Adjusting for nonresponse bias corrects overestimates of food allergy prevalence

Lianne Soller, BSc, MSc, PhD candidate[a],
Moshe Ben-Shoshan, MD, MSc[b],
Daniel W. Harrington, MA, PhD[c], Megan Knoll, MSc[a],
Joseph Fragapane, BEng, MD[a], Lawrence Joseph, PhD[a,d],
Yvan St. Pierre, MSc[a], Sebastien La Vieille, MD[e],
Kathi Wilson, PhD[f], Susan J. Elliott, PhD[g], and
Ann E. Clarke, MD, MSc[a,h]

Clinical Implications

- We are the first to demonstrate that adjustment for nonresponse can lead to important changes in food allergy prevalence. Clinicians must be cautious when interpreting the literature because most authors do not account for nonresponse.

TO THE EDITOR:

Nationwide estimates of food allergy prevalence are frequently based on telephone surveys, as this allows population-based sampling from geographically diverse regions. The most recent telephone surveys from the United States and Canada estimate that the prevalence of self-reported food allergy ranges between 8.1% and 9.1%.[1,2] However, such studies are often limited as they provide prevalence estimates for a limited number of allergies[3,4] and do not consider nonresponse bias,[1-6] which may result in an overrepresentation of certain demographic groups who may tend to report more allergies.

Given these limitations, we used data collected in the Canadian population-based SPAACE (Surveying Prevalence of food Allergy in All Canadian Environments) study, which inquired about allergies to several foods and obtained information from households who refused or could not be reached to complete the study. This allowed us to: (1) provide population-weighted prevalence estimates of allergy to any food and (2) explore the influence of nonresponse bias on prevalence by presenting a range of estimates using different assumptions about food allergy prevalence among nonresponders.

METHODS

Survey methodology

The SPAACE study was a random cross-Canada telephone survey conducted between September 2010 and 2011, which targeted vulnerable Canadians (ie, those of low income, New Canadians, and of self-reported Aboriginal identity) using 2006 Canadian Census data (refer to Supplement E1 in this article's Online Repository at www.jaci-inpractice.org).[7,8] Households were telephoned and the initial adult respondent was queried using the Food Allergy Prevalence Questionnaire (FAPQ) on whether any household member had an allergy to peanut, tree nut, fish, shellfish, sesame, milk, egg, wheat, and/or soy, or other foods.[7] Food allergy was defined as follows:

TABLE I. Weighted perceived and probable prevalence estimates of food allergy by age group

	Children under 18, % (95% CrI) (n = 4026)	Adults 18 and over, % (95% CrI) (n = 10,996)	All ages, % (95% CrI) (n = 15,022)
Perceived			
Peanut	2.4 (1.6, 3.2)	0.7 (0.5, 0.9)	1.1 (0.9, 1.3)
Tree nut	1.6 (1.0, 2.3)	1.2 (0.9, 1.5)	1.3 (1.0, 1.6)
Fish	1.0 (0.3, 1.8)	0.6 (0.4, 0.8)	0.7 (0.5, 0.9)
Shellfish	1.4 (0.6, 2.1)	1.9 (1.5, 2.2)	1.7 (1.4, 2.0)
Sesame	0.1 (0.0, 0.3)	0.2 (0.1, 0.3)	0.2 (0.1, 0.3)
Milk	0.7 (0.3, 1.1)	0.7 (0.5, 0.9)	0.7 (0.5, 0.9)
Egg	1.0 (0.6, 1.5)	0.5 (0.3, 0.7)	0.6 (0.4, 0.8)
Wheat	0.3 (0.0, 0.6)	0.4 (0.2, 0.6)	0.4 (0.2, 0.5)
Soy	0.1 (0.0, 0.3)	0.1 (0.0, 0.2)	0.1 (0.1, 0.2)
Other	2.2 (1.5, 3.0)	3.5 (3.0, 4.0)	3.2 (2.8, 3.6)
Any	6.9 (5.5, 8.2)	7.7 (6.9, 8.4)	7.5 (6.9, 8.1)
Probable*			
Peanut	2.2 (1.4, 2.9)	0.6 (0.4, 0.8)	1.0 (0.7, 1.2)
Tree nut	1.5 (0.9, 2.1)	1.0 (0.8, 1.3)	1.2 (0.9, 1.4)
Fish	0.9 (0.3, 1.6)	0.5 (0.3, 0.7)	0.6 (0.4, 0.8)
Shellfish	0.8 (0.4, 1.2)	1.6 (1.3, 2.0)	1.4 (1.2, 1.7)
Sesame	0.1 (0.0, 0.3)	0.2 (0.1, 0.3)	0.2 (0.1, 0.3)
Milk	0.2 (0.0, 0.3)	0.2 (0.1, 0.3)	0.2 (0.1, 0.3)
Egg	1.0 (0.5, 1.5)	0.5 (0.3, 0.6)	0.6 (0.4, 0.8)
Wheat	0.2 (0.0, 0.5)	0.2 (0.1, 0.4)	0.2 (0.1, 0.4)
Soy	0.1 (0.0, 0.3)	0.1 (0.0, 0.2)	0.1 (0.0, 0.2)

*We collected only detailed information about food allergy to the 9 common foods; therefore, probable estimates for other foods and any food could not be calculated.

(1) *Perceived*: individuals self-reporting any food allergy, and
(2) *Probable*: individuals self-reporting a convincing history[9,10] and/or a physician diagnosis of allergy to peanut, tree nut, fish, shellfish, sesame, milk, egg, wheat, and/or soy.

If the respondent refused to complete the FAPQ, the interviewer administered a much briefer Refusal Questionnaire (RQ) that queried if any household member had an allergy and if present, data on the household size, the respondent's education, the food(s) to which the individual was allergic, and whether the allergy was diagnosed by a doctor were collected.

Developing weighted estimates of prevalence

Point estimates and 95% credible intervals (CrIs) for the prevalence of perceived and probable allergy were weighted to account for the oversampling of vulnerable populations (refer to Supplement E2 in this article's Online Repository at www.jaci-inpractice.org).[7] Credible intervals are the Bayesian analogue to standard confidence intervals.

Developing nonresponse bias estimates

To develop nonresponse bias-adjusted estimates of prevalence of perceived allergy to any food, 4 groups were identified:

(1) *Full Participants*: households who completed the FAPQ,
(2) *Refusal Questionnaire Participants*: households who completed the RQ only,

TABLE II. Nonadjusted and bias-adjusted prevalence estimates of perceived allergy to any food

	Nonadjusted		Bias-adjusted			
Estimate number	Full participants (FP), % (95% CrI) (n = 15,022)	Refusal questionnaire participants (RQP), % (95% CrI) (n = 1393*)	Nonparticipants (NP), % (95% CrI) (n = 17,059*)		Never reached participants (NRP), % (95% CrI) (n = 8419*)	All participants, % (95% CrI) (n = 41,893)
					NRP same as NP	
1	6.4 (6.0, 6.8)	2.1 (1.4, 2.9)	NP half RQP	1.0 (0.7, 1.4)	1.1 (0.7, 1.5)	3.0 (2.8, 3.3)
2	6.4 (6.0, 6.8)	2.1 (1.4, 2.9)	NP same as RQP	2.1 (1.4, 2.8)	2.1 (1.5, 2.9)	3.7 (3.2, 4.2)
3	6.4 (6.0, 6.8)	2.1 (1.4, 2.9)	NP twice RQP	4.2 (2.8, 5.7)	4.3 (2.9, 5.9)	4.9 (4.1, 5.9)
					NRP mixture of FP, RQP, and NP	
4	6.4 (6.0, 6.8)	2.1 (1.4, 2.9)	NP half RQP	1.0 (0.7, 1.4)	3.5 (3.2, 3.8)	3.5 (3.2, 3.8)
5	6.4 (6.0, 6.8)	2.1 (1.4, 2.9)	NP same as RQP	2.1 (1.4, 2.8)	4.0 (3.6, 4.5)	4.0 (3.6, 4.5)
6	6.4 (6.0, 6.8)	2.1 (1.4, 2.9)	NP twice RQP	4.2 (2.9, 5.7)	5.1 (4.4, 6.0)	5.1 (4.4, 5.9)
					NRP same as FP	
7	6.4 (6.0, 6.8)	2.1 (1.4, 2.9)	NP half RQP	1.0 (0.7, 1.4)	6.4 (6.0, 6.9)	4.1 (3.8, 4.4)
8	6.4 (6.0, 6.8)	2.1 (1.4, 2.9)	NP same as RQP	2.1 (1.4, 2.8)	6.4 (6.0, 6.9)	4.5 (4.2, 4.9)
9	6.4 (6.0, 6.8)	2.1 (1.4, 2.9)	NP twice RQP	4.2 (2.8, 5.7)	6.4 (6.0, 6.9)	5.4 (4.8, 6.1)

*The number of people in all nonallergic households in the RQP group, and in all households in the NP and NRP groups, was imputed using the distribution of the number of people in each household in the FP group.

(3) *Nonparticipants*: households who were reached by telephone but refused to complete either questionnaire, and
(4) *Never Reached Participants*: households who could not be reached by telephone.

Food allergy data were available only from *Full* and *RQ Participants*. Multiple imputation (MI), the gold standard for adjusting for missing data,[11] was used to adjust the estimates for nonresponse bias that resulted from missing food allergy data within the *Nonparticipants* and the *Never Reached Participants* by using a model that included observed data (census tract [CT] and province of residence) to predict the missing data on the probability of food allergy.[12]

A range of assumptions regarding the prevalence of food allergy in the *Nonparticipants* and *Never Reached Participants* were investigated (refer to Supplement E3 in this article's Online Repository at www.jaci-inpractice.org). Compared with the prevalence in the *RQ Participants* living in the same CT, the prevalence in the *Nonparticipants* was assumed to be: (1) half, (2) equal to, and (3) twice as large as the *RQ Participants*.

Compared with the prevalence of those in the same CT, the prevalence among the *Never Reached Participants* was assumed to be: (1) equal to the *Nonparticipants*; (2) a weighted average of the *Full*, *RQ*, and *Nonparticipants*; and (3) equal to the *Full Participants*.

MI was implemented via a hierarchical logistic regression model with 4 levels: individual, household, CT, and province of residence. Weighting to account for the overrepresentation of vulnerable populations could not be done in this analysis because demographic information was only available for *Full Participants*. The analyses were performed using WinBUGS (version 1.4.3, MRC Biostatistics Unit, Cambridge, United Kingdom) (refer to Supplement E3 in this article's Online Repository at www.jaci-inpractice.org).

RESULTS

Participation rate

We telephoned 17,337 households, 14,113 of whom were actually reached. Of these 14,113 households, 1351 were ineligible due to a language barrier or unavailability of an adult. Of the 12,762 eligible households, 5734 households, representing 15,022 individuals, completed the FAPQ (45% response rate, or 5734 of 12,762) and were thus *Full Participants*, 524 households completed the RQ (an additional 4%, or 524 of 12,762) and were thus *RQ Participants*, and the remaining 6504 households answered the telephone but refused to provide any information (51%) and were thus *Nonparticipants*. An additional 3224 households were never reached, and were thus *Never Reached Participants*.

Prevalence estimates

Among *Full Participants*, the unweighted self-reported (perceived) prevalence of allergy to any food was 6.4% (6.0%, 6.8%). After weighting, this estimate increased to 7.5% (6.9%, 8.1%) (Table I).

Compared with the *Full Participants*, the unweighted perceived prevalence of allergy to any food was lower among the *RQ Participants* (6.4% [6.0%, 6.8%] vs 2.1% [1.4%, 2.9%]) (Table II). Applying the different assumptions regarding the prevalence of food allergy among the *Nonparticipants* and *Never Reached Participants*, 9 selection bias-adjusted estimates were obtained for the perceived prevalence of allergy to any food ranging from 3.0% (2.8%, 3.3%) to 5.4% (4.8%, 6.1%) (refer to Table II and Supplement E3 in this article's Online Repository at www.jaci-inpractice.org).

DISCUSSION

Comparison with previous studies

The unweighted perceived prevalence of food allergy in this study (6.4% [6.0%, 6.8%]) was less than that in our general population study conducted 2 years earlier (8.1% [7.5%, 8.7%]),[2] but these estimates are not directly comparable as our current study targeted vulnerable populations. The weighted perceived prevalence in the current study (7.5% [6.9%, 8.1%]) is also lower than that estimated in the NHANES study, a US population-based door-to-door survey conducted between 2007 and 2010 (9.0% [8.3%, 9.6%]).[13] The NHANES survey is weighted for nonresponse in general, but this weighting may not be sufficient to account for all possible nonresponse bias.[13] However, our weighted perceived prevalence in children (6.9% [5.5%, 8.2%]) is similar to that estimated by Gupta in a US population-based