.94 for milk. All accurately predicted cases had residuals within 2 SDs of the highest (1.0) and lowest (0.12) probability values,[30-32] so that probabilities ranged from 0.8 to 1.0 for positive challenges and from 0.35 to 0.12 for negative cases. This means that low-risk cases (sIgE and SPT below decision points) can also be accurately predicted. The optimum criterion for positive versus negative diagnosis corresponded to a cutoff point >0.55 for peanut and egg, and a cutoff point >0.52 for milk. Table E2 also shows satisfactory values for cutoff points of >0.4 and >0.6 for peanut, milk, and egg, 83.0 to 100.0 for sensitivity, and 79.0 to 100.0 for specificity. Finally, we compared the predictive accuracy of different diagnostic methods (the diagnostic model we had developed vs the more traditional methods of using SPT or sIgE singly or in combination) separately in peanut, milk, and egg by using data from groups B and C. By comparison with other diagnostic methods, our model (including 6 variables) showed a clear advantage in clinical prediction (Table V). Considering either of these 3 other models (sIgE only, SPT only, sIgE and SPT) resulted in lower predictive value across the 3 foods (92% accuracy vs 57% to 81%, respectively).

TABLE II. Demographic and clinical characteristics of group B, phase 2 (development and evaluation)

	Peanuts		Milk		Egg	
	n = 94	33 %	n = 58	20%	n = 137	47%
Group B	**Mean or n**	**% or SD**	**Mean or n**	**% or SD**	**Mean or n**	**% or SD**
Male	53	56%	42	55%	74	54%
Age (y)	7.5	4.2	5.5	4.1	5.2	4.3
SPT wheal (mm)	4.0	2.0	3.0	1.6	3.4	1.6
sIgE (kUa/L)	8.2	6.1	9.6	7.3	7.7	6.0
Total IgE	1070	1129	1074	1276	1052	1141
Positive sIgE/not previously eaten	41	44%	3	5%	40	29%
Skin or oral or gastrointestinal or upper respiratory only	12	13%	33	57%	38	28%
Upper respiratory and gastrointestinal or 2 systems	27	29%	13	22%	29	21%
Lower respiratory or 3 systems	10	11%	7	12%	18	13%
Cardiovascular or 4 systems	4	4%	2	3%	12	9%
Diagnosis positive	50	53%	33	56%	80	58%

Data represent means (SDs) or numbers (percentages) for each group. Groups A and B formed the retrospective development groups, whereas group C was the prospective validation group. There were no missing data.

TABLE III. Demographic and clinical characteristics of group C, phase 3 (validation)

	Peanuts		Milk		Egg	
	n = 30	42 %	n = 20	29%	n = 20	29%
Group C	**Mean or n**	**% or SD**	**Mean or n**	**% or SD**	**Mean or n**	**% or SD**
Male	23	74%	14	70%	13	65%
Age (y)	8	3.0	6.5	2.8	5.5	3.3
SPT wheal (mm)	5.4	2.1	5.0	2.0	5.2	2.5
sIgE (kUa/L)	8.2	10.5	8.7	12.0	8.1	9.0
Total IgE	1129	842	1043	1166	1209	983
Positive sIgE/not previously eaten	12	42%	0	0%	5	25%
Skin or oral or gastrointestinal or upper respiratory only	6	18%	12	12%	5	25%
Upper respiratory and gastrointestinal or 2 systems	7	23%	5	25%	5	25%
Lower respiratory or 3 systems	3	10%	2	10%	4	20%
Cardiovascular or 4 systems	2	7%	1	5%	1	5%
Diagnosis positive	17	57%	11	55%	10	50%

Data represent means (SDs) or numbers (percentages) for each group. Groups A and B formed the retrospective development groups, whereas group C was the prospective validation group. There were no missing data.

DISCUSSION

An *in vitro* or in silico test that could more accurately diagnose IgE-mediated FA without automatic recourse to food challenge would be of great help as a complementary decision tool for clinicians. The current study was therefore designed to evaluate the clinical usefulness and diagnostic value of the model, using routinely collected data, in predicting food challenge outcome in a diverse sample of patients considered suitable for food challenge. Highly accurate, allergen-specific algorithms were developed separately for each of the 3 most important food types (peanut, egg, milk), consisting of variables that contributed independently to the accuracy of the final model.

We identified 6 major independent indicators: sex, age, history of reaction, sIgE, total IgE minus sIgE, and SPT. Grade of symptoms or history of reactions, which corresponded to previously published systems of severity grading,[27-29] emerged as one of the strongest predictors. A child, for example, who is suspected of having peanut allergy and has had a history of symptoms will have a higher relative risk for a positive outcome than a child with no symptoms, controlling for IgE result, SPT wheal diameter, sex, and age. The symptoms themselves were first each evaluated independently (and in combination with each other) in a series of univariate models and also with each of the other indicators. Overall, our findings demonstrate the importance of taking multiple indicators (including sex and age) into account to enhance prognostic ability rather than simply using sIgE and/or SPT. Time since reaction did not contribute to the models during any phase because it was probably confounded by age. Age had a significant effect on the outcome, but sIgE, SPT, and total IgE may also change over time. Other variables that were considered and rejected included whether the challenge was open or a DBPCFC and whether a child had an allergy to more than 1 food.

The accuracy of any model represents the most important consideration for comparison of different diagnostic methods. We accurately predicted 37 of 38 positive cases and 30 of 32 negative cases. The predictive value of this new calculator is therefore excellent. The AUC was 0.97 for peanut, 0.95 for egg, and 0.94 for milk ($P < .0001$). Furthermore, although the same clinical variables are present in the models for peanut, milk, and egg, their relative contribution is unique to each food type. For example, serum specific IgE has a much greater risk ratio, relative to the other variables in the model, in milk compared with peanut or egg models. This underscores the importance of using an allergen-specific prognostic calculator.

TABLE IV. The estimators, relative risk and confidence intervals (Phase 2) presented separately for peanut, milk and egg ($P < .05$)

Allergen	Indicators	b-estimate	95% CI	Relative risk‡	95% CI
Peanut*					
	Male	4.60	±3.4	8.6	4.5-16.4
Grade of symptoms†	1	3.32	±3.2	4.4	2.8-7.1
	2	4.61	±3.5	8.8	4.6-17.0
	3	7.86	±4.6	12.8	6.0-26.2
	4	11.08	±7.8	18.4	10.1-34.3
	SPT (mm)	2.85	±1.3	3.8	2.7-5.4
	sIgE (kUa/L)	0.50	±0.2	1.3	1.0-1.9
	IgE-sIgE (kUa)	−0.002	±0.001	1.1	0.8-1.4
	Age (y)	−0.37	±0.2	0.9	0.7-1.2
	Constant	−11.63	± 8.2		
Egg*					
	Male	1.70	±1.5	2.2	1.5-3.1
Grade of symptoms†	1	1.40	±1.2	1.9	1.3-2.8
	2	2.08	±1.7	2.6	1.7-3.7
	3	2.74	±2.1	3.5	2.2-5.4
	4	3.76	±2.7	5.5	3.2-9.1
	SPT (mm)	0.29	±0.1	1.2	0.5-3.7
	IgE-sIgE (kUa)	−0.004	±0.001	1.1	0.7-2.1
	sIgE (kUa/L)	0.20	±0.1	1.1	0.6-2.3
	Age (y)	−0.15	±0.1	0.9	0.3-2.2
	Constant	−2.42	±1.4		
Milk*					
	Male	0.59	±0.3	1.4	1.1-1.6
Grade of symptoms†	1	0.48	±0.2	1.3	0.8-2.3
	2	2.36	±1.3	1.7	1.0-3.0
	3	7.92	±4.1	4.4	2.1-9.0
	4	8.51	±5.7	5.1	2.5-9.3
	SPT (mm)	0.35	±0.22	1.2	0.7-2.1
	IgE-sIgE (kUa)	−0.006	±0.002	1.1	0.7-2.0
	sIgE (kUa/L)	1.8	±0.8	2.6	1.2-6.8
	Age (y)	−0.15	±0.1	0.9	0.6-1.6
	Constant	−14.61	±9.3		

Data from group B.
*Food being challenged.
†Each of the 4 levels is has its own estimator with no symptoms (0) as the indicator: 1, skin or oral or gastrointestinal or upper respiratory only; 2, upper respiratory and gastrointestinal or 2 systems; 3, lower respiratory or 3 systems; 4, cardiovascular or 4 systems.
‡Corresponds to a 1-unit increase.

Although the generation of 95% positive predictive values for SPTs and serum specific IgE has improved patient treatment, only a quarter of peanut-sensitized children have a wheal of at least 8 mm to peanut.[12] The patients in whom this model would most likely be useful are those with no previous exposure or symptoms of grades 1 and 2, together with a positive SPT and an elevated sIgE. The advantages already outlined of probability values may enhance communication between physician and patient in addition to clinical decision-making. Furthermore, when a patient falls within an immunologic gray area, a probability of clinical allergy can still be calculated for each specific indicator. Although the model is inclusive in that it incorporates all age groups and severity of reactions, it is also sensitive to different food types (through allergen specific weightings) and can cater for potential discrepancies between IgE and SPT levels (through individual weighting from 0 upward on a continuous scale).

We recognize that there are limitations in the current study. As reflects normal clinical practice, there were relatively few DBPCFCs in our cohort, but there was no difference in the usefulness of the algorithm in open versus blind challenges (data not shown). The sample was demographically homogenous, and the majority of cases were children with a mean age of 7 years. In addition, 23% to 44% overall had a positive sIgE but had no previous symptoms to peanut, again reflecting clinical practice. We must also note that, because the quality of extracts used for SPT and serum IgE determinations has an impact in the performance of diagnostic tests, the model developed in this study may not perform similarly with different SPT extracts and *in vitro* IgE testing systems. The values derived from the subtraction of sIgE from total IgE may also be affected by variations produced by the ImmunoCAP system (Phadia, Uppsala, Sweden). Further validation of the model requires the prospective application of the algorithm in other pediatric centers in an experimental pre/post design. It should also be emphasized that we intend this to be a complementary decision tool for clinicians, not a definitive diagnostic tool. It is likely to be more useful in delaying a hospital challenge that is likely to be positive than in encouraging home challenge.

Although more complex than previous models or the use of decision points, this results in a better predictive accuracy for both patients and physicians. Moreover, we are now developing an electronic calculator that will perform the analyses in seconds and that can be used by health professionals at all levels in clinic and in daily practice. The enhanced choice that this will give parents

TABLE V. Comparison of 4 diagnostic methods for peanut, egg, and milk combined

Diagnostic method	Sensitivity	CI	Specificity	CI	Accuracy
Peanut					
sIgE only	75.0	60.5-87.1	46.0	31.4-62.1	61%
SPT only	86.0	73.2-93.9	67.0	50.8-79.5	75%
sIgE and SPT	88.0	74.6-94.8	75.0	54.7-87.9	81%
Cork/Southampton algorithm	96.0	90.0-99.0	90.0	82.0-94.0	93%
Egg					
sIgE only	72.0	56.7-82.2	43.0	25.1-60.2	57%
SPT only	84.0	70.5-93.5	64.0	42.4-80.3	74%
sIgE and SPT	86.0	72.4-93.7	72.0	54.0-85.1	79%
Cork/Southampton algorithm	93.0	88.0-97.0	90.0	83.0-94.0	91%
Milk					
sIgE only	77.0	57.7-91.4	48.0	29.5-63.2	64%
SPT only	85.0	70.5-92.3	63.0	44.9-77.5	74%
sIgE and SPT	86.0	74.9-92.1	74.0	53.7-88.9	79%
Cork/Southampton algorithm	93.0	85.0-98.0	89.0	80.0-94.0	91%

Data from groups B and C.

and patients also has future implications for the improvement of health-related quality of life in families with FA[34-37] and the health economy[38] in cost of performing food challenges.

The model also has applications to new research areas. For example, cost-effectiveness modeling suggests that great benefit would accrue from enhanced precision in delineating a stop/continue point during the maintenance phase in oral immunotherapy. It is possible that this point could be predicted with an enhanced degree of accuracy by using the present combination of variables, which would be available before immunotherapy is started. From this, health-related quality of life in patients and parents undergoing immunotherapy could be improved by higher levels of certainty in relation to length of time to completion.[39] Even though it is possible that some modifications may be needed for other cohorts, for the first time we have demonstrated the advantage of combining clinical variables that are routinely available prechallenge by systematically evaluating their influence on challenge outcome to improve the diagnosis of IgE-mediated FA.

Key messages

- **We developed and validated a predictive model composed of 6 clinical factors.**
- **The model showed a clear advantage in clinical prediction compared to sIgE only, SPT only, and sIgE and SPT.**
- **Findings have implications for delivery of FA–related health care, health-related quality of life, and more economic use of health service resources.**

REFERENCES

1. Anandan C, Gupta R, Simpson CR, Fischbacher C, Sheikh A. Epidemiology and disease burden from allergic disease in Scotland: analyses of national databases. J R Soc Med 2009;102:431-42.
2. Gupta R, Sheikh A, Strachan D, Anderson HR. Increasing hospital admissions for systemic allergic disorders in England: analysis of national admissions data. BMJ 2003;327:1142-3.
3. Gupta R, Sheikh A, Strachan DP, Anderson HR. Time trends in allergic disorders in the UK. Thorax 2007;62:91-6.
4. Simpson CR, Newton J, Hippisley-Cox J, Sheikh A. Incidence and prevalence of multiple allergic disorders recorded in a national primary care database. J R Soc Med 2008;101:558-63.
5. Hughes DA, Mills C. Food allergy: a problem on the increase. Biologist (London) 2001;48:201-4.
6. Hourihane JOB, 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.
7. Roberts S. Challenging times for food allergy tests: a discussion of their risks and benefits. Arch Dis Child 2005;90:564-6.
8. Sampson HA. Utility of food-specific IgE concentrations in predicting symptomatic food allergy. J Allergy Clin Immunol 2001;107:891-6.
9. 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.
10. Sampson HA, Ho DG. Relationship between food-specific IgE concentrations and the risk of positive food challenges in children and adolescents. J Allergy Clin Immunol 1997;100:444-51.
11. Roehr CC, Reibel S, Ziegert M, Sommerfeld C, Wahn U, Niggemann B. Atopy patch tests, together with determination of specific IgE levels, reduce the need for oral food challenges in children with atopic dermatitis. J Allergy Clin Immunol 2001;107:548-53.
12. Roberts GC, Lack G. Diagnosing peanut allergy with skin prick and specific IgE testing. J Allergy Clin Immunol 2005;115:1291-4.
13. Hill DJ, Heine RG, Hosking CS. The diagnostic value of skin prick testing in children with food allergy. Pediatr Allergy Immunol 2004;15:435-41.
14. Garcia-Ara C, Boyano-Martinez T, Diaz-Pena JM, Martin-Munos F, Reche-Frutos M, Martin-Esteban M. Specific IgE levels in the diagnosis of immediate hypersensitivity to cow's milk protein in the infant. J Allergy Clin Immunol 2001;107:185-90.
15. Knight AK, Shreffler WG, Sampson HA, Sicherer SH, Noone S, Mofidi S, et al. Skin prick test to egg white provides additional diagnostic utility to serum egg white-specific IgE antibody concentration in children. J Allergy Clin Immunol 2006;117:842-7.
16. Komata T, Soderstrom L, Borres M, Tachimoto H, Ebisawa M. The predictive relationship of food-specific serum IgE concentrations to challenge outcomes for egg and milk varies by patient age. J Allergy Clin Immunol 2007;119:1272-4.
17. Boyano MT, Garcia-Ara C, Diaz-Pena JM, Martin-Esteban M. Prediction of tolerance on the basis of quantification of egg-white specific IgE antibodies in children with egg allergy. J Allergy Clin Immunol 2002;110:304-9.
18. Hippisley-Cox J, Coupland C, Robson J, Sheikh A, Brindle P. Predicting risk of type 2 diabetes in England and Wales: prospective derivation and validation of QD Score. BMJ 2008;336:1475-82.
19. Hippisley-Cox J, Coupland C, Vinogradova Y, Robson J, Minhas R, Sheikh A, Brindle P. Predicting cardiovascular risk in England and Wales: prospective derivation and validation of QRISK2. BMJ 2008;336:1475-82.
20. Herman WH. Predicting risk for diabetes: choosing (or building) the right model. Ann Intern Med 2009;150:812-4.
21. Hui AY, Chan HL, Wong VW, et al. Identification of chronic hepatitis B patients by a simple non-invasive predictive model. Am J Gastroenterol 2005;100:616-23.

22. Szpurek D, Moszynski R, Smolen A, et al. Using logistic regression analysis in preliminary differential diagnosis of adnexal masses. Int J Gynecol Cancer 2005; 15:817-23.
23. Roberts G, Lack G. Food allergy, getting more out of your skin prick testing. Clin Exp Allergy 2000;30:1495-8.
24. Bock SA, Sampson HA, Atkins FM, et al. Double-blind placebo controlled food challenge (DBPCFC) as an office procedure: a manual. J Allergy Clin Immunol 1988;82:986-97.
25. 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.
26. Nowak-Wegrzyn A, Assa'ad AH, Bahna SL, Bock A, Sicherer SH, Teuber SS. Work Group report: oral food challenge testing. J Allergy Clin Immunol 2009;6:S365-83.
27. Ewan PW, Clark AT. Efficacy of a management plan based on severity assessment in longitudinal and case-controlled studies of 747 children with nut allergy: proposal for good practice. Clin Exp Allergy 2005;35:751-6.
28. Ewan PW, Clark AT. Long-term prospective observational study of patients with peanut and nut allergy after participation in a management plan. Lancet 2001; 357:111-5.
29. Fleischer D. The natural history of tree nut allergy. J Allergy Clin Immunol 2005; 115:S482-523.
30. Menard SW. Applied logistic regression analysis. Sage Publications UK, Inc; 2000.
31. Zhang J, Kai FY. What's the relative risk? A method of correcting the odds ratio in cohort studies of common outcomes. JAMA 1998;280:1690-1.
32. Johnson RA, Wichern DW. Applied multivariate statistical analysis. Upper Saddle River, NJ: Prentice-Hall; 2002.
33. Nakas CT, Yiannoutsos NT. Ordered multiple class receiver operating characteristic (ROC) analysis. Stat Med 2004;23:3437-49.
34. Cohen BL, Noone NS, Munoz-Furlong A. Parental burden in food allergy. J Allergy Clin Immunol 2004;114:1159-63.
35. DunnGalvin A, Gaffney A, Hourihane JO'B. Developmental pathways in food allergy: a new theoretical model. Allergy 2009;64:560-8.
36. DunnGalvin A, de Blok BMJ, Dubois A, Hourihane JO'B. Development and Validation of the Food Allergy Quality of Life – Parent Administered Questionnaire (FAQLQ-PF) for food allergic children aged 0-12 years. Clin Exp Allergy 2008;38:977-86.
37. Flokstra-de Blok BMJ, DunnGalvin A, Vlieg-Boerstra BJ, Oude Elberink JNG, Duiverman EJ, Hourihane JO, et al. Development and validation of the self-administered Food Allergy Quality of Life Questionnaire for adolescents. J Allergy Clin Immunol 2008;122:139-44, 144.e1-2.
38. Miles S, Fordham R, Mills C, Valovirta E, Mugford M. A framework for measuring costs to society of IgE-mediated food allergy. Allergy 2005;8:996-1003.
39. DunnGalvin A, Burks WJ, Dubois AEJ, Chang WC, Hourihane JO'B. Profiling families enrolled in food allergy immunotherapy studies. Pediatrics 2009;124: e503-9.

Algorithm was based on the logistic regression model: $\pi = ez/(1 + ez)$. π is probability, e is mathematical constant ($e = 2.71828$), z is linear combination of x and β, ie $z = \alpha + \beta 1x1 + \beta 2x2 + \beta 3x3 + \ldots + \beta qxq$, where α is constant, x is variable, β is the estimator; then the equation is expressed as the following: $\pi = \exp(\alpha + \beta 1x1 + \beta 2x2 + \beta 3x3 + \ldots + \beta qxq)/[1 + \exp(\alpha + \beta 1x1 + \beta 2x2 + \beta 3x3 + \ldots + \beta qxq)]$. In our model, x is the clinical predictor, β is the estimator and π is the probability of positive versus a negative diagnosis.

Where probability of food allergy diagnosis based on challenge is *P*FA, sex is split into male (1) and female (0) and symptoms are split into 4 grades (Tables I and III).

*patent pending

FIG E1. Logistic regression model and experimental algorithm.

TABLE E1. The proto-algorithm estimators, relative risk ratios, and CIs (phase 1) for all foods ($P < .05$)

Indicators	b-estimate	95% CI	RR	95% CI
High SPT*	4.10	±2.2	2.7	1.8-3.9
High IgE†	2.26	±1.2	1.7	1.2-2.5
Male	1.55	±1.1	1.6	1.1-2.3
Age <7 y	1.50	±1.0	1.0	0.9-1.1
History of reaction	2.40	±1.3	1.5	0.9-1.1
Constant	−3.07			

Data from group A.
*≥8 mm (peanut and milk), ≥7 mm (egg).
†≥15 kUa/L (peanut and milk), ≥7 kUa/L (egg).

TABLE E2. Diagnostic test criteria for peanut, egg, and milk

Criterion	Sensitivity	95% CI	Specificity	95% CI
Peanut AUC, 0.97 (CI, 0.82-0.99); $P < .0001$*				
>0.4	94.1	71.3-99.9	85.0	55.1-98.1
>0.5†	94.1	71.3-99.9	100.0	75.3-100.0
>0.6	83.0	57.0-96.2	100.0	75.3-100.0
Egg AUC, 0.95 (CI, 0.80-0.99); $P < .0001$*				
>0.4	100.0	69.2-100.0	80.0	44.5-98.0
>0.5†	100.0	69.2-100.0	90.0	56.0-98.0
>0.6	90.0	56.0-99.7	90.0	56.0-98.0
Milk AUC, 0.94 (CI, 0.74-0.99); $P < .0001$*				
>0.4	100.0	71.5-100.0	79.0	40.0-97.2
>0.5†	100.0	71.5-100.0	89.0	52.0-99.7
>0.6	91.0	59.0-99.8	89.0	52.0-99.7

Data from group C. AUC, SE, significance, and confidence levels for peanut, milk, and egg.
*Under the nonparametric assumption.
†Null hypothesis: true area ≥0.5