

Development of a tool predicting severity of allergic reaction during peanut challenge



R. Sharon Chinthrajah, MD ^{*,†,‡}; Natasha Purington, MS ^{*,§}; Sandra Andorf, PhD ^{*,†}; Jaime S. Rosa, MD, PhD ^{*,‡}; Kaori Mukai, PhD ^{*,||,¶}; Robert Hamilton, PhD ^{#,*,††}; Bridget Marie Smith, PhD ^{‡‡,|||}; Ruchi Gupta, MD, MPH ^{¶¶,##}; Stephen J. Galli, MD ^{*,||,§}; Manisha Desai, PhD ^{*,§,||}; Kari C. Nadeau, MD, PhD ^{*,†,‡}

^{*} Sean N. Parker Center for Allergy and Asthma Research, Stanford University, Stanford, California

[†] Department of Medicine, Stanford University School of Medicine, Stanford, California

[‡] Department of Pediatrics, Stanford University School of Medicine, Stanford, California

[§] School of Medicine, Quantitative Sciences Unit, Stanford University School of Medicine, Stanford, California

^{||} Department of Pathology, Stanford University School of Medicine, Stanford, California

[¶] Department of Microbiology and Immunology, Stanford University School of Medicine, Stanford, California

[#] Division of Allergy and Clinical Immunology, Johns Hopkins University School of Medicine, Baltimore, Maryland

^{**} Department of Medicine, Johns Hopkins University School of Medicine, Baltimore, Maryland

^{††} Department of Pathology, Johns Hopkins University School of Medicine, Baltimore, Maryland

^{‡‡} Department of Pediatrics, Feinberg School of Medicine, Northwestern University, Chicago, Illinois

^{|||} Center of Innovation for Complex Chronic Healthcare, Edward Hines Jr. VA Hospital, Hines, Illinois

^{¶¶} Department of Pediatrics and Medicine, Feinberg School of Medicine, Northwestern University, Chicago, Illinois

^{##} Ann & Robert H. Lurie Children's Hospital of Chicago, Chicago, Illinois

ARTICLE INFO

Article history:

Received for publication December 20, 2017.

Received in revised form April 18, 2018.

Accepted for publication April 23, 2018.

ABSTRACT

Background: Reliable prognostic markers for predicting severity of allergic reactions during oral food challenges (OFCs) have not been established.

Objective: To develop a predictive algorithm of a food challenge severity score (CSS) to identify those at higher risk for severe reactions to a standardized peanut OFC.

Methods: Medical history and allergy test results were obtained for 120 peanut allergic participants who underwent double-blind, placebo-controlled food challenges. Reactions were assigned a CSS between 1 and 6 based on cumulative tolerated dose and a severity clinical indicator. Demographic characteristics, clinical features, peanut component IgE values, and a basophil activation marker were considered in a multistep analysis to derive a flexible decision rule to understand risk during peanut OFC.

Results: A total of 18.3% participants had a severe reaction (CSS >4). The decision rule identified the following 3 variables (in order of importance) as predictors of reaction severity: ratio of percentage of CD63^{hi} stimulation with peanut to percentage of CD63^{hi} anti-IgE (CD63 ratio), history of exercise-induced asthma, and ratio of forced expiratory volume in 1 second to forced vital capacity (FEV₁/FVC) ratio. The CD63 ratio alone was a strong predictor of CSS ($P < .001$).

Conclusion: The CSS is a novel tool that combines dose thresholds and allergic reactions to understand risks associated with peanut OFCs. Laboratory values (CD63 ratio), along with clinical variables (exercise-induced asthma and FEV₁/FVC ratio) contribute to the predictive ability of the severity of reaction to peanut OFCs. Further testing of this decision rule is needed in a larger external data source before it can be considered outside research settings.

Trial Registration: [ClinicalTrials.gov](https://clinicaltrials.gov) identifier: NCT02103270.

© 2018 American College of Allergy, Asthma & Immunology. Published by Elsevier Inc. All rights reserved.

Reprints: Kari C. Nadeau, MD, PhD, Sean N. Parker Center for Allergy and Asthma Research at Stanford University, Stanford University School of Medicine, 269 Campus Dr, CCSR 3215, Stanford, CA 94305-5101; E-mail: knadeau@stanford.edu.

Disclosures: Authors have nothing to disclose.

Funding Sources: This study was supported by grants U19 AI104209 (Drs Chinthrajah, Mukai, Nadeau, and Galli) and T32 AR050942 (Dr Rosa) from the National Institutes of Health, Glen and Wendy Miller Family Foundation (Drs Gupta and Smith), and the Sean N. Parker Center for Allergy and Asthma Research at Stanford University.

Introduction

Epidemiologic studies indicate that the prevalence of food allergies, especially to peanuts, has increased worldwide in the last few decades.^{1–3} Food and peanut allergies are now estimated to affect approximately 6% to 10% and 1% to 3% of the overall population, respectively. Among children with food allergies, approximately 30% of food allergic individuals have allergies to

multiple foods.^{4–7} Emergency department visits and hospitalizations for food-induced anaphylaxis have significantly increased in recent years,⁸ and in the United States, these events now number approximately 200,000 per year, with further increases expected.⁹ The risk of a severe reaction is higher for those with peanut and tree nut allergies, and it is estimated that more than 50% of these patients experience at least one severe allergic reaction.⁴

The diagnostic approach to food allergies can be complex.^{10,11} Skin prick tests (SPTs) and measurements of food allergen specific IgE (sIgE) are commonly performed; however, these tests are not definitive because they lack specificity and result in a high number of false-positive results.^{10,11} The current gold standard for confirming or excluding a food allergy is the double-blind, placebo-controlled food challenge (DBPCFC).^{10,11} However, DBPCFCs carry the risk of systemic and potentially life-threatening reactions, including severe anaphylaxis, that require hospitalization or intensive care.¹²

The development of clinically useful tools for stratifying patients according to potential risk of severe reactions during peanut oral food challenge (OFC) would therefore be of great clinical value. Predictive models using multivariate logistic regression analysis already have been developed for egg, wheat, and milk allergy, and these are promising in efforts to stratify risk for severity of allergic reaction during peanut OFC with these foods.^{13–15} For peanut allergy, certain biomarkers have been associated with reaction severity during peanut OFCs, such as isolated sensitization to Ara h 2, epitope diversity (ie, extent of IgE binding to the combination of Ara h 1, 2, and 3), and results of changes in levels of basophil CD63 or CD203c staining in basophil activation tests (BATs) on ex vivo activation of basophils with peanut allergen.^{16–19} In this study, we sought to develop a predictive algorithm to identify individuals with peanut allergy who are at higher risk for severe reactions during standardized peanut OFC. To do this, we tested data obtained from participants eligible for a large phase 2 randomized clinical study using peanut oral immunotherapy, including well-characterized clinical data and the participant's responses to validated and standardized peanut OFCs, as well as results of complete blood cell counts, SPTs, BATs, measurements of peanut sIgE and specific IgG4 (sIgG4), total IgE (tIgE), and peanut component resolved diagnostics (CRDs) (sIgE to Ara h 1, 2, 3, 8, and 9). This permitted us to define biomarkers, which we incorporated into a unique food challenge severity score (CSS) and a predictive algorithm to identify those individuals at higher risk for severe reactions to a standardized peanut OFC.

Methods

Study Population and Preliminary Visit

All aspects of the study were authorized by the Stanford University School of Medicine Institutional Review Board and were overseen in collaboration with National Institute of Allergy and Infectious Diseases and the DAIT (Data Safety Monitoring Board). A total of 120 patients between the ages of 7 and 55 years with a convincing history of peanut allergy were recruited to undergo standardized DBPCFCs at the Sean N. Parker Center for Allergy and Asthma Research at Stanford University as part of screening for clinical trial enrollment (Peanut Oral Immunotherapy: Safety, Efficacy and Discovery [POISED]; NCT02103270). A detailed medical history, physical examination, spirometry, SPT, sIgE and tIgE testing, BATs, and CRDs were completed at screening. Patients with a prior history of a peanut allergic reaction that required intubation or eliciting hypotension were excluded, whereas patients with previous reactions to peanut that required epinephrine for other severe symptoms were eligible.

DBPCFCs

Standardized DBPCFCs to peanut and oat (placebo) were performed according to validated guidelines.^{11,20,21} Normal physical examination findings and forced expiratory volume in 1 second (FEV₁) of 80% or greater were confirmed before the start of each OFC. The protocol consisted of 7 escalating doses (5, 20, 50, 100, 100, 100, and 125 mg) of the food protein in flour form, concealed in either applesauce or pudding, ingested by the participant every 15 minutes, as tolerated. Vital signs and pertinent physical examination findings were recorded every 15 minutes or more frequently at the discretion of the clinician. If a positive allergic reaction was observed during OFC,^{20,22} the OFC was immediately discontinued and appropriate treatments were administered as per guidelines and published procedures.^{20,23,24} Reaction severity was scored according to a combination of the Bock and Prachtall grading scale (range, 0–3)^{22,24} and assigned a severity clinical indicator (SCI) of negative or positive. A positive SCI categorized those participants who had any of the following reactions: severe urticaria (involving more than two-thirds of the body surface area), severe or disabling abdominal pain, vomiting and diarrhea in severe abdominal symptoms, any wheezing, any voice changes (airway obstruction), or hypotension. A negative SCI score was assigned for milder reactions. None of the patients in this cohort experienced voice changes or hypotension. To evaluate for sensitivity to peanut protein quantitatively, the cumulative tolerated dose (CTD; the amount of peanut ingested before treating the participant for obstructive symptoms) was also quantified.

CSSs

We developed a clinical severity scoring system, with scores between 1 and 6. This new scoring system combines 2 important factors, the SCI and the CTD, as detailed in Table 1 (eg, CSS 1 indicates a mild reaction on ingestion of a large amount of peanut, and CSS 6 indicates a severe reaction following ingestion of a small amount of peanut).

BATs

BATs were performed as previously described.²⁵ Whole blood was collected in heparin tubes, kept at 4°C for 24 hours, and then stimulated for 30 minutes with peanut (prepared from Byrd Mill peanut flour) as previously described²⁵ (peanut concentrations of 1,000, 100, 10, 1, and 0.1 ng/mL) at 37°C. Anti-IgE, interleukin 3, or RPMI media alone were used as controls. Cells were stained with CD203c, CD63, CD123, and HLA-DR. Basophils were gated as CD123⁺HLA-DR⁺. Basophil expression of CD63 and CD203c was analyzed with FlowJo (Ashland, OR). In total, we assessed 15 BAT variables as presented in eTable 1. As previously reported, the following forms of CD203c and CD63 were evaluated: the ratio of CD203c stimulation with peanut to CD203c RPMI (negative control), percentage of CD63^{hi} stimulation with peanut, and the ratio of percentage of CD63^{hi}

Table 1
Distribution of Challenge Severity Score

Cumulative tolerated Dose, mg	Severity challenge indicator ^a	Challenge severity score	No. (%) of patients
175–375	Negative	1	15 (12.5)
25–75	Negative	2	44 (36.7)
<25	Negative	3	39 (32.5)
175–375	Positive	4	9 (7.5)
25–75	Positive	5	12 (10)
<25	Positive	6	1 (0.8)

^aThe severity challenge indicator is a measure of reaction severity.

stimulation with peanut to percentage of CD63^{hi} anti-IgE.²⁶ The ratio of the percentages of CD63^{hi} basophils after stimulation with varying concentrations of peanut vs anti-IgE (CD63 ratio) were also included to assess the magnitude of the response.²⁵

Statistical Analysis and Variable Selection Process

Descriptive statistics, such as means, medians, SDs, and interquartile ranges, were used to compare baseline characteristics by CSS. Because only one patient was assigned a CSS of 6, scores 5 and 6 were collapsed when reporting demographics. Graphical tools were used to assess distributional aspects. Log transformations were applied when normality assumptions were violated. Kruskal-Wallis and Fisher exact tests were used to draw comparisons on CSS by continuous and categorical characteristics, respectively.

Our variable selection decisions are outlined in eFigure 1. Among the 94 available clinical, laboratory, and demographic features, self-reported overall and detailed allergic rhinitis variables were excluded ($n = 7$) because of nonconcordance against clinical testing (eTable 2) because our interest was in arriving at an objective decision rule. Furthermore, absolute counts were included for analysis from the complete blood cell count (excluded $n = 4$ variables).

We then took a 3-staged approach to arriving at our model. In the first stage, 1 BAT variable of 15 possible BAT variables was pre-selected using an area under the curve (AUC) because most laboratories would have resources to measure 1 BAT (CD203c or CD63) after stimulation with 1 concentration of peanut. Logistic regression models were fit to SCI as a function of each log-transformed BAT, and the BAT with the highest AUC was selected to be included as a potential predictor in the second stage of the process.

In the second stage, variables of importance for CSS were identified from the subset of 69 clinical, laboratory, and demographic features, including the preselected BAT from the first stage, using conditional random forest (RF) approaches.²⁷ Conditional RF methods use a permutation-based significance testing procedure to select variables as opposed to the default approach that uses the Gini statistic, avoiding the bias of selecting variables that have many possible splits found in traditional RF methods.²⁸ Variables were chosen as

relevant if their importance value was above the absolute value of the lowest negative-scoring variable.²⁹

In the third stage, a clinically interpretable decision rule using classification and regression trees (CART) was derived from the subset of variables deemed important via RF in the second stage. CART is a tree-based learning technique for classifying observations.³⁰

To compare the results derived from variable selection via RF and prediction via CART to more widely used statistical methods, the least absolute shrinkage and selection operator (LASSO) was applied to the 69 preselection features. LASSO is a regression analysis method that performs both variable selection and regularization to gain higher predictive ability of the final fitted model.³¹ The root mean square error (RMSE) from CART and the LASSO regression model were compared.

Two supplementary analyses were also performed. First, CSS was fit as a function of log-transformed CD63 ratio alone in a linear regression model, and the change in CSS and P value were reported. Then, 4 logistic regression models were fit to an indicator for whether a subject reacted to the peanut OFCs at or before 175 mg as a function of the CD63 ratio, Ara h 2, sIgE, and SPT, respectively. A cutoff of 175 mg of peanut protein was chosen to be comparable to other published literature, looking at a dose less than a full peanut (240 mg).^{32,33} The Youden index was used to determine a threshold for each predictor, which maximized the AUC.³⁴ The threshold and AUC (95% confidence interval [CI]) were reported for each of the 5 predictors.

The entire 3-stage process and LASSO were internally validated using repeated 10-fold cross-validation as commonly performed.³⁵ All analyses were conducted using R, version 3.2.2.^{36,37} Additional details of statistical methods can be found in the eMethods in the Supplement.

Results

Total Study Population

The median participant age was 11 years. Baseline demographics are reported in Table 2. Eighty-two percent of the cohort was younger than 18 years; 33% of the cohort was female. Participants

Table 2
Baseline Demographics^a

Demographic	Challenge severity score					Total (N = 120)	P value ^c
	1 (n = 15)	2 (n = 44)	3 (n = 39)	4 (n = 9)	5 and 6 ^b (n = 13)		
Age at enrollment, median (IQR), y	9 (4)	11 (7)	11 (13)	13 (18)	9 (2)	11 (7)	.10
Child (<18 years old)	14 (93.3)	38 (86.4)	28 (71.8)	6 (66.7)	12 (92.3)	98 (81.7)	.15
Female	6 (40.0)	8 (18.2)	17 (43.6)	4 (44.4)	4 (30.8)	39 (32.5)	.10
Race							
White	14 (93.3)	34 (77.3)	21 (53.9)	7 (77.8)	9 (69.2)	85 (70.8)	.04
Asian	3 (20.0)	11 (25.0)	22 (56.4)	2 (22.2)	4 (30.8)	42 (35.0)	.02
Black or African American	0	3 (6.8)	0	0	0	3 (2.5)	.48
Native Hawaiian or Pacific Islander	0	0	1 (2.6)	0	0	1 (0.8)	.63
Not Hispanic or Latino	15 (100.0)	42 (95.5)	39 (100.0)	9 (100.0)	12 (92.3)	117 (97.5)	.43
BMI, mean (SD)	17.8 (3.2)	18.6 (4.7)	20.3 (6.0)	20.5 (3.4)	17.7 (2.5)	19.1 (4.8)	.12
No. of food allergies, mean (SD)	3.1 (2.5)	3.9 (2.5)	4.0 (2.8)	3.0 (2.3)	3.6 (2.7)	3.7 (2.6)	.71
Patient with food allergy to							
>1 food	11 (73.3)	33 (75.0)	27 (69.2)	5 (55.6)	8 (61.5)	84 (70.0)	.72
Asthma and allergy history							
Asthma	6 (40.0)	26 (59.1)	30 (76.9)	7 (77.9)	11 (84.6)	80 (66.7)	.04
Age at onset, mean (SD), y	2.9 (1.9)	3.6 (2.7)	3.7 (3.2)	4.5 (3.3)	3.2 (2.1)	3.6 (2.8)	.91
AD	11 (73.3)	33 (75.0)	28 (71.8)	4 (44.4)	11 (84.6)	87 (72.5)	.37
Age at onset, mean (SD), y	0.8 (1.4)	0.9 (1.2)	2.6 (5.8)	0.5 (0.4)	0.6 (0.8)	1.4 (3.5)	.83
AR	11 (73.3)	34 (77.3)	28 (71.8)	8 (88.9)	9 (69.2)	90 (75.0)	.84
Age at onset, mean (SD), y	5.5 (2.2)	5.1 (4.0)	5.7 (3.6)	5.9 (3.3)	3.3 (1.6)	5.2 (3.5)	.19

Abbreviations: AD, atopic dermatitis; AR, allergic rhinitis; BMI, body mass index (calculated as weight in kilograms divided by the square of height in meters).

^aData are presented as number (percentage) of patients unless otherwise indicated.

^bCollapsed scores 5 and 6 because there was only 1 patient in the challenge severity score 6 category.

^cKruskal-Wallis and Fisher exact tests were used to calculate P values for continuous and categorical variables by challenge severity score, respectively.

had comorbid atopic conditions, including other food allergies (70%), asthma (67%), atopic dermatitis (73%), and allergic rhinitis (AR) (75%).

Characteristics of the Study Population and CSS Outcomes

OFCs were discontinued after objective signs or symptoms of an allergic reaction were observed. Symptoms ranged from mild urticaria to more severe symptoms, such as wheezing or diffuse urticaria. The mean (SD) CSS was 2.7 (1.2). Most patients (82%) were assigned a CSS of 1 to 3, indicating less severe reactions in the DBPCFC (Table 1). All challenge reactions were treated with no sequelae. We observed no biphasic reactions. Eighteen children (aged 7–17 years) and 4 adults (aged 18–53 years) had more severe scores (CSS >3). Of the 22 participants with a more severe CSS, 50% had wheezing, 45% had severe abdominal pain, and 5% had severe urticaria (eTable 3). Each of the 120 participants developed an objective allergic reaction to a CTD of 500 mg or less of peanut protein during their DBPCFC. The median CTD was 75.0 mg.

A higher proportion of white patients were in the lowest CSS group (CSS 1, 93.3%; $P = .04$) (Table 2). There were more patients with a history of asthma in the highest CSS group (CSS 5–6, 84.6%; $P = .04$); however, per our protocol, DBPCFCs were only conducted if FEV₁ was 80% or greater. Consequently, the mean FEV₁ was 95.5%, and the mean ratio of FEV₁ to forced vital capacity FVC was 84% (Table 3).

There was no evidence to suggest that the peanut SPT mean wheal size, sIgE, tIgE, CRD tests, or complete blood cell counts were significantly different across CSS levels (Table 3).

Preselection of BAT

The percentage of CD63^{hi} response to 10 ng/mL of peanut divided by percentage of CD63^{hi} anti-IgE (CD63 ratio) was chosen among all BATs to be further considered in RF based on having the highest AUC (0.67) (eTable 1).

Predictive Algorithm for the CSS

Of the set of 69 features, 4 features were identified as important predictors of the CSS using RF selection and the threshold criteria previously described (Fig 1). The CD63 ratio, history of exercise-induced asthma (EIA), FEV₁/FVC ratio at DBPCFC, and absolute blood basophil count were then entered into a CART model, with 3 appearing in the final decision rule (Fig 2). Our analysis demonstrated that the CD63 ratio, a history of EIA, and the FEV₁/FVC ratio at DBPCFC are important predictors of reaction severity. Participants with a CD63 ratio greater than 5.49 were at highest risk of experiencing a severe reaction with a median CSS of 4. For CD63 ratios between 0.47 and 5.49, a history of EIA may increase the risk of a severe reaction. For CD63 ratios of 0.47 or less, an FEV₁/FVC

Table 3
Clinical and Laboratory Characteristics

Characteristic	Challenge severity score					Total (N = 120)	P value ^b
	1 (n = 15)	2 (n = 44)	3 (n = 39)	4 (n = 9)	5 and 6 ^a (n = 13)		
FEV ₁ at DBPCFC, mean (SD), %	94.5 (10.1)	97.5 (12.9)	94.7 (11.2)	91.1 (7.6)	94.7 (13.1)	95.5 (11.7)	.72
FEV ₁ /FVC at DBPCFC, mean (SD)	0.86 (0.07)	0.86 (0.07)	0.82 (0.07)	0.80 (0.08)	0.82 (0.06)	0.84 (0.07)	.02
Peanut wheal size, mean (SD), mm	12.2 (7.45)	13.3 (6.4)	13.4 (5.6)	12.7 (5.8)	13.5 (7.1)	13.3 (6.1)	.96
White blood cell count, mean (SD), K/ μ L							
Absolute basophil count	0.05 (0.03)	0.04 (0.03)	0.04 (0.02)	0.04 (0.02)	0.04 (0.01)	0.04 (0.02)	.90
Absolute neutrophil count	3.3 (1.2)	2.9 (1.3)	3.2 (1.4)	3.7 (1.1)	3.2 (1.8)	3.1 (1.4)	.17
Absolute lymphocyte count	2.4 (0.7)	2.2 (0.6)	2.2 (0.6)	2.4 (0.8)	2.2 (0.5)	2.2 (0.6)	.78
Absolute monocyte count	0.5 (0.2)	0.5 (0.2)	0.5 (0.2)	0.5 (0.1)	0.5 (0.2)	0.5 (0.2)	.92
Absolute eosinophil count	0.5 (0.2)	0.4 (0.3)	0.4 (0.2)	0.4 (0.3)	0.4 (0.2)	0.4 (0.2)	.61
CD203c basophil activation test result, mean (SD) ^c							
PF 0.1	1.28 (0.40)	1.14 (0.15)	1.20 (0.18)	1.42 (0.24)	1.17 (0.17)	1.20 (0.22)	.01
PF 1	1.28 (0.27)	1.28 (0.22)	1.34 (0.25)	1.49 (0.27)	1.38 (0.24)	1.32 (0.24)	.09
PF 10	1.52 (0.52)	1.66 (0.36)	1.66 (0.40)	1.88 (0.42)	1.72 (0.40)	1.67 (0.40)	.11
PF 100	1.74 (0.66)	1.85 (0.49)	1.86 (0.51)	2.02 (0.44)	1.91 (0.36)	1.86 (0.50)	.44
PF 1,000	1.73 (0.63)	1.82 (0.48)	1.75 (0.47)	1.97 (0.44)	1.86 (0.28)	1.80 (0.47)	.39
CD63 ^{hi} basophil activation test result, median (IQR) ^c							
PF 0.1	0.48 (1.02)	0.60 (1.24)	0.54 (0.81)	0.45 (0.66)	0.90 (1.18)	0.55 (1.25)	.75
PF 1	0.31 (0.62)	1.66 (4.35)	2.21 (3.24)	2.06 (3.35)	4.83 (5.83)	1.92 (4.20)	.01
PF 10	3.33 (6.63)	10.70 (25.97)	14.20 (18.37)	12.60 (15.80)	24.40 (19.80)	11.90 (20.83)	.004
PF 100	10.50 (25.61)	24.00 (31.68)	20.50 (30.99)	9.59 (18.70)	36.80 (22.50)	22.40 (29.48)	.02
PF 1,000	6.82 (22.54)	14.85 (25.04)	13.50 (13.63)	11.30 (17.68)	14.90 (16.79)	13.80 (18.85)	0.24
PF10/CD63 ^{hi} anti-IgE	0.46 (1.59)	0.70 (1.42)	1.51 (1.84)	2.42 (7.57)	1.93 (5.07)	1.21 (1.82)	.01
Antibodies, median (IQR)							
Peanut IgG4, μ g/L	430.0 (1,870.0)	1,110.0 (2,345.0)	640.0 (930.0)	450.0 (620.0)	1600.0 (1 m220.0)	830 (1,600)	.05
Peanut IgE, ng/L	77.1 (277.7)	178.9 (412.5)	193.0 (478.0)	21.9 (217.3)	383.1 (614.6)	192.9 (458.0)	.13
Total IgE, ng/L	1,342.0 (2,364.4)	1,321.3 (2,526.6)	1,149.2 (1,766.6)	592.9 (1,461.6)	1,351.8 (2,249.7)	1,284.7 (2,379)	.49
Peanut components, median (IQR), ng/L							
Ara h 1	27.3 (166.3)	21.4 (115.1)	57.1 (191.2)	3.6 (15.5)	179.6 (212.5)	34.9 (177.8)	.18
Ara h 2	39.3 (107.8)	70.6 (209.5)	87.1 (197.9)	12.8 (140.9)	196.7 (216.9)	85.0 (207.1)	.09
Ara h 3	0.9 (37.2)	3.2 (22.2)	9.7 (32.2)	1.1 (5.3)	25.9 (39.3)	5.7 (38.6)	.31
Ara h 8	0.0 (2.5)	0.0 (2.5)	0.0 (1.6)	0.0 (0.6)	0.0 (0.4)	0.0 (1.8)	.95
Ara h 9	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	.16

Abbreviations: BAT, basophil activation test; DBPCFC, double-blind, placebo-controlled food challenge; FEV₁, forced expiratory volume in 1 second; FVC, forced vital capacity; IQR, interquartile range.

^aCollapsed scores 5 and 6 because there was only 1 patient was in the challenge severity score 6 category.

^bKruskal-Wallis and Fisher exact tests were used to calculate P values for continuous and categorical variables by challenge severity score, respectively.

^cPF 100 indicates values for basophils stimulated with peanut flour at the 100-ng/mL concentration and so on.

Table 4
Baseline Characteristic Differences by CART Terminal Nodes

Characteristic	Node ^a				
	1 (n = 11)	2 (n = 23)	3 (n = 25)	4 (n = 54)	5 (n = 7)
Continuous Variables, Mean (SD)					
Peanut SPT result	4.27 (1.31)	5.78 (2.81)	4.88 (1.84)	5.08 (2.12)	5.50 (1.89)
FEV ₁ at DBPCFC	92.18 (9.53)	99.4 (13.87)	91.24 (8.85)	96.5 (11.59)	94.57 (14.30)
FEV ₁ /FVC ratio at DBPCFC	0.75 (0.05)	0.88 (0.05)	0.81 (0.07)	0.85 (0.06)	0.80 (0.10)
No. of food allergies	3.00 (3.22)	4.00 (2.59)	4.36 (2.66)	3.65 (2.43)	2.57 (2.15)
BMI	18.63 (3.91)	20.12 (6.79)	19.44 (4.48)	18.32 (3.89)	21.09 (6.24)
CD63 ratio	0.20 (0.16)	0.20 (0.17)	1.73 (1.45)	2.01 (1.12)	10.72 (4.41)
Continuous Variables, Median (IQR)					
Age at enrollment, y	11 (9–16)	9 (8–14.5)	12 (9–16)	10.5 (8–14)	13 (10.5–2.5)
Peanut diagnosis age, y	2 (1–3)	2 (1–3)	1 (0.8–2)	1.5 (1–2)	1.5 (1–3.5)
Peanut allergy duration, y	9.2 (6.8–13.3)	8 (6–13)	9.3 (7.5–13.3)	8.7 (7–12)	11.5 (9.8–29.5)
Ara h 1, ng/L	1.2 (0–3.6)	0.6 (0–26.0)	52.5 (3.0–135)	82.6 (10.2–224)	134.7 (60.0–232)
Ara h 2, ng/L	6.3 (1.8–14.0)	8.9 (1.4–64.8)	75.6 (25.1–169)	152.9 (52.6–409)	193.5 (104.1–439)
Ara h 3, ng/L	0 (0–1.3)	0.4 (0–3.0)	4.1 (0.5–40.3)	18.2 (3.0–46.4)	25.9 (14.0–43.4)
Ara h 8, ng/L	0 (0–1.3)	0.3 (0–12.2)	0 (0–0.7)	0 (0–1.5)	0.4 (0.1–2.2)
Ara h 9, ng/L	0 (0–12.1)	0 (0–0.2)	0	0	0.5 (0–2.6)
Peanut IgG4, µg/L	480 (255–715)	360 (210–1,520)	940 (350–1,600)	910 (402–2,388)	2,460 (1,055–4,535)
Peanut IgE, ng/L	14.4 (2.9–65.8)	47.3 (8.5–219)	192.8 (36.1–312)	316.0 (147–767)	446.5 (218–1,252)
Total IgE, ng/L	1,342.0 (708–3,370)	1,193.2 (603–2,302)	1,281 (683–2,403)	1,339.6 (699–3,226)	3,991.8 (473–8,808)
Categorical Variables, %					
Male	81.8	82.6	68	61.1	42.9
white	90.9	73.9	64	68.5	71.4
Asian	18.2	43.5	32	37	28.6
Black	0	4.3	4	1.9	0
Non-Hispanic	100	95.7	100	96.3	100
Peanut diagnosis only	54.5	17.4	24	29.6	57.1
Tree nut diagnosis	36	70	68	59	43
History of asthma	63.6	52.2	100	53.7	100
History of AR	72.7	87	84	66.7	71.4
History of AD	63.6	73.9	84	72.2	42.9

Abbreviations: AD, allergic dermatitis; AR, allergic rhinitis; BMI, body mass index (calculated as weight in kilograms divided by the square of height in meters); CART, classification and regression trees; DBPCFC, double-blind, placebo-controlled food challenge; FEV₁, forced expiratory volume in 1 second; FVC, forced vital capacity; IQR, interquartile range; SPT, skin prick test.

^aThe cross-validation root mean square error was 1.16.

ratio of 0.8 or less may result in slightly higher risk of a severe reaction compared with participants with higher FEV₁/FVC ratios. To assess the validity of this model, a repeated 10-fold cross-validation was used and yielded an RMSE of 1.16 for the 3-staged modeling process.

Features stratified by the terminal nodes of the CART are given in Table 4. Notably, node 5, which has a median CSS of 4, is characterized by a population of participants with the highest IgE values (total, peanut, Ara h 1, 2, 3, 8, and 9) and highest CD63 ratios.

To determine the predictive ability of the CD63 ratio using standard regression techniques without the presence of other features, we also fit a linear regression model for CSS as a function of log-transformed CD63 ratio. We found that the CD63 ratio was a significant predictor of CSS alone, with a cross-validated RMSE of 1.08 (eTable 4).

LASSO Regression

After including all 69 preselected features into a LASSO model, history of EIA and the CD63 ratio were selected as statistically significant predictors of CSS (eTable 5). The repeated 10-fold cross-validation RMSE for this model was 2.47.

Discussion

The DBPCFC is currently the gold standard for diagnosing food allergies and is most often performed in a research setting where the patients are likely to be allergic. In these allergic patients, although a positive outcome confirming food allergy is likely on DBPCFC, severe symptoms can be life-threatening.^{10–12} Identification

of possible prognostic indicators that are associated with severe outcomes during a peanut OFC may aid in appropriate risk stratification and management when a OFC is being considered by a trained professional. Previous studies for peanut allergy have focused on clinical features and certain tests, such as SPTs,³⁸ sIgE,¹¹ or CRDs, that can assist in distinguishing between the absence and presence of a food allergy.³⁹ Many investigators have already shown that sIgE to Ara h 2, as opposed to sIgE to whole peanut, has greater specificity for predicting reactivity during peanut OFC (ranging from 85% to 95%, depending on the cohort).^{40–46}

Few studies have examined biomarkers to predict the severity of a reaction. Our approach was to examine demographics and clinical and laboratory data in relation to a potential new predictor tool using a CART statistical approach. BATs can be a useful tool in predicting the severity of reaction during peanut OFC in peanut-sensitized individuals,^{17,19,47} despite earlier studies to the contrary.⁴⁸ Other authors have published evidence on severity scores; however, these did not incorporate history of EIA, CD63 ratio, or FEV₁/FVC ratio at DBPCFC.^{17,19} Some groups have also tried to use clinical characteristics to predict increased adverse events during oral immunotherapy (OIT). One study found baseline allergic rhinitis and peanut SPT wheal size were significant predictors of higher overall adverse event rates. SPT wheal size predicted increased gastrointestinal adverse events, and AR predicted increased systemic reactions during oral immunotherapy.⁴⁹

Our study was a prospective study of a well-characterized population of peanut allergic children and adults who underwent DBPCFCs. Patients with a prior history of severe reactions to peanut were included, except for those patients with a history of a peanut allergic reaction that required intubation or eliciting hypotension.

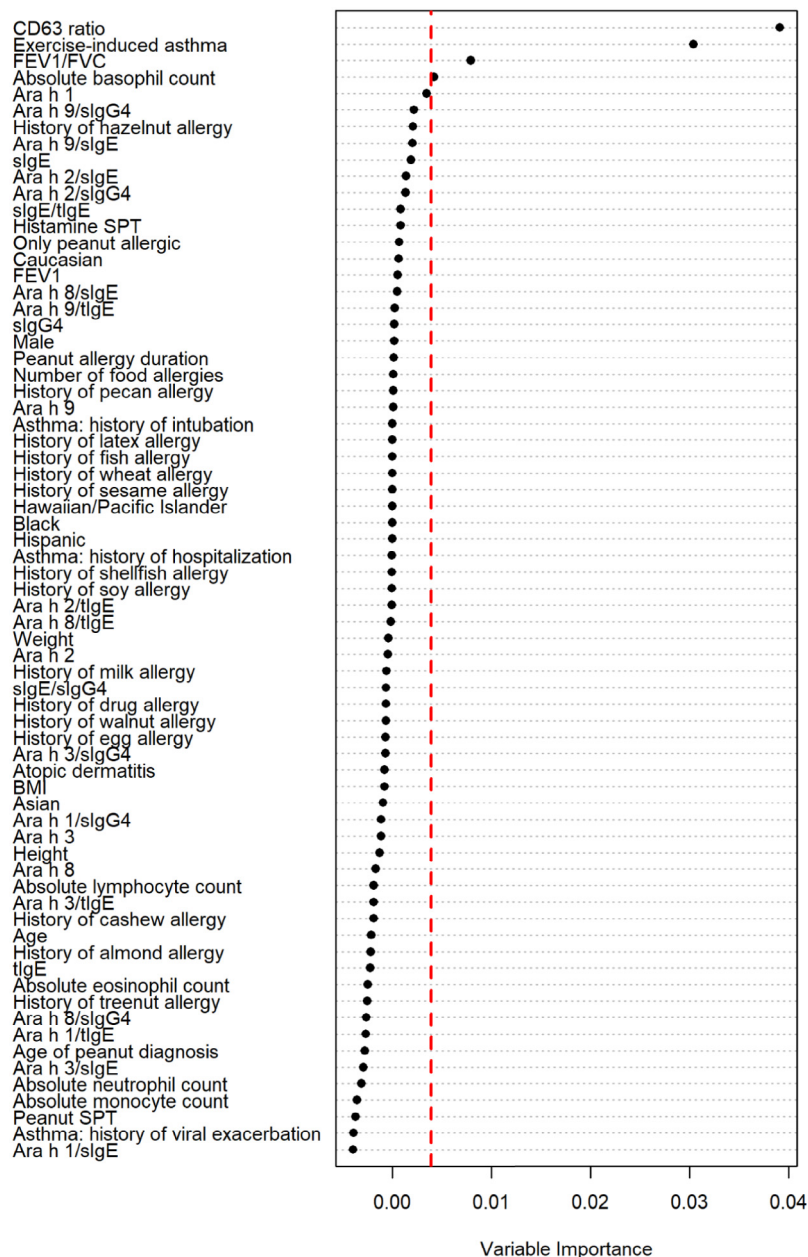


Figure 1. Random forest plot highlights variables that were deemed important: the CD63 ratio, a history of exercise-induced asthma (EIA), ratio of forced expiratory volume in 1 second (FEV1) to forced vital capacity (FVC) at double-blind, placebo-controlled food challenge (DBPCFC), and the absolute blood basophil count. Included variables are listed on the y-axis ($n = 69$), with corresponding variable importance on the x-axis. Variables were deemed important predictors of the severity score if their importance value was larger than the absolute value of the lowest negative-scoring variable (to the right of the dashed red line). BMI, body mass index (calculated as weight in kilograms divided by the square of height in meters); slgE, specific IgE; slgG4, specific IgG4; SPT, skin prick test; tlgE, total IgE.

We aimed to develop a possible method to predict severe allergic reactions because management of these reactions is labor-intensive, requires increased observation, and can be life-threatening.

There is no universal consensus to characterize what constitutes a severe reaction during peanut OFC. Although other grading systems to assign ordinal values to severity of individual symptoms have been developed, not all symptoms should be assessed based on similar grading. For instance, severe pruritus grade 3 has a different clinical significance compared with wheezing grade 3.⁵⁰ Because there is no universal agreement on the criteria for administering epinephrine during peanut OFC, the SCI was created to characterize symptoms that most practitioners would treat with epinephrine. The CSS is novel in that it not only groups similarly severe

symptoms in a more clinically meaningful way using SCI, but it incorporates the amount of food allergen ingested.^{20,51} Furthermore, CSS provides multilevel phenotyping of clinical symptom severity on an ordinal scale that captured the allergic features of the reaction. For example, a patient who is able to consume the equivalent of 1 peanut (240 mg) and develops limited hives would be assigned a CSS of 1, whereas a patient who can only consume a low dose of peanut (5 mg) and develops wheezing would receive a CSS of 6. Clinically, these are important distinctions that have heretofore not been successfully characterized.

To our knowledge, this is the first study to use machine learning procedures (ie, RF and CART), which incorporate a combination of clinical and laboratory-based parameters to develop an internally

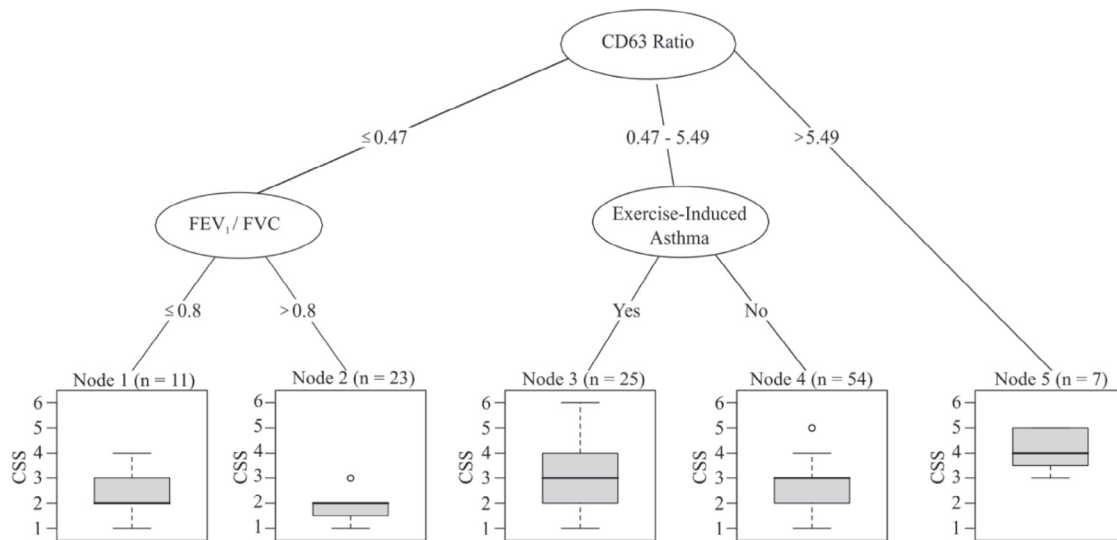


Figure 2. Classification and regression trees were used to identify 3 features to classify participants and assess risk of severity of reaction (challenge severity score [CSS]): the CD63 ratio, a history of exercise-induced asthma, and ratio of forced expiratory volume in 1 second (FEV₁) to forced vital capacity (FVC) at double-blind, placebo-controlled food challenge.

validated *decision rule* that could potentially be used by clinicians to better understand a patient's risk of having a severe allergic reaction during peanut OFC. Unlike CART, previous models using standard regression procedures are not well equipped to *discover*, in an agnostic fashion, complex interactions among features that are not prespecified using dichotomous branching logic and optimal cut-off values among the features.

The basophil activation ratio of CD63 was prominently featured in our predictive model, similar to work done by Santos et al.¹⁷ The ratio of CD63-positive basophils after stimulation with peanut vs anti-IgE was selected in the CART model in 3 different ranges to be an important initial discriminator of the CSS. A CD63 ratio of greater than 5.49 identified a subset of patients who had the highest predicted CSS. The absolute basophil count was chosen as 1 of 4 predictors in the RF model but then excluded in the CART model, likely because no significant cut points could be identified to risk stratify participants.

In the predictive model (Fig 2), not only was a history of EIA important, but the CART analysis also selected the FEV₁/FVC ratio on the day of peanut OFC. Patients with asthma are known to have a higher risk for more severe allergic reactions.^{52,53} Thus, it is not recommended to conduct a peanut OFC in a patient with respiratory symptoms or an active respiratory infection. Current research protocols specify an FEV₁/FVC ratio cutoff of 80% or higher before the start of peanut OFCs; this should also be considered for peanut OFCs in the allergist's office as well.

Interestingly, antibody tests, including sIgE, sIgG4, sIgE/sIgG4 ratios, and peanut CRDs were similar in peanut allergic participants who had mild or severe reactions and thus were not identified as important variables relative to the ones ultimately selected, particularly, the CD63 ratio. However, our data indicate that the participants with the highest median CSS of 4 (Fig 2 and Table 4) had on average higher IgEs (total, peanut, and CRDs) compared with participants in other nodes. Even though these biomarkers had higher expression in node 5 compared with other nodes, they were not as predictive as the CD63 ratio (Fig 1). In addition, we fit several logistic regression models to determine whether any known predictors of severity (ie, SPT, Ara h 2, and sIgE) were better than the CD63 ratio and found that the CD63 ratio was the superior biomarker (eTable 6).

With more clinical trials studying different therapies in peanut allergic individuals, our tool could be used to aid the researcher in phenotyping those patients at risk for a more severe reaction at DBPCFC. We propose to currently apply this predictive decision rule in clinical research only. However, if a patient has a concurrent illness,⁴⁹ these coexisting conditions can increase the likelihood of positive outcomes during peanut OFC. Our DBPCFCs were performed in controlled settings in which we excluded illnesses or other events known to put participants at risk; these are the same precautions one would take in an outpatient clinic or clinical research study.

To provide a comparison between CART and a more widely used method, we fit a LASSO regression model using the 69 features chosen in the preselection process. We found that a history of EIA and a higher CD63 ratio were significantly associated with higher CSS. Furthermore, the RMSE of the LASSO model was higher compared with that of the CART (2.44 vs 1.16), suggesting that the CART method was a superior predictive model. Importantly, while the cross-validated RMSE of the linear regression model that only includes the CD63 ratio was lower than that obtained using CART (1.08 vs 1.16), interpreting results from the CART model is arguably more intuitive than interpretation resulting from a standard regression model. From the regression model, a 10% increase in the CD63 ratio was associated with a 0.02-unit increase in CSS (eTable 4). Conversely, using CART, we see how CSS is predicted to increase when the CD63 ratios exceed a threshold.

We present early findings from a well-characterized, relatively large, peanut allergic cohort that underwent standardized OFCs as part of a large, prospective phase 2 study. The decision rule developed using the CSS may be considered as a predictive aid in clinical research settings to determine the risk of severe reaction during peanut OFC. However, further testing of this decision rule is needed in a larger validation cohort. One of the limitations of this study is that this study did not include children (who may have no clear history of EIA), and this model cannot, at this time, be applied in younger children. The severity predictors we discovered in our CART and RF analyses highlight the importance of the CD63 ratio, which may aid the researcher in deciding to conduct peanut OFCs.

Acknowledgments

We thank all the participants and their families who are part of this phase 2 study. We also thank Dr Vanitha Sampath for her critical read, review, and edits, Dr Robert Tibshirani for his review and input on the statistical methods, and Dr Marshall Plaut for his critical read, review, and input to the manuscript.

Supplementary Data

Supplementary data related to this article can be found at <https://doi.org/10.1016/j.anai.2018.04.020>.

References

- [1] Rinaldi M, Harnack L, Oberg C, Schreiner P, St Sauver J, Travis LL. Peanut allergy diagnoses among children residing in Olmsted County, Minnesota. *J Allergy Clin Immunol*. 2012;130:945–950.
- [2] Sicherer SH, Munoz-Furlong A, Godbold JH, Sampson HA. US prevalence of self-reported peanut, tree nut, and sesame allergy: 11-year follow-up. *J Allergy Clin Immunol*. 2010;125:1322–1326.
- [3] Jackson KD, Howie LD, Akinbami LJ. Trends in allergic conditions among children: United States, 1997–2011. *NCHS Data Brief*. 2013;1–8.
- [4] Gupta RS, Springston EE, Warrier MR, et al. The prevalence, severity, and distribution of childhood food allergy in the United States. *Pediatrics*. 2011;128:e9–e17.
- [5] Solter L, Ben-Shoshan M, Harrington DW, et al. Overall prevalence of self-reported food allergy in Canada. *J Allergy Clin Immunol*. 2012;130:986–988.
- [6] Nwaru BI, Hickstein L, Panesar SS, Roberts G, Muraro A, Sheikh A. Prevalence of common food allergies in Europe: a systematic review and meta-analysis. *Allergy*. 2014;69:992–1007.
- [7] Osborne NJ, Koplin JJ, Martin PE, et al. Prevalence of challenge-proven IgE-mediated food allergy using population-based sampling and predetermined challenge criteria in infants. *J Allergy Clin Immunol*. 2011;127:668–676, e661–662.
- [8] Dyer LA, Lau CH, Smith TL, Smith BM, Gupta RS. Pediatric emergency department visits and hospitalizations due to food-induced anaphylaxis in Illinois. *Ann Allergy Asthma Immunol*. 2015;115:56–62.
- [9] Clark S, Espinola J, Rudders SA, Banerji A, Camargo CA Jr. Frequency of US emergency department visits for food-related acute allergic reactions. *J Allergy Clin Immunol*. 2011;127:682–683.
- [10] American College of Allergy, Asthma, & Immunology. Food allergy: a practice parameter. *Ann Allergy Asthma Immunol*. 2006;96:S1–S68.
- [11] Sampson HA, Aceves S, Bock SA, et al. Food allergy: a practice parameter update-2014. *J Allergy Clin Immunol*. 2014;134:1016–1025, e1043.
- [12] Cianferoni A, Khullar K, Saltzman R, et al. Oral food challenge to wheat: a near-fatal anaphylaxis and review of 93 food challenges in children. *World Allergy Organ J*. 2013;6:14.
- [13] Sugiura S, Matsui T, Furuta T, Sasaki K, Kando N, Ito K. Development of a prediction model for severe wheat allergy. *Pediatr Allergy Immunol*. 2018;29:93–96.
- [14] Sugiura S, Matsui T, Nakagawa T, et al. Development of a prediction model of severe reaction in boiled egg challenges. *Allergol Int*. 2016;65:293–299.
- [15] Sugiura S, Sasaki K, Matsui T, Nakagawa T, Kando N, Ito K. Development of a prediction model for a severe reaction in cow's milk challenges. *Allergol Int*. 2017;66:493–494.
- [16] Flinterman AE, Knol EF, Lencer DA, et al. Peanut epitopes for IgE and IgG4 in peanut-sensitized children in relation to severity of peanut allergy. *J Allergy Clin Immunol*. 2008;121:737–743, e710.
- [17] Santos AF, Du Toit G, Douiri A, et al. Distinct parameters of the basophil activation test reflect the severity and threshold of allergic reactions to peanut. *J Allergy Clin Immunol*. 2015;135:179–186.
- [18] Shreffler WG, Beyer K, Chu TH, Burks AW, Sampson HA. Microarray immunoassay: association of clinical history, in vitro IgE function, and heterogeneity of allergenic peanut epitopes. *J Allergy Clin Immunol*. 2004;113:776–782.
- [19] Song Y, Wang J, Leung N, et al. Correlations between basophil activation, allergen-specific IgE with outcome and severity of oral food challenges. *Ann Allergy Asthma Immunol*. 2015;114:319–326.
- [20] 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;82:986–997.
- [21] Boyce JA, Assa'ad A, Burks AW, 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:S1.
- [22] Sampson HA, Gerth van Wijk R, Bindslev-Jensen C, et al. Standardizing double-blind, placebo-controlled oral food challenges: American Academy of Allergy, Asthma & Immunology-European Academy of Allergy and Clinical Immunology PRACTALL consensus report. *J Allergy Clin Immunol*. 2012;130:1260–1274.
- [23] Lieberman P, Nicklas RA, Oppenheimer J, et al. The diagnosis and management of anaphylaxis practice parameter: 2010 update. *J Allergy Clin Immunol*. 2010;126:477–480, e471–442.
- [24] Nowak-Węgrzyn 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:S365–S383.
- [25] Mukai K, Gaudenzio N, Gupta S, et al. Assessing basophil activation by using flow cytometry and mass cytometry in blood stored 24 hours before analysis. *J Allergy Clin Immunol*. 2017;139:889–899, e811.
- [26] Santos AF, Lack G. Basophil activation test: food challenge in a test tube or specialist research tool? *Clin Transl Allergy*. 2016;6:10.
- [27] Strobl C, Boulesteix AL, Kneib T, Augustin T, Zeileis A. Conditional variable importance for random forests. *BMC Bioinformatics*. 2008;9:307.
- [28] Strobl C, Boulesteix A-L, Zeileis A, Hothorn T. Bias in random forest variable importance measures: illustrations, sources and a solution. *BMC Bioinformatics*. 2007;8:25.
- [29] Strobl C, Malley J, Tutz G. An introduction to recursive partitioning: rationale, application, and characteristics of classification and regression trees, bagging, and random forests. *Psychol Methods*. 2009;14:323–348.
- [30] Breiman L, Friedman J, Olshen R, Stone CJ. Classification and Regression Trees. Belmont, CA: Wadsworth; 1984.
- [31] Tibshirani R. Regression shrinkage and selection via the Lasso. *J R Stat Soc B*. 1996;58:267–288.
- [32] Varshney P, Jones SM, Scurlock AM, et al. A randomized controlled study of peanut oral immunotherapy: clinical desensitization and modulation of the allergic response. *J Allergy Clin Immunol*. 2011;127:654–660.
- [33] Bird JA, Spergel JM, Jones SM, et al. Efficacy and safety of AR101 in oral immunotherapy for peanut allergy: results of ARC001, a randomized, double-blind, placebo-controlled phase 2 clinical trial. *J Allergy Clin Immunol Pract*. 2018;6:476–485, e3.
- [34] Lopez-Raton M, Rodríguez-Álvarez MX, Suárez CC, Sampedro FG. OptimalCutpoints: an R package for selecting optimal cutpoints in diagnostic tests. *J Stat Softw*. 2014;61:1–36.
- [35] Kohavi R. A study of cross-validation and bootstrap for accuracy estimation and model selection; 1995. <https://pdfs.semanticscholar.org/0be0/d781305750b37acb35fa187febd8db67bfcc.pdf>. Accessed August 12, 2017.
- [36] Hothorn T, Hornik K, Zeileis A. Unbiased recursive partitioning: a conditional inference framework. *J Comp Graph Stat*. 2006;15:651–674.
- [37] R Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing; 2015. <https://www.r-project.org/>. Accessed December 3, 2017.
- [38] Sporik R, Hill D, Hosking C. Specificity of allergen skin testing in predicting positive open food challenges to milk, egg and peanut in children. *Clin Exp Allergy*. 2000;30:1541–1546.
- [39] Lack G. Clinical practice: food allergy. *N Engl J Med*. 2008;359:1252–1260.
- [40] Beyer K, Grabenhenrich L, Hartl M, et al. Predictive values of component-specific IgE for the outcome of peanut and hazelnut food challenges in children. *Allergy*. 2015;70:90–98.
- [41] Codreanu F, Collignon O, Roitel O, et al. A novel immunoassay using recombinant allergens simplifies peanut allergy diagnosis. *Int Arch Allergy Immunol*. 2010;154:216–226.
- [42] Dang TD, Tang M, Choo S, et al. Increasing the accuracy of peanut allergy diagnosis by using Ara h 2. *J Allergy Clin Immunol*. 2012;129:1056–1063.
- [43] Lieberman JA, Glaumann S, Batelson S, Borres MP, Sampson HA, Nilsson C. The utility of peanut components in the diagnosis of IgE-mediated peanut allergy among distinct populations. *J Allergy Clin Immunol Pract*. 2013;1:75–82.
- [44] Nicolaou N, Poorafshar M, Murray C, et al. Allergy or tolerance in children sensitized to peanut: prevalence and differentiation using component-resolved diagnostics. *J Allergy Clin Immunol*. 2010;125:191–197, e191–113.
- [45] Klemans RJ, Broekman HC, Knol EF, et al. Ara h 2 is the best predictor for peanut allergy in adults. *J Allergy Clin Immunol Pract*. 2013;1:632–638, e631.
- [46] Klemans RJ, Otte D, Knol M, et al. The diagnostic value of specific IgE to Ara h 2 to predict peanut allergy in children is comparable to a validated and updated diagnostic prediction model. *J Allergy Clin Immunol*. 2013;131:157–163.
- [47] Ocmant A, Mulier S, Hanssens L, et al. Basophil activation tests for the diagnosis of food allergy in children. *Clin Exp Allergy*. 2009;39:1234–1245.
- [48] Blumchen K, Beder A, Beschoner J, et al. Modified oral food challenge used with sensitization biomarkers provides more real-life clinical thresholds for peanut allergy. *J Allergy Clin Immunol*. 2014;134:390–398, e394.
- [49] Virkud YV, Burks AW, Steele PH, et al. Novel baseline predictors of adverse events during oral immunotherapy in children with peanut allergy. *J Allergy Clin Immunol*. 2017;139:882–888, e885.
- [50] US Department of Health and Human Services. Common Terminology Criteria for Adverse Events (CTCAE), version 4.03; 2010. https://evs.nci.nih.gov/ftp1/CTCAE/CTCAE_4.03_2010-06-14/QuickReference_8.5x11.pdf. Accessed December 6, 2017.
- [51] Yanagida N, Okada Y, Yukiko H, et al. Food allergy and anaphylaxis – 2054: easy-to-use severity grading system for treatment of symptoms induced by oral food challenge. *World Allergy Organ J*. 2013;6:P137.
- [52] Bock SA, Muñoz-Furlong A, Sampson HA. Fatalities due to anaphylactic reactions to foods. *J Allergy Clin Immunol*. 2001;107:191–193.
- [53] Bock SA, Muñoz-Furlong A, Sampson HA. Further fatalities caused by anaphylactic reactions to food, 2001–2006. *J Allergy Clin Immunol*. 2007;119:1016–1018.

Supplementary Data

eMethods. Supplemental Statistical Analysis

To evaluate whether self-reported histories of allergies were accurate, a concordance analysis was conducted against clinical tests. A specific IgE level of 0.35 kUA/L or greater or skin prick test result of 3 mm or greater (subtracting out negative control values) was considered a positive clinical history of allergy. Logistic regression models were fit to each self-reported allergy history as a function of their respective clinical history; areas under the curve and *P* values were obtained.

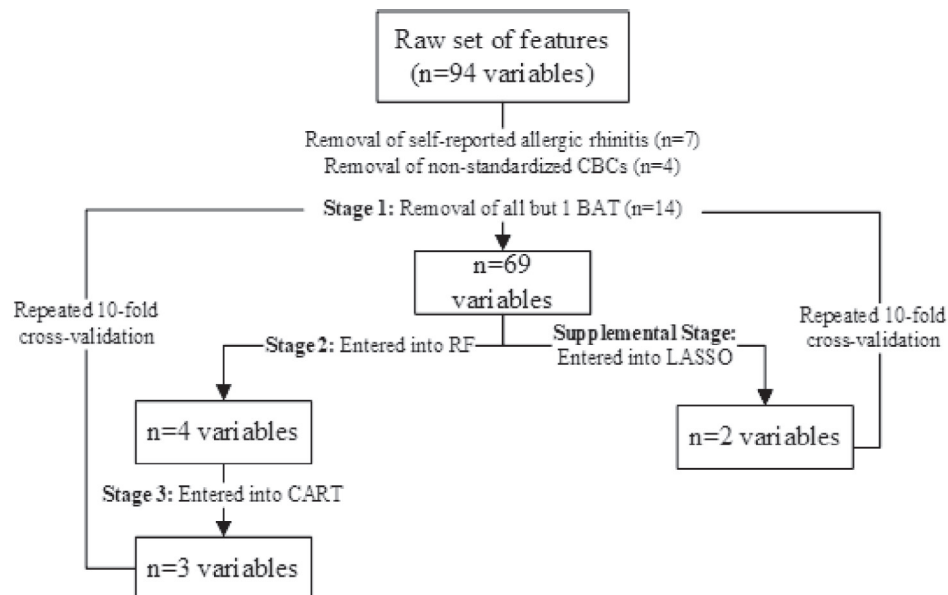
First-Stage Details

Because no assumptions on normality are needed in random forest and classification and regression trees, the chosen basophil activation tet was entered into random forest and classification and regression trees on its raw scale.

Second-Stage Details

Random forest is a general technique of building classification rules that relies on resampling techniques of the original data with and without each feature to yield importance of each variable considered by incorporating what is called the out-of-bag error (ie, the increase in error when a feature is removed from consideration). Conditional variable importance measures were calculated using the *varpimp* function from the *party* package.^{30–32} To ensure the robustness of our variable importance results, the number of trees was set to 1,000 (*cforest_unbiased* [*ntreentree* = 1,000]).

Root mean square error is a model fit statistic used to measure the differences between values predicted by a model and the observed values. It can be interpreted as the SD of the unexplained variance and is in the same units as the outcome. Although there is no clear threshold for root mean square error in determining the goodness of fit, lower values generally indicate better model fit.



eFigure 1. Variable selection and modeling consort diagram. BAT, basophil activation test; CART, classification and regression trees; CBC, complete blood cell count; LASSO, least absolute shrinkage selection operator; RF: random forest.

eTable 1
Area Under the Curve (AUC) Analysis for Basophil Activation Tests (BATs)

BAT (log scale)	AUC	P value
CD203c PF 0.1/CD203c RPMI	0.66	.09
CD203c PF 1/CD203c RPMI	0.67	.03
CD203c PF 10/CD203c RPMI	0.62	.13
CD203c PF 100/CD203c RPMI	0.60	.21
CD203c PF 1,000/CD203c RPMI	0.62	.16
CD63 ^{hi} PF 0.1	0.52	.31
CD63 ^{hi} PF 1	0.61	.33
CD63 ^{hi} PF 10	0.63	.24
CD63 ^{hi} PF 100	0.57	.83
CD63 ^{hi} PF 1,000	0.52	.91
CD63 ^{hi} PF 0.1/CD63 ^{hi} anti-IgE	0.53	.58
CD63 ^{hi} PF 1/CD63 ^{hi} anti-IgE	0.66	.03
CD63^{hi} PF 10/CD63^{hi} anti-IgE	0.67	.01
CD63 ^{hi} PF 100/CD63 ^{hi} anti-IgE	0.66	.01
CD63 ^{hi} PF 1,000/CD63 ^{hi} anti-IgE	0.59	.04

AUC from logistic regression was used to determine which BAT value was most associated with the severity challenge indicator. Bolded feature was selected for further analysis.

eTable 2
Concordance Results of Self-reported Allergies vs Clinical Testing^a

Variable	No. of patients	AUC	P value
Food allergen			
Almond	80	0.56	.19
Cashew	82	0.73	<.001
Hazelnut	79	0.58	.07
Pecan	72	0.66	.01
Walnut	84	0.63	.01
Sesame	25	0.60	.96
Egg	41	0.68	.04
Fish	25	0.95	.95
Shellfish	33	0.81	.01
Milk	32	0.68	.11
Soy	27	0.58	.49
Wheat	28	0.81	.94
Allergic rhinitis			
Dust mite	115	0.65	.003
Grass	115	0.58	.05
Tree	114	0.55	.25
Weeds	28	0.50	.97
Mold	27	0.53	.82
Pet	30	0.51	.89

Abbreviation: AUC, area under the curve.

^aLogistic regression techniques were used to determine whether self-reported allergies were concordant with clinical testing.

eTable 3
Severe Challenge Severity Score (CSS) (CSS >3) by Symptoms

Symptom	No. (%) of patients (n = 22)
Severe abdominal pain	10 (45)
Severe urticaria	1 (5)
Any wheezing	11 (50)

eTable 4
Linear Regression Model Result

Feature	Unit change in CSS	SE	P value	CV RMSE
Log(CD63 ratio) ^a	0.21	0.05	<.001	1.08

Abbreviations: CSS, challenge severity score; CV RMSE, cross-validated root mean square error.

^aCD63 ratio is the percentage of CD63^{hi} basophils stimulated with peanut protein at 10 ng/mL divided by the percentage of CD63^{hi} basophils stimulated with anti-IgE. Interpretation of coefficient is a 10% increase in CD63 ratio associated with a $0.21 \times \log(1.1) = 0.02$ unit increase in CSS.

eTable 5
LASSO Regression Model Results

Selected feature	Unit change in CSS	CV RMSE
Exercise-induced asthma	0.17	2.47
CD63 ratio ^a	0.09	

Abbreviation: CSS, challenge severity score; CV RMSE, cross-validated root mean square error; LASSO, least absolute shrinkage and selection operator.

^aCD63 ratio is the percentage of CD63^{hi} basophils stimulated with peanut protein at 10 ng/mL divided by the percentage of CD63^{hi} basophils stimulated with anti-IgE. The LASSO regression techniques were used to explore associations between the 45 preselected features and the severity score.

eTable 6
AUC Comparison of CD63 Ratio With Known Predictors of Severity

Feature	Threshold ^a	AUC (95% CI)
CD63 ratio ^b	2.29	0.68 (0.55–0.80)
Ara h 2	111.02	0.67 (0.55–0.79)
Peanut IgE	192.76	0.65 (0.51–0.79)
Skin prick test	12	0.52 (0.38–0.67)

Abbreviation: AUC, area under the curve.

^aThreshold is based on the Youden index.

^bCD63 ratio is the percentage of CD63^{hi} basophils stimulated with peanut protein at 10 ng/mL divided by the percentage of CD63^{hi} basophils stimulated with anti-IgE. Outcome is reacting at or before 175 mg of peanut protein.