

Food allergy-related quality of life after double-blind, placebo-controlled food challenges in adults, adolescents, and children

Jantina L. van der Velde, MD,^{a,b} Bertine M. J. Flokstra-de Blok, PhD,^{b,c} Hans de Groot, MD, PhD,^d Joanne N. G. Oude-Elberink, MD, PhD,^{b,e} Marjan Kerkhof, PhD,^{b,f} Eric J. Duiverman, MD, PhD,^{a,b} and Anthony E. J. Dubois, MD, PhD^{a,b} Groningen and Delft, The Netherlands

Background: Currently, the longitudinal validity (validity over time) and responsiveness (ability to measure change over time) of the Food Allergy Quality of Life Questionnaire–Adult Form (FAQLQ-AF), the Food Allergy Quality of Life Questionnaire–Teenager Form (FAQLQ-TF), and the Food Allergy Quality of Life Questionnaire–Child Form (FAQLQ-CF) are unknown.

Additionally, the self-reported impact of a double-blind, placebo-controlled food challenge (DBPCFC) on health-related quality of life (HRQL) in adults (≥18 years of age), adolescents (13–17 years of age), and children (8–12 years of age) is unknown.

Objective: The aims of this study were to assess the longitudinal validity and responsiveness of the FAQLQ-AF, FAQLQ-TF, and FAQLQ-CF and to assess the impact of a DBPCFC on HRQL.

Methods: Two hundred twenty-one participants suspected of food allergy were included from Dutch allergy centers. Participants undergoing a DBPCFC (experimental group) completed the FAQLQ and Food Allergy Independent Measure (FAIM) 1 month before (baseline) and 6 months after (follow-up) a DBPCFC. Participants not undergoing a DBPCFC (control group) completed the questionnaire package twice with a 7-month interval.

Results: HRQL scores improved after a DBPCFC, with greater improvements in HRQL scores after a negative outcome (food allergy ruled out) than a positive outcome (food allergy confirmed), demonstrating responsiveness of the FAQLQs.

Significant correlations were shown between the change (follow-up minus baseline) in FAQLQ and FAIM scores supporting longitudinal validity of these questionnaires: FAQLQ-AF (Pearson correlation coefficient = 0.71, $P < .001$), FAQLQ-TF (Pearson correlation coefficient = 0.35, $P = .018$), and FAQLQ-CF (Pearson correlation coefficient = 0.51, $P < .001$).

Conclusions: Our findings demonstrate the longitudinal validity and responsiveness of the FAQLQs. Greater

improvements in HRQL scores were shown after a negative outcome than after a positive outcome. (J Allergy Clin Immunol 2012;■■■:■■■-■■■.)

Key words: Double-blind, placebo-controlled food challenge, food allergy, health-related quality of life, longitudinal validation, responsiveness

Food is essential for life and important in social activities. Consequently, living with food allergy might seriously disrupt the daily life of patients with food allergy.¹ They are often afraid of allergic reactions on accidental exposure and are continuously faced with dietary and social restrictions.² Health-related quality of life (HRQL) instruments can be used as outcome measures to evaluate this impact of food allergy and its subsequent interventions on HRQL.³ HRQL is of particular interest as an outcome measure for food allergy because symptoms of this disease are infrequent and mortality is low.

Recently, the self-administered Food Allergy Quality of Life Questionnaire–Adult Form (FAQLQ-AF), Food Allergy Quality of Life Questionnaire–Teenager Form (FAQLQ-TF), and Food Allergy Quality of Life Questionnaire–Child Form (FAQLQ-CF) were developed for measuring HRQL in patients with food allergy.^{2,4,5} The FAQLQ-AF, FAQLQ-TF, and FAQLQ-CF were shown to be valid, reliable, and discriminative instruments.^{2,4,6} However, HRQL instruments that will be used as outcome measures in clinical trials must correlate over time with other relevant measures (longitudinally validity) and must be able to measure small but relevant HRQL changes over time (responsiveness).^{3,7} Currently, the longitudinal validity and responsiveness of the FAQLQ-AF, FAQLQ-TF, and FAQLQ-CF are unknown.

A double-blind, placebo-controlled food challenge (DBPCFC), which is the gold standard for diagnosing food allergy, is likely to cause HRQL changes over time, especially if the test result is negative (food allergy ruled out). Until now, only 1 study has been published on the impact of food challenges on HRQL. This study showed that parent-proxy-reported HRQL of children with food allergy (0–12 years) improves after a food challenge.³ Thus far, no studies have been published on the impact of a DBPCFC on HRQL in adolescents (13–17 years of age) and adults (≥18 years of age). Because these age groups are developmentally different, the experience of food allergy and its subsequent interventions might differ for each group. Additionally, no studies have been published on the impact of a DBPCFC on HRQL scores in children (8–12 years of age) from the child's perspective, whereas it is well known that there generally is child-parent disagreement on the child's HRQL.^{8,9} Finally, no studies have been published

From ^athe Department of Pediatrics, Division of Pediatric Pulmonology and Pediatric Allergy, ^bthe GRIAC Research Institute, ^cthe Department of General Practice, ^dthe Department of Allergy, and ^ethe Department of Epidemiology, University of Groningen, University Medical Center Groningen, Groningen, and ^fthe Department of Pediatric Allergology, Reinier de Graaf Groep, Delft.

Supported by the European Union through the EuroPrevall Project (FOOD-CT-2005-514000), by the Nutricia Research Foundation, and by the Stichting Astma Bestrijding. Disclosure of potential conflict of interest: The authors declare that they have no relevant conflicts of interest.

Received for publication October 10, 2011; revised May 15, 2012; accepted for publication May 30, 2012.

Corresponding author: Jantina L. van der Velde, MD, Department of Pediatrics, Division of Pediatric Pulmonology and Pediatric Allergy, University Medical Center Groningen, PO Box 30.001, 9700 RB Groningen, The Netherlands. E-mail: j.l.van.der.velde@umcg.nl

0091-6749/\$36.00

© 2012 American Academy of Allergy, Asthma & Immunology

<http://dx.doi.org/10.1016/j.jaci.2012.05.037>

Abbreviations used

DBPCFC:	Double-blind, placebo-controlled food challenge
FAIM:	Food Allergy Independent Measure
FAQLQ-AF:	Food Allergy Quality of Life Questionnaire–Adult Form
FAQLQ-CF:	Food Allergy Quality of Life Questionnaire–Child Form
FAQLQ-TF:	Food Allergy Quality of Life Questionnaire–Teenager Form
HRQL:	Health-related quality of life
MID:	Minimal important difference
NNT:	Number needed to treat
SEM:	Standard error of measurement

comparing HRQL in participants who undergo a DBPCFC with control participants.

Therefore, the aims of this study were to investigate the longitudinal validity and responsiveness of the FAQLQs and to investigate the self-reported impact of a DBPCFC on HRQL of adults, adolescents, and children with food allergy.

METHODS

Participants and procedures

Adults (≥18 years of age), adolescents (13–17 years of age), and children (8–12 years of age) who were awaiting a clinically indicated DBPCFC were included from Dutch allergy centers in Groningen and Voorburg between May 2007 and January 2011. The only criterion used for referral for a DBPCFC was a clinical suspicion of food allergy. IgE measurements, skin prick tests, or prior history of anaphylaxis were not routinely used in our decision to refer for a DBPCFC. The 2 centers collaborate on uniform methodology of DBPCFC procedures.^{10,11} A DBPCFC consists of a placebo and an active challenge, which were randomly administered on separate days with at least a 2-week interval. Participants and any subjects with patient contact were blinded to the sequence of the 2 challenges.

All patients visiting the clinic with suspected food allergy were placed on the waiting list for a DBPCFC on a first-come, first-served basis. Participants were included in the experimental group or control group based on the expected duration of time spent on the waiting list for a DBPCFC. Participants who were expected to have a DBPCFC within 6 months were included in the experimental group and completed the questionnaire package at home 1 month before (baseline) and 6 months after (follow-up) a DBPCFC (placebo and active challenge) was performed. Participants who were not expected to have a DBPCFC within 6 months were included in the control group and completed the questionnaire package twice with a 7-month interval without undergoing a DBPCFC (no placebo or active challenge). Additionally, participants recruited through Dutch food allergy support organizations were included in the control group when they reported a physician-diagnosed food allergy (skin prick test, IgE).

Participants were excluded from the study if diagnostic/therapeutic interventions or accidental ingestions resulting in clinical allergic reactions (Mueller classification grade III/IV) had taken place between baseline and follow-up measurement or when less than 85% of the questions were completed. Possible differences between descriptive characteristics of participants (Table I) recruited in the participating centers and between participants from the control and experimental groups were examined by using χ^2 (nominal data) and Mann-Whitney (ordinal, continuous data) tests.

This study was approved by the local medical ethics review commission, who deemed that permission from the commission was not required.

Questionnaires

FAQLQs. The Dutch FAQLQ-AF, FAQLQ-TF, and FAQLQ-CF are self-reported, disease-specific instruments for measuring the effect of food allergy on the adult's (≥18 years of age),² adolescent's (13–17 years of age),⁵ and

child's (8–12 years of age) HRQL,⁴ respectively. The FAQLQ-AF contains 29 items and 4 domains (Risk of Accidental Exposure, Emotional Impact, Allergen Avoidance and Dietary Restrictions, and Food Allergy–related Health). The FAQLQ-TF contains 23 items and 3 domains (Allergen Avoidance and Dietary Restrictions, Risk of Accidental Exposure, and Emotional impact). The FAQLQ-CF contains 24 items and 4 domains (Risk of Accidental Exposure, Emotional Impact, Allergen Avoidance, and Dietary Restrictions). Each item is scored on a 7-point scale. The total FAQLQ score of each questionnaire is the mean of all items and ranges from 1 (minimal impairment in HRQL) to 7 (maximal impairment in HRQL).

Food Allergy Independent Measure. The Food Allergy Independent Measure (FAIM),¹² a self-report instrument, reflects the participant's perceived food allergy severity and food allergy–related risk. The FAIM consists of 4 expectation-of-outcome questions, which capture the participants' perceived expectation of the chance of accidental exposure and of what will happen after accidental exposure. These expectations are likely to be the source/predictors of HRQL changes in patients with anaphylactic disorders.¹³ Additionally, the FAIM contains 2 questions that reflect disease severity. Each question was scored on a 7-point scale. The total FAIM score is the mean of all items and ranges from 1 (limited severity perception) to 7 (greatest severity perception). Instruments based on the method of expectation-of-outcome questions have proved to be useful independent measures for evaluating the construct validity of HRQL instruments^{2,4,5,14–16} and are used in this study to evaluate the longitudinal validity of the FAQLQs.

Statistical analysis

Data were analyzed with SPSS software for Windows (version 18.0; SPSS, Chicago, Ill).

Change in HRQL and perceived disease severity after DBPCFC. Changes in HRQL (follow-up minus baseline FAQLQ scores) were tested for significance (Wilcoxon signed-rank test). *P* values of .05 or less were considered statistically significant. The overall difference in HRQL changes was calculated as follows:

$$\text{HRQL change for the experimental group} - \text{HRQL change for the control group.}$$

After the outcome of the DBPCFC, participants in the experimental group (Exp_total) were split into 3 groups: (1) positive outcome (ie, food allergy in question confirmed [Exp_pos]); (2) negative outcome (ie, food allergy in question ruled out [Exp_neg]); and (3) questionable outcome (ie, food allergy not confirmed/not ruled out [Exp_quest]). Additionally, a subgroup analysis was performed on participants with a negative outcome and no remaining other food allergy (Exp_neg_NRFA). Changes in perceived disease severity (follow-up minus baseline FAIM scores) were analyzed in the same way.

Relevance of changes in HRQL after DBPCFCs. Both a statistical and clinical method were used to evaluate whether a change in HRQL was considered relevant.

First, the smallest change in HRQL that is considered statistically important is called the standard error of measurement (SEM) and is calculated as follows:

$$\text{SEM} = \sigma_x \sqrt{1 - r_{xx}}.^{17,18}$$

σ_x represents the SD of the FAQLQ baseline measurement, and r_{xx} represents the reliability or the intraclass correlation coefficient of the FAQLQs.^{6,14}

Second, the smallest change in HRQL that is considered clinically important by participants is called the minimal important difference (MID).¹⁹ The MID has been estimated to be approximately 0.5 in several HRQL questionnaires using a 7-point scale^{3,17,19,20} and can be used to calculate the therapeutic value of an intervention. The therapeutic value of an intervention is defined as the number of participants who need to undergo the intervention for 1 participant to have a clinically important improvement over and above that which he or she would have experienced with the control intervention. The therapeutic value of a DBPCFC was calculated by using the methodology of number needed to treat (NNT).²¹

TABLE I. Participants' characteristics

	Age group		
	≥18 y, n* (%† [nexp‡])	13-17 y, n (% [nexp])	8-12 y, n (% [nexp])
No.	73 (100% [53])	71 (100% [46])	77 (100% [57])
Sex			
Female	58 (79% [44])	35 (49% [23])	29 (38% [22])
Male	15 (21% [9])	36 (51% [23])	48 (62% [35])
Age (y), mean (SD)	35.8 (14.9)	15.2 (1.4)	10.4 (1.6)
DBPCFC			
Peanut	20 (27% [20§])	23 (32% [23])	25 (32% [25])
Nut	12 (16% [12])	9 (13% [9])	18 (23% [18])
Milk	2 (3% [2])	8 (11% [8])	7 (9% [7])
Egg	3 (4% [3])	3 (4% [3])	5 (6% [5])
Wheat	8 (11% [8])	—	1 (1% [1])
Soy	4 (5% [4])	3 (4% [3])	1 (1% [1])
Sesame	4 (5% [4])	—	—
No. of food allergies			
1 Food	17 (23% [13])	12 (17% [9])	20 (26% [13])
2 Foods	14 (19% [12])	22 (31% [13])	26 (34% [17])
3 Foods	12 (16% [9])	15 (21% [9])	12 (16% [10])
>3 Foods	30 (41% [19])	22 (31% [15])	19 (25% [17])
Symptoms			
Cardiovascular	39 (53% [29])	34 (48% [21])	19 (25% [13])
Respiratory	67 (92% [49])	61 (86% [42])	58 (75% [41])
Gastrointestinal	44 (60% [31])	43 (61% [28])	41 (53% [26])
Skin	53 (73% [38])	50 (70% [30])	46 (60% [32])
Ever experienced anaphylaxis?			
Yes	62 (85% [46])	55 (77% [35])	56 (73% [41])
No	11 (15% [7])	16 (23% [11])	21 (27% [16])

*Number of participants (sum of experimental and control groups).

†Percentage of participants (sum of experimental and control groups).

‡Number of participants in the experimental group.

§DBPCFCs not performed in participants of the control group.

Percentages of participants whose symptoms improve, remain stable, and deteriorate were calculated as follows: change in HRQL scores of less than −0.5 (important improvement in HRQL), between −0.5 and 0.5 (unchanged HRQL), and greater than 0.5 (important deterioration in HRQL). The proportion of participants who fared better after a DBPCFC or who fared better in the control group were calculated from the proportion of participants with an important improvement in HRQL, unchanged HRQL, or an important deterioration in HRQL (see Table E1 in this article's Online Repository at www.jacionline.org). The overall proportion of participants who benefit from a DBPCFC is the difference of the 2 abovementioned proportions. The NNT of a DBPCFC is the reciprocal of the proportion of participants benefiting from DBPCFCs.

Factors influencing HRQL changes. Multiple linear regression analysis was performed to correct for factors possibly influencing changes in HRQL. The change in FAQLQ scores was used as the dependent variable. The outcome of the DBPCFC (positive vs control/negative vs control) and possible confounders (type of recruitment, center of recruitment, age, baseline HRQL scores, and number of food allergies) were used as independent variables.

Responsiveness of the FAQLQs. Changes in HRQL scores were calculated and tested for significance (see "Change in HRQL scores after DBPCFC" below) to investigate the ability of the FAQLQs to measure relevant changes over time (responsiveness).

Longitudinal validity of the FAQLQs. Pearson correlation coefficients were calculated between the change in FAQLQ scores and the change in FAIM scores to assess the longitudinal validity of the FAQLQs. Moderate correlations were expected.²²

RESULTS

Participants

Questionnaire packages were sent to 392 participants and returned by 299 participants (response rate, 76%; see Table E2 in this article's Online Repository at www.jacionline.org). Follow-up questionnaire packages were returned by 235 participants (response rate, 79%). Fourteen participants were excluded (reasons listed in Table E2). Therefore, 221 participants were eligible for analysis. Participants' characteristics are shown in Table I. There were no significant differences between participants in the control group and the experimental group, except for the mean age of adults (control group, 49.3 years; experimental group, 32.4 years; $P = .002$). Specifically, there were no differences between participants recruited from the 2 participating centers (experimental group) or between participants recruited by advertisement or from the clinic (control group, data not shown).

Change in HRQL after DBPCFC

There were no significant differences in baseline HRQL (FAQLQ) scores between the experimental and control groups (adults, $P = .956$; adolescents, $P = .622$; and children, $P = .908$) or between patients recruited from the 2 participating centers (adolescents, $P = .782$; children, $P = .384$) or between participants recruited by advertisement or from the clinic (adolescents, $P = .777$), except for adults (5.68 vs 3.78, $P = .009$).

For adults, HRQL scores improved significantly after a DBPCFC when all outcomes of the test were combined (baseline, 4.35; follow-up, 3.81; $P = .001$; $n = 53$), whereas the HRQL scores of participants in the control group did not change significantly (baseline, 4.35; follow-up, 4.24; $P = .421$; $n = 20$; Table II). A great improvement in HRQL scores was shown after a negative outcome (food allergy ruled out, $n = 24$) and when there was no remaining other food allergies (baseline, 4.20; follow-up, 1.97; $P = .046$; $n = 6$). A smaller significant improvement in HRQL scores was shown after a positive outcome (baseline, 4.56; follow-up, 4.21; $P = .036$; $n = 24$). After a questionable outcome, there was no significant change in HRQL (baseline, 3.93; follow-up, 4.16; $P = .225$; $n = 5$). For children, similar results were shown. For adolescents, HRQL scores improved significantly after a negative DBPCFC outcome (baseline, 3.68; follow-up, 2.71; $P = .013$; $n = 13$). However, HRQL scores did not improve significantly after a positive outcome (baseline, 3.93; follow-up, 3.92; $P = .905$; $n = 29$), whereas the HRQL scores of participants in the control group improved significantly (baseline, 4.03; follow-up, 3.69; $P = .018$; $n = 25$).

Relevance of changes in HRQL after DBPCFC

To assess the relevance of changes in HRQL scores from a statistical and clinical point of view, the SEM was calculated (FAQLQ-AF, 0.25; FAQLQ-TF, 0.16; and FAQLQ-CF, 0.44), and the MID of 0.5 was used. The MID is the smallest difference in HRQL that is considered to be clinically important by patients. The overall improvement in HRQL scores after a negative DBPCFC outcome exceeded the SEM and MID for adults, adolescents, and children and was therefore considered statistically and clinically relevant (Table III). The overall differences in HRQL after a positive DBPCFC outcome only reached the SEM and were therefore considered to be statistically relevant. The overall improvement in HRQL scores nearly reached the MID for children.

TABLE II. Mean FAQLQ and FAIM scores after DBPCFC

Age (y), no. (exp/con)	FAQLQ scores			FAIM scores		
	≥18 (53/20)	13-17 (46/25)	8-12 (57/20)	≥18 (53/20)	13-17 (46/25)	8-12 (57/20)
DBPCFC outcome						
Exp_Total,* no.	53	46	57	53	46	57
Baseline (SD)	4.35 (1.15)	3.89 (1.03)	3.80 (1.45)	3.84 (0.92)	3.48 (0.86)	3.35 (0.97)
Follow-up (SD)	3.81 (1.35)	3.65 (1.32)	3.25 (1.39)	3.31 (1.01)	3.33 (1.05)	3.19 (1.02)
P value	.001	.169	.002	.001	.264	.292
Exp_Pos,† no.	24	29	31	24	29	31
Baseline (SD)	4.56 (0.93)	3.93 (1.12)	3.75 (1.44)	3.91 (0.78)	3.66 (0.89)	3.35 (0.97)
Follow-up (SD)	4.21 (1.21)	3.92 (1.15)	3.36 (1.28)	3.71 (0.92)	3.76 (0.88)	3.33 (0.89)
P value	.036	.905	.041	.295	.535	.857
Exp_Neg,‡ no.	24	13	23	24	13	23
Baseline (SD)	4.24 (1.36)	3.68 (0.90)	3.93 (1.47)	3.89 (1.08)	3.10 (0.83)	3.38 (0.96)
Follow-up (SD)	3.36 (1.47)	2.71 (1.29)	3.14 (1.54)	2.91 (1.03)	2.35 (0.90)	3.02 (1.21)
P value	.009	.013	.018	<.001	.013	.153
Exp_Neg_NRFA,§ no.	6	4	4	6	4	4
Baseline (SD)	4.20 (1.85)	3.71 (0.75)	3.39 (1.51)	4.17 (1.42)	2.75 (0.67)	3.54 (1.09)
Follow-up (SD)	1.97 (1.41)	1.58 (0.50)	2.55 (1.51)	2.00 (0.62)	1.33 (0.24)	2.88 (1.71)
P value	.046	.068	.465	.042	.068	.380
Exp_Quest, no.	5	4	3	5	4	3
Baseline (SD)	3.93 (1.01)	4.32 (0.69)	3.21 (1.70)	3.30 (0.70)	3.42 (0.29)	3.11 (0.59)
Follow-up (SD)	4.16 (0.77)	4.72 (1.02)	2.97 (1.62)	3.31 (0.67)	3.35 (0.36)	3.11 (0.59)
P value	.225	.273	.593	1.00	.593	1
Control group, no.	20	25	20	20	25	20
Baseline (SD)	4.35 (1.52)	4.03 (1.28)	3.77 (1.27)	3.79 (1.14)	3.75 (1.07)	3.31 (0.98)
Follow-up (SD)	4.24 (1.53)	3.69 (1.19)	3.84 (1.20)	3.67 (1.18)	3.79 (0.85)	3.37 (1.01)
P value	.421	.018	.763	.195	.884	.751

Scores in boldface represent significant values ($P < .05$).

*Experimental group (exp).

†Experimental group, positive DBPCFC outcome.

‡Experimental group, negative DBPCFC outcome.

§Experimental group, negative DBPCFC outcome and no remaining food allergy.

||Experimental group, questionable DBPCFC outcome.

TABLE III. NNT of a DBPCFC

DBPCFC outcome	Overall difference in HRQL	Proportion better on DBPCFC	Proportion better in control group	Proportion benefiting after DBPCFC	NNT
Negative					
Adults	-0.76 ($P = .023$)	0.47	0.19	0.28	3.56
Adolescents	-0.62 ($P = .040$)	0.41	0.17	0.23	4.28
Children	-0.87 ($P = .023$)	0.43	0.19	0.24	4.18
Positive					
Adults	-0.24 ($P = .240$)	0.34	0.23	0.11	8.75
Adolescents	0.34 ($P = .078$)	0.19	0.43	-0.24	-4.15
Children	-0.47 ($P = .065$)	0.45	0.19	0.26	3.83

Overall difference = Change in experimental group (follow-up - baseline HRQL scores) - change in control group. A negative overall difference represents an improvement in HRQL scores.

The proportion of participants with a clinically important improvement in the HRQL score (< -0.5) after a negative DBPCFC outcome was higher than in the control group (Fig 1 and see Table E1 in this article's Online Repository at www.jacionline.org). The proportion of participants whose HRQL scores were the same (-0.5 to 0.5) after a negative outcome was lower than in the control group. The proportions of participants with an important deterioration in HRQL score (> 0.5) were 13% (adults), 8% (adolescents), and 13% (children) after a negative DBPCFC outcome compared with 25% (adults), 8% (adolescents), and 0% (children) in the control group. The NNT of a negative DBPCFC outcome was 4 (Table III). Thus for every

4 participants who have a negative DBPCFC outcome, 1 participant has a clinically important improvement in HRQL.

After a positive DBPCFC outcome, the NNT was higher (indicating lower effectiveness) compared with that seen in adults with a negative DBPCFC outcome (Table III). Remarkably, adolescents in the control group fared better than adolescents after a positive DBPCFC outcome.

Factors influencing HRQL changes

A negative DBPCFC outcome resulted in improvement in HRQL scores compared with that seen in the control group after

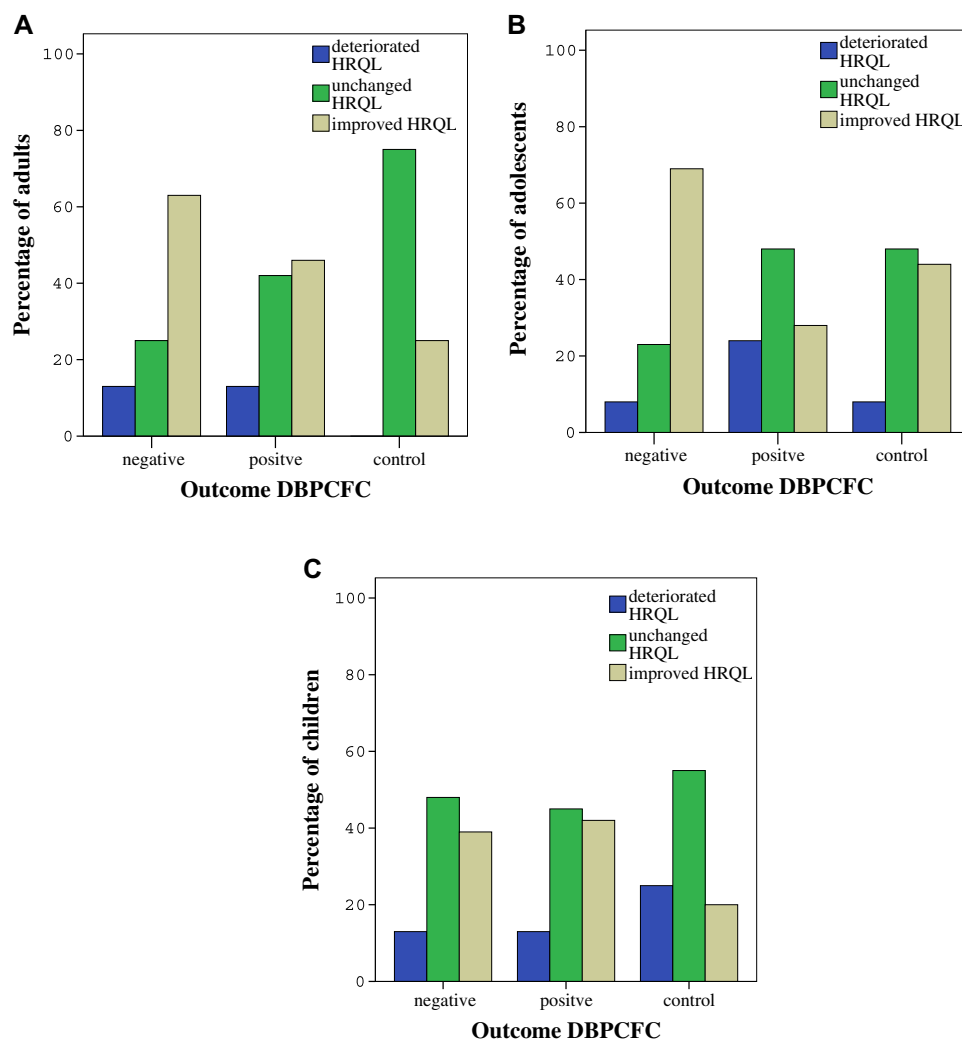


FIG 1. A-C, Percentages of participants whose HRQL scores deteriorate (change >0.5), shows no change (change between -0.5 and 0.5), or improves (change <-0.5) after a DBPCFC in adults (Fig 1, A), adolescents (Fig 1, B), and children (Fig 1, C). **D-F,** Percentages of participants whose FAIM scores deteriorate, show no change, or improve in adults (Fig 1, D), adolescents (Fig 1, E), and children (Fig 1, F).

correction for confounding (Table IV). A positive DBPCFC outcome resulted in improvement in HRQL scores after correction for confounding in children. Having more than 2 food allergies was associated with deterioration in HRQL scores for children. Higher FAQLQ baseline scores (poorer HRQL) contributed significantly to improvement in HRQL for all ages.

Change in perceived disease severity (FAIM score) after DBPCFC

There were no significant differences in baseline FAIM scores between participants in the control and experimental groups (adults, $P = .701$; adolescents, $P = .199$; and children, $P = .986$) or between participants recruited from the 2 participating centers (adolescents, $P = .410$; children, $P = .875$) or between participants recruited by means of advertisement or from the clinic (adults, $P = .148$; adolescents, $P = .393$).

For adults, FAIM scores improved significantly after a DBPCFC when all outcomes of the test were combined (baseline, 3.84; follow-up, 3.31; $P = .001$; Table II). A great improvement in perceived disease severity was shown after a negative outcome

(baseline, 3.89; follow-up, 2.91; $P < .001$). No improvement in perceived disease severity was shown after a positive outcome (baseline, 3.91; follow-up, 3.71; $P = .295$). No change in perceived disease severity was shown in the control group (baseline, 3.79; follow-up, 3.67; $P = .195$). Similar results were shown for adolescents and children. The percentages of participants with a clinically important change in FAIM score are shown in Fig 1, D-F.

Responsiveness of the FAQLQs

All FAQLQs were able to measure changes exceeding the SEM and MID of the FAQLQs. This supports the responsiveness of all FAQLQs.

Longitudinal validation of the FAQLQs

Significant correlations were shown between the change in FAQLQ scores and the change in FAIM scores for the FAQLQ-AF (Pearson correlation coefficient = 0.71, $P < .001$), FAQLQ-TF (Pearson correlation coefficient = 0.35, $P = .018$), and FAQLQ-CF (Pearson correlation coefficient = 0.51, $P < .001$). This supports the longitudinal validity of the FAQLQs.

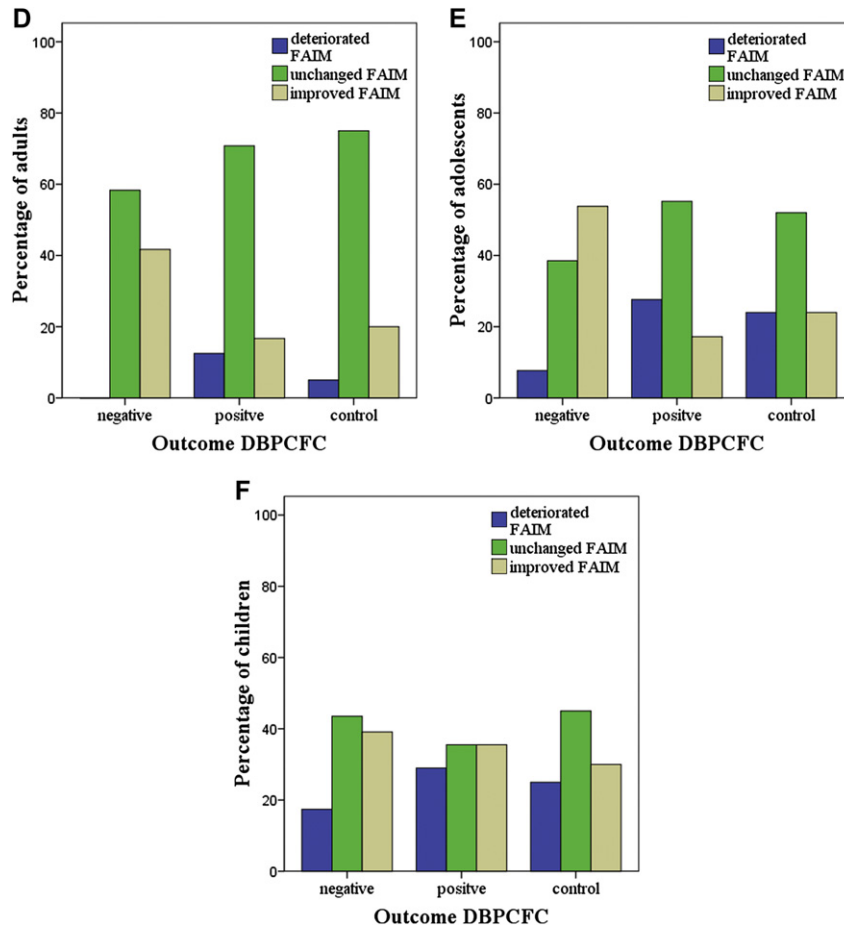


FIG 1. (Continued)

TABLE IV. Factors influencing change in HRQL (follow-up minus baseline scores) after positive and negative DBPCFC outcomes using multiple linear regression analysis

	Adults			Adolescents			Children		
	B*	P value	95% CI	B	P value	95% CI	B	P value	95% CI
Test outcome									
Positive (control vs positive)	-0.16	.68	-0.91 to 0.59	0.34	.43	-0.52 to 1.20	-0.94	.02	-1.74 to -0.18
Negative (control vs negative)	-0.74	.03	-1.42 to -0.06	-0.70	.011	-1.56 to 0.17	-1.19	.00	-1.94 to -0.45
Age (y)	-0.01	.42	-0.03 to 0.01	-0.08	.23	-0.22 to 0.05	0.06	.42	-0.09 to 0.21
No. of food allergies (≤ 2 vs > 2)	0.39	.14	-0.13 to 0.92	0.31	.16	-0.13 to 0.74	0.50	.04	0.02 to 0.99
FAQLQ baseline (1 = best HRQL to 7 = worst HRQL)	-0.26	.02	-0.47 to -0.05	-0.27	.01	-0.45 to -0.08	-0.37	.00	-0.55 to -0.20
Center recruitment (Voorburg vs Groningen)	NA	NA	NA	-0.44	.24	-1.17 to 0.30	0.38	.24	-0.26 to 1.01
Type of recruitment (clinic vs advertisement)	-0.21	.71	-0.89 to 1.31	-0.52	.11	-1.16 to 0.12	NA	NA	NA

NA, Not applicable.

*A negative coefficient represents an improvement in the HRQL score. *P* values of .05 or less were considered statistically significant.

DISCUSSION

This study examined, for the first time, the longitudinal validity and responsiveness of the FAQLQ-AF, FAQLQ-TF, and FAQLQ-CF and the self-reported effect of a DBPCFC on HRQL in patients with food allergy. All FAQLQs were able to measure relevant changes in HRQL scores and correlated significantly to measures of perceived disease severity. These findings thus support the responsiveness and longitudinal validity of the FAQLQ-AF, FAQLQ-TF, and FAQLQ-CF and the suitability of these instruments as outcome measures.

Additionally, this study shows that adults, adolescents, and children benefit from a negative DBPCFC outcome, and children even benefit if the outcome is positive. The NNT of a negative DBPCFC outcome was 4. Thus for every 4 participants who have a negative DBPCFC outcome, 1 participant has a clinically important improvement in HRQL. These findings thus support the large impact of food allergy on HRQL and illustrate which patients benefit from DBPCFCs and its subsequent changes in management, from a HRQL point of view.

Overall, our findings are in agreement with those of previous literature, which showed that parent-proxy reports on the child's HRQL improved significantly after a food challenge in children aged 0 to 12 years and that larger improvements in HRQL were shown after a negative outcome than after a positive outcome.³ Additionally, other studies on parents' perceptions of their child's food challenge²³⁻²⁶ showed reduced parental concerns^{25,26} and parental anxiety²⁴ after food challenges, irrespective of the outcome of the challenge. It was hypothesized that these improvements in HRQL/well-being were caused by the fact that a definitive diagnosis provides a sense of certainty. This hypothesis was based on findings that the main concerns of patients with food allergy and their families were uncertainty and lack of information rather than fear of the challenge procedure³ and that a greater certainty of the diagnosis after a positive or negative outcome was perceived as positive in children with egg allergy.²⁶ This hypothesis is further supported by our data showing no improvement in HRQL scores after challenges with an uncertain outcome. Thus by providing a sense of certainty, a food challenge might have a positive effect on the HRQL score, irrespective of the outcome.^{3,26}

There is much discussion on the interpretation of HRQL changes.^{19,27-32} Different methods can be used to determine a minimal important change (ie, the change in HRQL that is considered clinically relevant). Such methods can be distribution or anchor based. Distribution-based methods, such as the SEM, rely on scoring methods of an instrument and the distribution of the results.²⁹ Anchor-based methods use an external criterion (anchor) to interpret whether a particular magnitude of change is meaningful for patients/clinicians. Because it is necessary to know whether the observed change is important from the patient's/clinician's perspective, anchor-based methods are still preferred in assessing the MID.¹⁹ In our study both a distribution-based method (SEM) and an anchor-based method (MID) were used. We found a statistically and clinically relevant change in HRQL scores after a negative DBPCFC outcome and for children after a positive DBPCFC outcome.

Although HRQL scores of adolescents improve significantly after a negative DBPCFC outcome, some remarkable findings were shown for adolescents. First, adolescents' HRQL scores did not improve after a positive outcome. Several factors might explain this finding. Because different issues relating to food allergy might have different effects at different stages of life,^{2,4,5} the experience of food allergy-related interventions, such as DBPCFCs, might differ for each age group as well. It might be that the effect of the feeling of certainty of diagnosis after a DBPCFC on HRQL, as is suggested by studies of parents of children with food allergy,²³⁻²⁶ is less relevant for adolescents themselves because it is overshadowed by the effect of peer comparison and peer pressure^{3,33,34} on the adolescent's perception of living with food allergy. A second remarkable finding was that the HRQL scores of adolescents in the control group improved significantly, although less than the MID of 0.5. Despite this, the overall improvement in HRQL scores after a negative DBPCFC outcome (improvement in the experimental group minus improvement in the control group) was statistically and clinically significant (0.62, $P = .040$). However, the nature of factors associated with HRQL changes in adolescents requires further investigation because no predictive factors were identified in this study.

Another unexpected finding was that the HRQL scores of some participants deteriorated after a negative DBPCFC outcome, even

though this only occurred in 7 participants. Further characterization of these participants showed that some participants' characteristics tended to be more common in participants whose HRQL scores deteriorated after a negative DBPCFC outcome compared with participants whose HRQL scores did not deteriorate. These factors are number of food allergies (4/2, $P = .062$), percentage of peanut challenges (57%/21%, $P = .079$) and percentage of participants who did not eat the tested food despite the negative DBPCFC outcome (57%/33%, $P = .108$). We hypothesize that these factors might influence deterioration in HRQL scores after a negative challenge result. However, this needs further investigation with larger numbers of patients.

This study also has some limitations. Although there was sufficient power to determine the longitudinal validity and responsiveness of the FAQLQ-AF, FAQLQ-TF, and FAQLQ-CF and to determine the overall effect of a DBPCFC on HRQL, the power to study the effect of a DBPCFC in the different subgroups (age specific and DBPCFC outcome specific) is moderate. Therefore, the subgroup analyses should be interpreted with caution.

Second, the majority of participants had peanut or nut allergy, which might limit the generalizability of these findings to patients with other food allergies.

Third, because of logistic and ethical reasons, we assigned participants to the control and experimental groups based on the expected duration of time spent on the waiting list. Therefore, we corrected for possible confounders, such as method of recruitment, age, and baseline HRQL scores. Although a true randomized trial still is methodologically preferred, it might be difficult to generalize results obtained from participants willing to be randomized to postponement of a DBPCFC to most other patients with food allergy.

In summary, our results support the longitudinal validity and responsiveness of the FAQLQ-AF, FAQLQ-TF, and FAQLQ-CF and the suitability of these instruments as outcome measures. Additionally, a DBPCFC is considered a time-consuming process and has previously been discussed in the literature as being potentially burdensome to patients. However, our study shows that children, adolescents, and adults benefit from a negative DBPCFC outcome and its subsequent changes in management, and children do so even if the outcome is positive. Thus, a DBPCFC is not only the gold standard for diagnosing food allergy, but also improves patient HRQL for specified groups. This provides clinicians with an additional reason to promote the use of DBPCFCs. Information from self-reported FAQLQs will help to optimize such interventions from an HRQL point of view. Further research is needed to study the predictors of HRQL changes after DBPCFCs.

Clinical implications: The FAQLQs are longitudinally valid and responsive instruments for measuring food allergy-related quality of life. Clinicians should promote the use of DBPCFCs because food allergy-related quality of life improves after DBPCFCs for specified groups.

REFERENCES

1. Flokstra-de Blok BM, Dubois AE, Vlieg-Boerstra BJ, Oude Elberink JN, Raat H, DunnGalvin A, et al. Health-related quality of life of food allergic patients: comparison with the general population and other diseases. *Allergy* 2010;65:238-44.
2. Flokstra-de Blok BM, van der Meulen GN, DunnGalvin A, Vlieg-Boerstra BJ, Oude Elberink JN, Duiverman EJ, et al. Development and validation of the food allergy quality of life questionnaire-adult form. *Allergy* 2009;64:1209-17.

3. DunnGalvin A, Cullinane C, Daly DA, Flokstra-de Blok BM, Dubois AE, Hourihane JO. Longitudinal validity and responsiveness of the food allergy quality of life questionnaire–parent form in children 0–12 years following positive and negative food challenges. *Clin Exp Allergy* 2010;40:476–85.
4. Flokstra-de Blok BM, DunnGalvin A, Vlieg-Boerstra BJ, Oude Elberink JN, Duiverman EJ, Hourihane JO, et al. Development and validation of a self-administered food allergy quality of life questionnaire for children. *Clin Exp Allergy* 2009;39:127–37.
5. Flokstra-de Blok BM, DunnGalvin A, Vlieg-Boerstra BJ, Oude Elberink JN, 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, e1–2.
6. van der Velde JL, Flokstra-de Blok BM, Vlieg-Boerstra BJ, Oude Elberink JN, Schouten JP, DunnGalvin A, et al. Test-retest reliability of the food allergy quality of life questionnaires (FAQLQ) for children, adolescents and adults. *Qual Life Res* 2009;18:245–51.
7. Oude Elberink JN, De Monchy JG, Van Der Heide S, Guyatt GH, Dubois AE. Venom immunotherapy improves health-related quality of life in patients allergic to yellow jacket venom. *J Allergy Clin Immunol* 2002;110:174–82.
8. van der Velde JL, Flokstra-de Blok BM, DunnGalvin A, Hourihane JO, Duiverman EJ, Dubois AE. Parents report better health-related quality of life for their food-allergic children than children themselves. *Clin Exp Allergy* 2011;41:1431–9.
9. van der Velde JL, Flokstra-de Blok BM, Hamp A, Knibb RC, Duiverman EJ, Dubois AE. Adolescent-parent disagreement on health-related quality of life of food-allergic adolescents: who makes the difference? *Allergy* 2011;66:1580–9.
10. Vlieg-Boerstra BJ, Herpertz I, Pasker L, van der Heide S, Kukler J, Jansink C, et al. Validation of novel recipes for double-blind, placebo-controlled food challenges in children and adults. *Allergy* 2011;66:948–54.
11. Vlieg-Boerstra B, Dubois A. Published recipes for double-blind placebo-controlled food challenges. *Allergy* 2005;60:1212.
12. van der Velde JL, Flokstra-de Blok BM, Vlieg-Boerstra BJ, Oude Elberink JN, DunnGalvin A, Hourihane JO, et al. Development, validity and reliability of the food allergy independent measure (FAIM). *Allergy* 2010;65:630–5.
13. Oude Elberink JN. Significance and rationale of studies of health-related quality of life in anaphylactic disorders. *Curr Opin Allergy Clin Immunol* 2006;6:298–302.
14. DunnGalvin A, de BlokFlokstra BM, Burks AW, Dubois AE, Hourihane JO. Food allergy QoL questionnaire for children aged 0–12 years: Content, construct, and cross-cultural validity. *Clin Exp Allergy* 2008;38:977–86.
15. Cohen BL, Noone S, Munoz-Furlong A, Sicherer SH. Development of a questionnaire to measure quality of life in families with a child with food allergy. *J Allergy Clin Immunol* 2004;114:1159–63.
16. Oude Elberink JN, de Monchy JG, Golden DB, Brouwer JL, Guyatt GH, Dubois AE. Development and validation of a health-related quality-of-life questionnaire in patients with yellow jacket allergy. *J Allergy Clin Immunol* 2002;109:162–70.
17. Kocks JW, Tuinenga MG, Uil SM, van den Berg JW, Stahl E, van der Molen T. Health status measurement in COPD: the minimal clinically important difference of the clinical COPD questionnaire. *Respir Res* 2006;7:62.
18. Norman GR, Sloan JA, Wyrwich KW. Interpretation of changes in health-related quality of life: the remarkable universality of half a standard deviation. *Med Care* 2003;41:582–92.
19. Juniper EF, Guyatt GH, Willan A, Griffith LE. Determining a minimal important change in a disease-specific quality of life questionnaire. *J Clin Epidemiol* 1994;47:81–7.
20. Deyo RA, Diehr P, Patrick DL. Reproducibility and responsiveness of health status measures. Statistics and strategies for evaluation. *Control Clin Trials* 1991;12(suppl):142S–58S.
21. Guyatt GH, Juniper EF, Walter SD, Griffith LE, Goldstein RS. Interpreting treatment effects in randomised trials. *BMJ* 1998;316:690–3.
22. Bullinger M, Anderson R, Cella D, Aaronson N. Developing and evaluating cross-cultural instruments from minimum requirements to optimal models. *Qual Life Res* 1993;2:451–9.
23. Strinnholm A, Brulin C, Lindh V. Experiences of double-blind, placebo-controlled food challenges (DBPCFC): a qualitative analysis of mothers' experiences. *J Child Health Care* 2010;14:179–88.
24. Zijlstra WT, Flinterman AE, Soeters L, Knulst AC, Sinnema G, L'Hoir MP, et al. Parental anxiety before and after food challenges in children with suspected peanut and hazelnut allergy. *Pediatr Allergy Immunol* 2010;21:e439–45.
25. Nguyen M, Wainstein BK, Hu W, Ziegler JB. Parental satisfaction with oral peanut food challenges; perception of outcomes and impact on management of peanut allergy. *Pediatr Allergy Immunol* 2010;21:1119–26.
26. Kemp AS, Allen CW, Campbell DE. Parental perceptions in egg allergy: does egg challenge make a difference? *Pediatr Allergy Immunol* 2009;20:648–53.
27. Guyatt GH, Osoba D, Wu AW, Wyrwich KW, Norman GR. Clinical Significance Consensus Meeting Group. Methods to explain the clinical significance of health status measures. *Mayo Clin Proc* 2002;77:371–83.
28. Schunemann HJ, Akl EA, Guyatt GH. Interpreting the results of patient reported outcome measures in clinical trials: the clinician's perspective. *Health Qual Life Outcomes* 2006;4:62.
29. Turner D, Schunemann HJ, Griffith LE, Beaton DE, Griffiths AM, Critch JN, et al. The minimal detectable change cannot reliably replace the minimal important difference. *J Clin Epidemiol* 2010;63:28–36.
30. Revicki D, Hays RD, Cella D, Sloan J. Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. *J Clin Epidemiol* 2008;61:102–9.
31. Wyrwich KW, Tierney WM, Wolinsky FD. Using the standard error of measurement to identify important changes on the asthma quality of life questionnaire. *Qual Life Res* 2002;11:1–7.
32. Rejas J, Pardo A, Ruiz MA. Standard error of measurement as a valid alternative to minimally important difference for evaluating the magnitude of changes in patient-reported outcomes measures. *J Clin Epidemiol* 2008;61:350–6.
33. DunnGalvin A, Gaffney A, Hourihane JO. Developmental pathways in food allergy: a new theoretical framework. *Allergy* 2009;64:560–8.
34. DunnGalvin A, Hourihane JB. Chapter 3 developmental trajectories in food allergy: a review. *Adv Food Nutr Res* 2009;56:65–100.

TABLE E1. Calculation of the proportion of patients showing improvement (HRQL change ≤ -0.5), no change (HRQL change between -0.5 and 0.5), or deterioration (HRQL change >0.5) in HRQL scores after DBPCFC

Control group	Negative DBPCFC outcome (food allergy ruled out)			Positive DBPCFC outcome (food allergy confirmed)		
	Improved (u)	Unchanged (v)	Deteriorated (w)	Improved (x)	Unchanged (y)	Deteriorated (z)
Adults	(0.63)	(0.25)	(0.13)	(0.46)	(0.42)	(0.13)
Adolescents	(0.69)	(0.23)	(0.08)	(0.28)	(0.48)	(0.24)
Children	(0.39)	(0.48)	(0.13)	(0.42)	(0.45)	(0.13)
Improved (a)	(au)	(av)	(aw)	(ax)	(ay)	(az)
Adults (0.25)	0.16	0.06	0.03	0.11	0.10	0.03
Adolescents (0.44)	0.30	0.10	0.03	0.12	0.21	0.11
Children (0.20)	0.08	0.10	0.03	0.08	0.09	0.03
Unchanged (b)	(bu)	(bv)	(bw)	(bx)	(by)	(bz)
Adults (0.75)	0.47	0.19	0.09	0.34	0.31	0.09
Adolescents (0.48)	0.33	0.11	0.04	0.13	0.23	0.12
Children (0.55)	0.22	0.26	0.07	0.23	0.25	0.07
Deteriorated (c)	(cu)	(cv)	(cw)	(cx)	(cy)	(cz)
Adults (0.00)	0.00	0.00	0.00	0.00	0.00	0.00
Adolescents (0.08)	0.06	0.02	0.01	0.02	0.04	0.02
Children (0.25)	0.10	0.12	0.03	0.10	0.11	0.03

The proportion of patients who fared better after a negative DBPCFC outcome is the sum of *bu*, *cu*, and *cv*. The proportion of patients who fared better in the control group is the sum of *av*, *aw*, and *bw*. The proportion of patients whose outcome is the same irrespective of a positive diagnosis is the sum of *au*, *bv*, and *cw*. Proportions after a positive DBPCFC outcome were calculated in the same way.

TABLE E2. Flow chart participant recruitment

	All ages	≥18 y	13-17 y	8-12 y
Recruited participants, no.	392	137	117	138
Returned first packages, no.	299	92	96	111
Returned second packages, no.	235	79	77	79
Excluded participants, no.	14	6	6	2
<85% of questions completed	4	1	1	2
Management intervention	6	2	4	—
Did not complete DBPCFC	3	3	—	—
Severe anaphylactic reaction before follow-up measurement	1	—	1	—
Included, no.	221	73	71	77
Experimental (Voorburg/Groningen), no.	156 (20/136)	53 (0/53)	46 (6/40)	57 (14/43)
Control, no. (clinic/food allergy support organizations)	65 (42/23)	20 (14/6)	25 (9/16)	20 (19/1)