Heliyon 10 (2024) e33545

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

Egg protein exposure estimation in risk assessment for Japanese food allergy labeling

Hiroshi Akiyama [a,b,*], Yoshinari Suzuki [b,1], Reiko Adachi [c], Momoka Kadokura [a], Asaka Takei [a], Masayoshi Tomiki [a,d], Kosuke Nakamura [b], Rie Ito [a], Yusuke Iwasaki [a], Clare Mills [e], Yukihiro Ohya [f], Tatsuki Fukuie [f]

[a] *Hoshi University, School of Pharmacy and Pharmaceutical Sciences, Department of Analytical Chemistry, 2-4-41 Ebara, Shinagawa-ku, Tokyo, 142-8501, Japan*
[b] *National Institute of Health Sciences, Division of Foods, 3-25-26 Tonomachi, Kawasaki-ku, Kawasaki, Kanagawa, 210-9501, Japan*
[c] *National Institute of Health Sciences, Division of Biochemistry, 3-25-26 Tonomachi, Kawasaki-ku, Kawasaki, Kanagawa, 210-9501, Japan*
[d] *Morinaga Institute of Biological Science, Inc., 2-1-1 Shimosueyoshi, Tsurumi-ku, Yokohama, Kanagawa, 230-8504, Japan*
[e] *The University of Surry, School of Biosciences, Guildford, Surrey, GU2 7XH, UK*
[f] *National Center for Child Health and Development, Allergy Center, 2-10-1 Okura, Setagaya-ku, Tokyo, 157-8535, Japan*

A R T I C L E I N F O

Keywords:
Labeling
Food allergy
Risk assessment
Bayesian estimation
Egg protein

A B S T R A C T

To assess the risk of food allergies in foods processed under the Japanese food labeling system, estimating exposure to hidden allergens is necessary. We assessed exposure to egg protein in foods processed according to the Japanese food labeling system. First, we estimated the concentration distribution of egg protein by Bayesian methods using data from the literature and the measurement of food products with precautional declarations in the labeling margin. We then estimated the food-intake portion-size distribution under two scenarios: soft drink consumption as an example of single, high-intake consumption, and confections, which are frequently consumed by children, as a realistic example of low-intake consumption. Finally, we estimated the distribution of unexpected intake of egg proteins in the form of single consumption. The mean exposure to egg protein under the high-intake scenario was estimated to be 0.0164 mg for 1–15-year-olds, 0.0171 mg for 4–15-year-olds, 0.0181 mg for 7–15-year-olds, and ≥0.0188 mg for 16-year-olds. The mean exposure to egg protein under the low-intake scenario was estimated to be 0.0018 mg for 1–15-year-olds, 0.0019 mg for 4–15-year-olds, 0.0020 mg for 7–15-year-olds, and ≥0.0022 mg for 16-year-olds. Compared to the reference dose of 2.0 mg proposed by the Joint the Food and Agriculture Organization (FAO)/World Health Organization (WHO) Expert Committee, the risk of onset of food allergies due to egg protein contamination from foods without egg labeling is considered to be extremely low under the current Japanese food labeling system.

1. Introduction

Food allergies are immunological reactions induced by ingesting foods that contain allergens and that manifest as symptoms, with even very small amounts (several milligrams) capable of causing serious health problems, such as anaphylactic shock [1]. In recent

* Corresponding author. Department of Analytical Chemistry, School of Pharmacy and Pharmaceutical Science, Hoshi University, Japan.
E-mail address: h-akiyama@hoshi.ac.jp (H. Akiyama).
[1] These two authors contributed equally to this work.

https://doi.org/10.1016/j.heliyon.2024.e33545
Received 29 December 2023; Received in revised form 22 June 2024; Accepted 23 June 2024
Available online 24 June 2024

years, the number of patients with food allergies has been increasing worldwide [1]. In Japan, the prevalence of food allergies is estimated to be 5–10 % in infants, 5 % in young children, and 4.5 % in school children [2]. Accidental ingestion of food containing allergens in patients can cause anaphylactic symptoms, such as decreased blood pressure, dyspnea, and even death. As globalization progresses in areas such as trade and cross-cultural exchange, food distribution and eating habits are gradually changing and diversifying. Consequently, verifying the effectiveness of the current Japanese risk management system based on scientific knowledge is important for ensuring food safety.

Precautionary allergen labeling (PAL), indicated by the statement "May contain X″ followed by the types of potential allergens, is permitted in Western countries. The PAL statements such as "Manufactured on Shared Equipment with X″ or " Made in the Same Facility as X″ or "Packaged on Shared Equipment with X″ are also included. However, the likelihood of contamination by potential allergens in such cases is typically so low that such warnings would not be practical for most applications and/or providing this information would result in an excess of PAL. Further, consumers with allergies find PAL confusing and unreliable [3]. The use of PAL statement such as May contain X is prohibited by the Japanese food allergy labeling system. However, the declarations such as such as "This manufacturing facility produces products containing allergenic ingredients" or "Manufactured using the same equipment used to produce the following allergenic ingredient(s)", are recommended when there is a risk of tentative contamination due to the same factory production lines being used to process foodstuffs, which stipulates that cautionary declaration be provided outside the label margin. The cautionary declaration outside the label margin is similar to the PAL statement allowed in other countries. However, this cautionary declaration is not considered to be labeling in Japan, because the declaration is written outside the margin. Therefore, if the food products with the cautionary declaration were detected over 10 mg/kg and it is proven that eggs are regularly contaminated, the product could be taken administrative measures or punished in Japanese labeling system. Accordingly, the PAL system used in other countries employed in Western countries differs from the Japanese labeling system.

To mitigate the risk of anaphylactic symptoms, the Japanese government requires that specific food ingredients be labeled on food items based on the Food Labeling Act (Law No. 70 of 2013) [4]. Although the methods used for scientifically evaluating foods containing allergens was not defined in that act, the Basic Act on Countermeasures against Allergic Diseases (Act No. 98 of 2014) was enacted in 2015. The latter act states that "The Japanese government will scientifically verify the labeling of foods containing allergens, and will appropriately review the labeling based on the results" (Article 15). In addition, there is a need to consider the scientific basis upon which the labeling system is based in Japan in a way that considers the trends in international evaluation standards.

Since the susceptibility of individuals with food allergies varies markedly among consumers, estimating the probability of foodstuffs inducing allergic reactions is important around the world [4–10]. The risk of individuals with food allergies ingesting foods containing allergens through cross-contamination depends not only on the amounts of allergens that are potentially present in the product, but also on the quantities consumed. In addition, the severity of symptoms also depends on the amount of allergen required to elicit a patient-specific reaction; this threshold is referred to as the evoking dose (ED). While the amounts of allergens consumed can be estimated based on food consumption surveys, estimating the threshold requires clinical information obtained from patients. Therefore, it is important to perform risk assessments based on the probability of triggering allergic symptoms and the probability of exposure. To conduct a quantitative risk assessment based on the estimation of the probability of triggering allergic reactions in individuals with food allergies in Japan, it is necessary to estimate the exposure to potential allergens in the food ingested at a single sitting.

In developing a risk assessment strategy, an assessment method that is based on quantifying the risk level from a clinical standpoint has been internationally developed for the purpose of estimating the minimum ED [10,11]. The calculation is estimated using the benchmark dose method, and is based on the results of oral food challenge tests. Data on major allergens, such as peanuts, eggs, and milk, for which there are labeling obligations are being compiled by the Codex Alimentarius Commission (a dedicated body within the Joint FAO/WHO Food Standards Programme). Since these data are mainly from Europe and the United States, data on Asians are somewhat limited. Furthermore, considerable progress has been made in recent years on developing evaluation methods to reduce the use of PAL [12]. For example, several parametric models have been used as interval-censored, survival analysis tools to scientifically estimate the relationship between EDs and their associated probabilities in food allergy patient populations. In addition, some risk assessment studies including the exposure to peanut or sesame proteins have been reported [12–16]. However, the risk assessment studies on exposure to egg protein for food allergy labeling have not yet been conducted.

This study focusses on eggs, a globally recognized allergenic ingredient with the highest frequency of inducing food allergies. We estimated the concentration distribution of egg protein by Bayesian estimation based on data in the literature and the concentration of egg protein in processed foods. These foods lacked labels explicitly mentioning that the foods contained egg as a raw material, but included a precautionary declaration outside the labeling margin. We then estimated the distribution of unexpected intake of egg proteins using food consumption patterns. High-intake scenarios were defined as beverages with large consumption volumes, while the low-intake scenario focused on confections, which are frequently consumed by children, a high-risk group for food allergies.

2. Materials and methods

2.1. Food consumption data

Food consumption data were collected by a nationwide dietary survey conducted during Japanese fiscal year (FY) 2005–FY 2007 [17]. The survey was conducted over three discontinuous days within a period of up to 12 days in four different seasons: May to June, August to September, November to December, and February to March. Participants living in 25 areas throughout Japan were selected, and the data obtained in the paper-based survey were collected and checked individually by registered dietitians. The average amount of daily food consumption by each age group (1–6-year-olds, 7–14-year-olds, 15–19-year-olds, and ≥20-year-olds) was summarized to calculate the average food intake. A total of 189 foodstuffs were classified into seven food categories (I, seasonings and beverages; II, cereals; III, potatoes, legumes and nuts; IV, fish/shellfish, meat and hens' eggs; V, oils/fats and milk/milk products; VI, sugar and confections/savories; and VII, fruits, vegetables and seaweed).

2.2. Monitoring by local governments

No reports on monitoring the validity of food allergen labeling by local governments were published before 2003. However, from 2004 onwards, annual reports were published by local government research institutes (Hokkaido, Miyagi, Tokyo, Chiba, Kanagawa, Okayama, Yamaguchi, and Kochi prefectures, and Sapporo, Hiroshima, and Fukuoka cities), as well as reports in academic journals (group of three local governments: Kanagawa Prefecture, Kawasaki, and Yokohama cities), by Chiba Prefecture, and by the National Institute of Health Sciences. When comparing the aggregated results for 2003–2008, 2009–2013, and 2014–2019, the concentration and frequency of allergen contamination were the lowest in 2014–2019. This is due to the fact that the allergen labeling system has become more widely adopted since it was launched in FY 2001.

Data for the detection of eggs in processed foods in which egg was not declared were carefully examined and recounted using the aggregated results of 2014–2019; the data from this period are considered to most closely reflect the current situation. Data from Hokkaido, Tokyo, Kanagawa, and Okayama prefectures that were published in the annual reports of research institutions [18–23] showed that the total number of food samples tested was 121. Since each sample was assayed using two different analytical kits, the total number of samples assayed was 242. Egg allergen protein concentrations were also determined for 16 samples of processed foods, of which 4 samples were assayed using both of the analytical kits and 12 samples were tested using only one, which gave a total number of samples of 20.

These egg protein data were then tabulated and aggregated into a dataset containing 242 samples (121 samples × 2 analytical kits). Of these 242 samples, the sample size for sweets was 56 (28 samples × 2 analytical kits).

We used a dataset to determine the prior distributions for the Bayesian estimation (BE) method; however, these datasets also included many non-detected (ND) values. To analyze the data, including the ND values, we employed the maximum likelihood estimation (MLE) method. We derived a likelihood function [$L(Y|\theta)$] by using the probability density function [$f(Y|\theta)$] for the quantified values, and the cumulative distribution function [$F(\mathrm{LB},\mathrm{UB}|\theta)$] between the lower bound (LB) and the upper bound (UB) for the ND values. The likelihood function was as follows:

$$L(Y|\theta) = \prod_{i=1}^{N_{\mathrm{obs}}} f(Y_i|\theta) \times \prod_{j=1}^{N_{\mathrm{cen}}} F\left(\mathrm{LB}_j, \mathrm{UB}_j \middle| \theta\right),$$

where Y is the observed value, θ is the distribution parameter vector, N_{obs} is the quantified sample size, and N_{cen} is sample size of ND. We used 0 and the reporting limit for LB and UB, respectively.

2.3. Determination of hidden egg protein using ELISA

We randomly purchased 92 processed foods (chocolate; crackers; rice crackers; cakes; crisps; *daifuku*, which is a Japanese sweet made using red bean filling wrapped in rice cake; cookies; marshmallows; biscuits; jelly; *okaki*, which is a sweet made from sticky rice; and drinks containing *Lactobacillus*) with labels that did not specifically state that the products contained egg as a raw material, but that did declare that eggs may be included outside the label margin. The protein concentration of each protein in the samples was determined using a FASPEK II kit (Morinaga Biochemical Research Institute, Yokohama, Japan). Although the original limit of quantification (LOQ) for ELISA measurements reported by the manufacturer was 0.31 ppm, the signal intensity for 0.1 ppm egg protein was sufficiently higher than that of the blank. Therefore, we employed 0.1 ppm as the limit of detection (LOD). The quantification was performed numerically by extrapolating the standard curve between the LOD and LOQ.

2.4. Bayesian estimation (BE)

The posterior distribution $f_{\mathrm{post}}(\theta|Y)$ of the distribution parameters obtained by BE can be expressed using Bayes' theorem, as follows:

$$f_{\mathrm{post}}(\theta|Y) = \frac{L(Y|\theta) f_{\mathrm{pri}}(\theta)}{f_{\mathrm{obs}}(Y)},$$

where $f_{\mathrm{pri}}(\theta)$ is the prior distribution of the distribution parameter θ and $f_{\mathrm{obs}}(Y)$ is the probability density function of the observed value Y. Since the observed value Y does not change after observation, it can be regarded as a constant and the formula can be simplified as follows:

$$f_{\mathrm{post}}(\theta|Y) \propto L(Y|\theta) f_{\mathrm{pri}}(\theta).$$

in other words, the BE method can be interpreted as the MLE method weighted based on prior knowledge.

The MLE method estimates the distribution parameter with the highest likelihood. Since the method maximizes the probability of generating data that are identical to the sample, the issue of overfitting needs to be considered [24]. In addition, the MLE method is not recommended for small sample sizes (e.g., n < 50) [25], primarily because small sample sizes have a large effect due to sample bias. The BE method can be interpreted as a means of weighting the results of the MLE method based on the prior distribution.

For the BE of the egg protein concentration in the processed foods, the log-normal distribution, gamma distribution, and Weibull distribution were all examined as probability density functions that take positive real numbers. The Cauchy distribution expressed by the following probability density function was used as the prior distribution:

$$f(Y|y_0, \gamma) = \frac{1}{\pi} \frac{\gamma}{(Y - y_0)^2 + \gamma^2}.$$

The Cauchy distribution is a distribution with heavy tails, and Gelman proposed the use of the Cauchy distribution as a weak informative prior [26]. We adopted $(y_0, \gamma) = (\widehat{\theta}_s^{MLE}, \widehat{\theta}_s^{MLE})$ for shape parameters (e^{σ}, α, and m of log-normal, gamma, and Weibull distributions, respectively), where $\widehat{\theta}_s^{MLE}$ means the estimator of the shape parameter using MLE for monitoring data in local government research. For the location parameter, we adopted $(y_0, \gamma) = (\widehat{\mu}^{MB}, \widehat{\mu}^{MB})$ for the theoretical population mean, where $\widehat{\mu}^{MB}$ is a mean value calculated by substituting half of the LOD into non-detects.

In the BE for the distribution of daily consumption of confections, we examined the log-normal, gamma, Weibull, and generalized gamma distributions as probability density functions which take non-negative real numbers. Since the histograms of beverage consumption show a bimodal distribution, we examined mixed log-normal, mixed gamma, and mixed Weibull distributions as probability density functions. We adopted the uniform distribution U(-∞, ∞) as a non-informative prior distribution in the BE method, because the sample size was sufficient.

The BE calculations were mostly performed using the R package, rstan (ver. 2.16.2) [27] under the following conditions: warmup = 2000, sampling = 500, thinning = 1, chain = 4. A total of 2000 Monte Carlo samples were obtained as the posterior distribution. By calculating the log-likelihood for each record, we calculated the widely applicable information criterion (WAIC) for each model [28], and adopted the model with the lowest WAIC as the optimal model. The WAIC was calculated using the R package loo (ver. 2.4.1) [29].

2.5. Exposure assessment

We performed simulations following two intakes, each with different combinations of egg protein concentration, and a single consumption distribution: a low-intake scenario for confectionery items and a high-intake scenario for beverages. In the low-intake scenario, we estimated egg protein exposure from confectionery items, which is a food group with a high intake frequency among children. In the high-intake scenario, we estimated the egg protein exposure for the beverages with largest single consumption. In the low-intake scenario, both egg protein concentration and consumption were estimated based on confectionery data. On the other hand, in the high-intake scenario, the consumption amount was estimated using the beverage data, but the concentration was estimated based on the data for all foods including the beverage data. The low-intake scenario is considered to be realistic in that children consume sweets. On the other hand, the probability that beverages contain egg-derived allergens is low compared to other foods, so the high-intake scenario is intended to be an excessive estimate compared to the actual exposure level, although the beverage of canned corn soup might be contaminated with egg protein if the factory line is used to prepare other foodstuffs.

For the egg protein concentration, 50 random numbers were generated from 2000 Monte Carlo samples obtained using the BE method, and the posterior prediction distribution was obtained using a total of 100,000 random numbers. As food allergies are an immediate immune response, we did not consider dairy intake in the analysis. Therefore, the Food Safety Commission of Japan (2016) proposed an estimation method that uses the median of the probability density distribution as the value for single consumption [30]. Therefore, we adopted the median value as the value for single consumption for the estimation results of the probability density distribution. We obtained 2000 median values for food consumption as posterior distributions. Since the posterior distribution of the median value was bell-shaped, we generated 100,000 random numbers from the normal distribution $N(\widehat{\mu}_{MED}^{BE}, \widehat{\sigma}_{MED}^{BE})$ and used them for PS, where $\widehat{\mu}_{MED}^{BE}$ and $\widehat{\sigma}_{MED}^{BE}$ are the mean and the standard deviation of the posterior distribution of the median, respectively. Finally, the product of the concentration and single consumption was taken as the exposure to egg protein.

3. Results and discussion

3.1. Estimation of protein level of hidden egg protein in processed foods on prior evaluation

After the launch of the Japanese food allergen labeling system in 2001, the results of food allergen protein concentration tests in processed foods without allergen labeling conducted by local government research institutes were collected. The detection rate of egg protein concentration in confectionery items was 1.8 % (1/56), and the detection rate of egg protein concentration in all processed products was 8.2 % (20/243). The estimation parameters using the MLE method are shown in Table 1. The estimated distribution of the egg protein concentration in confectionery products showed a small relative standard deviation and an almost constant value. This was attributed to the low detection ratio and the fact that the detected value (0.1 ppm) was smaller than many LOQs (1–8 ppm). On the other hand, the estimated distribution of egg protein concentration for all processed foods showed a wide distribution with high relative standard deviations (410–620 %).

3.2. Egg protein concentration measurements

We then examined the hidden egg protein concentration in processed foods lacking specific labels for egg used as a raw material but with caution declarations outside the labeling margin; the results are shown in Fig. 1. As expected, the concentrations of egg protein in the processed foods that we examined accounted for less than 0.31 ppm; only dried fruit food showed a value of 0.37 ppm. Since the samples below the LOD accounted for 64 % of the total 92 samples, the results suggest that the food labeling system used for allergens in Japan is effective.

3.3. Bayesian estimates of egg protein concentration distribution in processed foods with precautionary declarations

Based on a comparison of the estimated WAIC for the egg protein concentration distribution, a gamma distribution was adopted to assess the egg protein concentration in confectionery items, and a Weibull distribution was then employed to assess the egg protein concentration in all processed foods (Table S1). Table 2 shows the BE results for the distribution parameters of the egg protein concentration. The mean egg protein concentration in confectionery items was estimated to be 0.062 ± 0.018 mg/g, and that in all processed foods was estimated to be 0.066 ± 0.027 mg/g. Finally, Fig. 2 shows the estimated histograms of egg protein concentration under the low intake scenario (confectionary items) (Fig. 2A) and the high intake scenario (beverages) (Fig. 2B), which are obtained by a two-dimensional Monte Carlo simulation using the probability density distribution with the smallest WAIC.

3.4. Estimation of intake for one meal

For confectionary items, the histograms showing daily consumption showed a unimodal distribution (Fig. 3A). According to comparisons of the estimated WAICs (Table S2), we adopted a gamma distribution for the distribution of confectionary consumption for target populations aged 1–15, 4–15, and 7–15 years old, and a generalized gamma distribution for the target population aged 16 years old or above. The estimated posterior distributions of the median value are shown in Fig. 4A. The estimated single consumption (median value of daily consumption) for confectionery items was 27.4 ± 0.3 g for the 1–15-year-old age group (Table 3).

The daily intake of beverages showed a bimodal distribution (Fig. 3B). The low side of the peak mainly consists of alcohol drinks (such as wine and sake) which were speculated as consumption of alcohol-containing seasonings. Since it is not possible to determine the lowest value of the seasoning, it was considered that a mixed distribution should be assumed for the probability density distribution of soft drinks, and that single consumption should be estimated from the results on the high side of the peak. Based on comparisons of the estimated WAICs (Table S2), we employed a mixed log-normal distribution for estimating daily beverage consumption for each target population. Fig. 4B shows a histogram of the posterior distribution of the median value of the peak for higher consumption of beverages. The estimated single consumption (median of daily consumption) of beverages was 264 ± 2 g for the 1–15-year-old age group (Table 3).

Although the estimated single consumption for beverages in the 1–15-year-old age group in this study is difficult to compare with that in other countries, the estimated single consumption of beverages in Japan appears to be larger than the median value (141 g) of the daily consumption among adolescents in 24 countries in the European region [31].

3.5. Egg protein exposure distribution in Japanese processed foods

By multiplying the allergen concentration distribution by the single consumption distribution for each food, we estimated the egg protein exposure distribution under low- and high-intake scenarios. Percentile values of the exposure distribution of egg protein are shown in Table 4, and Fig. 5 shows that histograms of egg protein exposure under (Fig. 5A) the low intake scenario and (Fig. 5B) the high intake scenario for different target populations. Applying MCS, we analyzed the mean, standard deviation, and percentile values of the distribution of egg protein exposure using a probabilistic assessment with an estimation of uncertainty.

Under the low-intake scenario, the mean exposure to egg protein was estimated to be 0.0018 mg for 1–15-year-olds, 0.0019 mg for 4–15-year-olds, 0.0020 mg for 7–15-year-olds, and 0.0022 mg for individuals aged 16 years old or above. Under the high-intake scenario, the mean exposure to egg protein was estimated to be 0.0164 mg for 1–15-year-olds, 0.0171 mg for 4–15-year-olds,

Table 1
Estimates obtained using a maximum likelihood function for egg protein concentration reported by local government research institutes.

Distribution	Low intake			High intake		
	Lognormal	Gamma	Weibull	Lognormal	Gamma	Weibull
Parameter 1[a]	0.10	2839	306	0.060	0.060	0.306
Parameter 2[b]	1.10	3.5×10^{-5}	0.10	6.80	12.3	0.051
Mean	0.10	0.10	0.10	0.374	0.743	0.438
SD	0.010	0.02	0.0004	2.32	3.03	2.28
Loglikelihood	3.73	5.36	7.02	−54.9	−73.4	−60.4

[a] Parameter 1 means e^{μ} for lognormal, α for gamma, or m for Weibull distribution.
[b] Parameter 2 means e^{σ} for lognormal, β for gamma, or η for Weibull distribution.

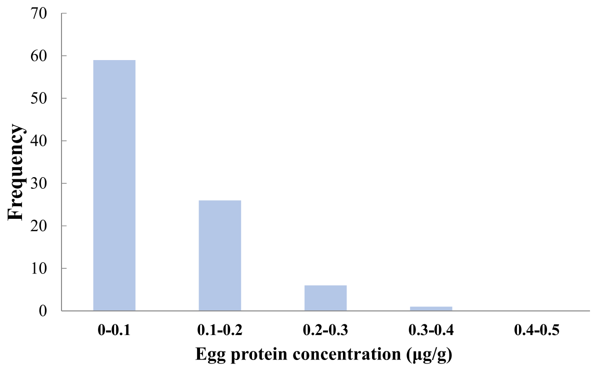

Fig. 1. Measurement of hidden egg protein concentration in processed foods without specific labels for egg as a raw material, but with cautionary declaration outside the labeling margin.
ND: not detected, <0.1 μg/g.

Table 2
Estimated results for the distribution of egg protein concentration in processed food obtained using the Bayesian estimation method.

Distribution	Low intake	High intake
	Gamma	Weibull
Parameter 1[a]	0.438 ± 0.328	0.669 ± 0.207
Parameter 2[b]	7.30 ± 5.21	0.043 ± 0.019
Mean	0.066 ± 0.027	0.062 ± 0.018
SD	0.156 ± 0.068	0.116 ± 0.088

[a] Parameter 1 means α for gamma distribution or m for Weibull distribution.
[b] Parameter 2 means β for gamma distribution or η for Weibull distribution.

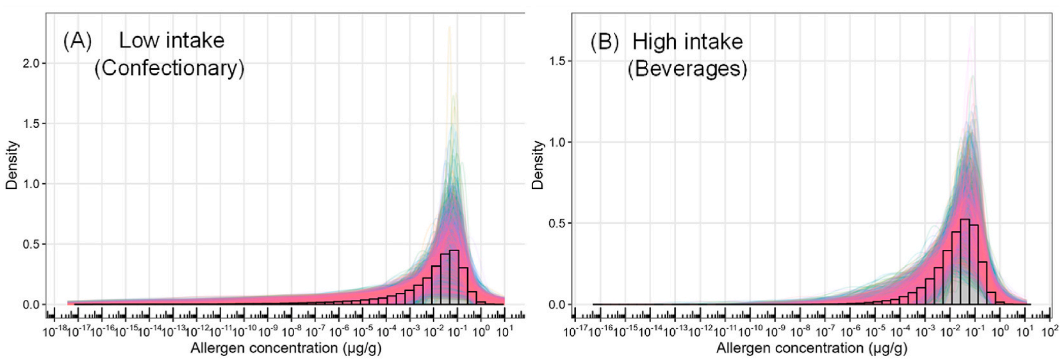

Fig. 2. Estimated histograms of egg protein concentration under (A) the low intake scenario (confectionary items) and (B) the high intake scenario (beverages). Fifty random numbers were generated by two-dimensional Monte Carlo simulations from 2000 parameters estimated by Bayesian estimation. Colors indicate density curves obtained from 2000 parameters.

0.0181 mg for 7–15-year-olds, and 0.0188 mg for individuals aged 16 years old or above.

3.6. Risk assessment for egg protein exposure

Recently, an *ad hoc* Joint FAO/WHO Expert Committee reached a consensus on reference doses (RfD) for priority allergenic foods [9]. The RfD recommendation for eggs is 2.0 mg total protein from the allergenic source. If a person ingested hidden egg protein in a certain processed food under the high-intake scenario under the Japanese allergen labeling system, then the mean intake would be 0.019 mg per single consumption, and the 97.5-percentile intake would be as high as 0.1 mg. The average intake and the 97.5-percentile intake for the highest age group are lower than the international RfD of 2.0 mg.

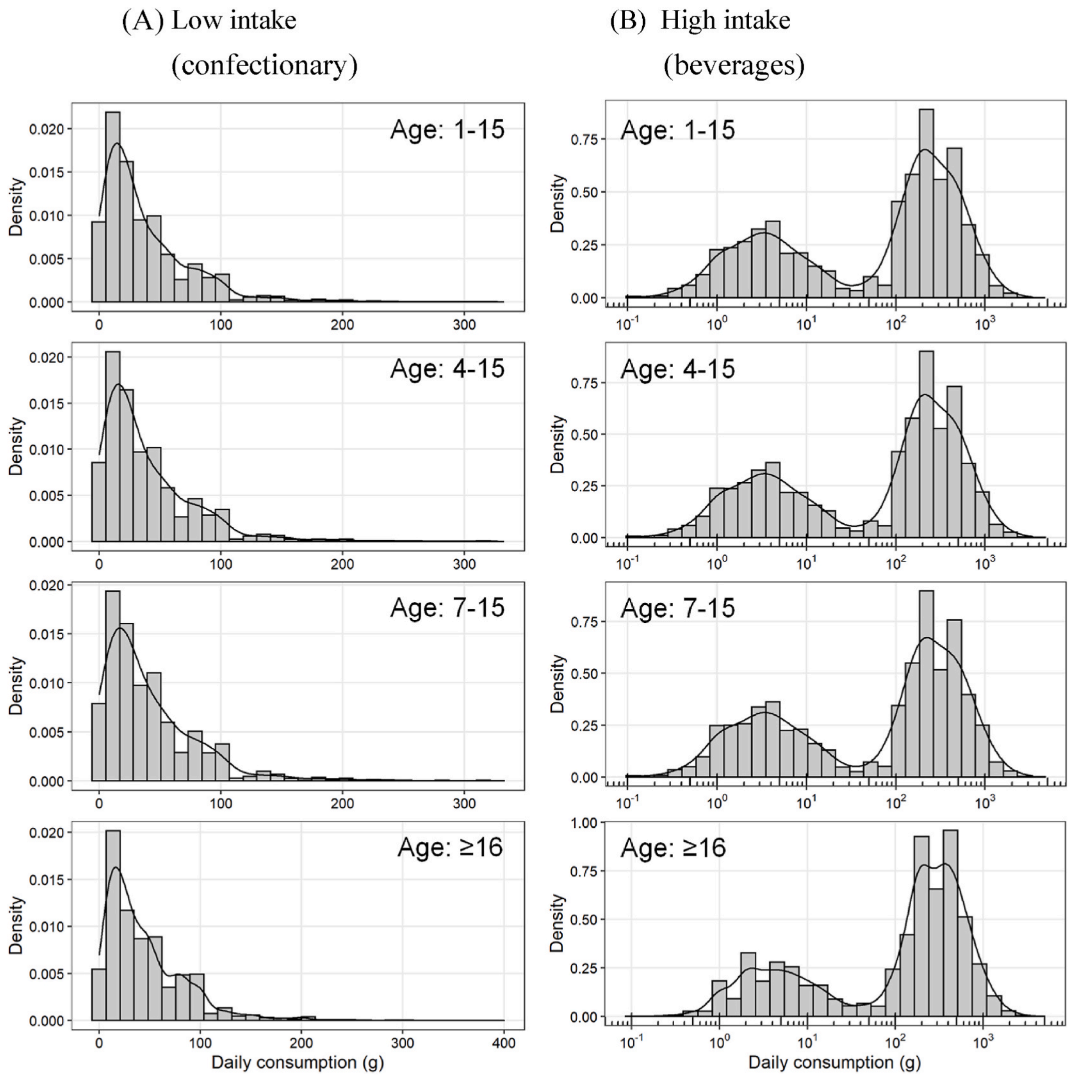

Fig. 3. Histograms of daily consumption of processed food under (A) the low intake scenario (confectionary items) and (B) the high intake scenario (beverages) for different target populations (1–15, 4–15, 7–15, ≥16 years old).

3.7. Future perspectives

Non-detected values frequently occur for chemical concentrations in food products due to the inherent limitations of analytical methods, characterized by LOD and the LOQ. Data that lack certain values, which fall below the LOD and LOQ, are referred to as 'left-censored'. Left-censored data complicate the calculation of summary statistics and are a major source of uncertainty in exposure assessments. Although European Food Safety Authority (EFSA) [32] and Environmental Protection Agency (EPA) [33] have raised the issue of statistical treatment of left-censored data in exposure estimation, numerous studies have persisted using the substitution method. Bayesian estimation has been reported to have less bias, small uncertainly, and good reliability compared to the non-Bayesian methods (substitution, non-parametric, semi-parametric, and parametric) in the context of left-censored data analysis [34–38].

Conversely, Wheeler et al. [11] have not only demonstrated the application of Bayesian estimation for survival regression, which accounts for censoring, but have also advocated for the use of Model Averaging to account for the uncertainty associated with model selection. In the present study, to narrow the technical divide between estimating exposure distribution and susceptibility distribution, we initially estimated the concentration distribution using a Bayesian method that considers left-censoring and generated posterior predictive distributions. Given that the WAICs were comparatively close across models (Table S1), incorporating Bayesian Model Averaging to consider the uncertainty in selecting probability density distributions is anticipated to yield more accurate predictions of allergen concentrations. This aspect warrants further investigation.

Furthermore, in this study, values derived from the calibration curve were assigned to measured values that were equal to or greater than the LOD and less than the LOQ. Due to the substantial uncertainties, employing cumulative probabilities for values

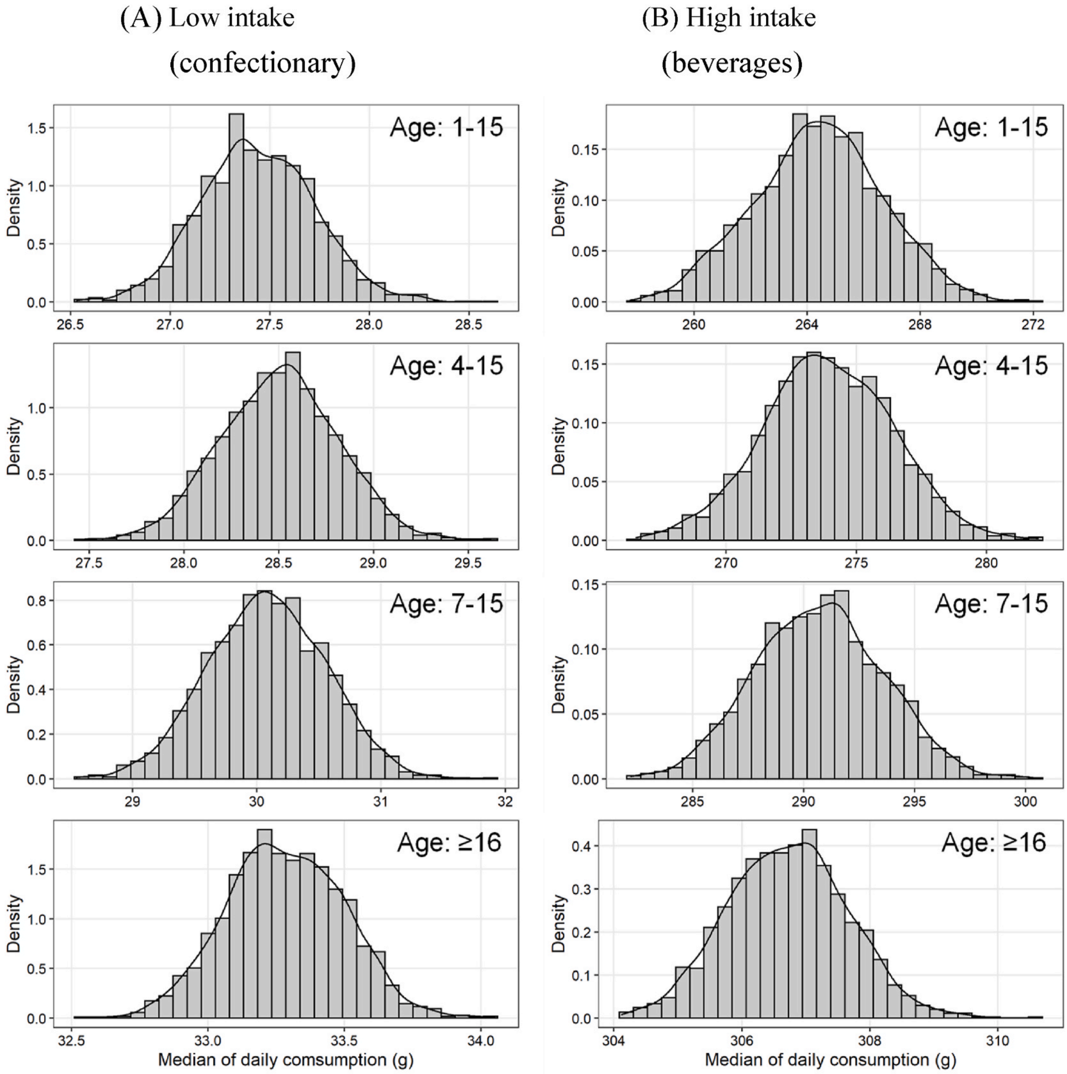

Fig. 4. Estimated posterior distribution of median of daily food consumption under (A) the low intake scenario and (B) the high intake scenario for different target populations (1–15, 4–15, 7–15, ≥16 years old).

Table 3
Estimated median value for daily food consumption (g) obtained using the Bayesian estimation method.

Age group	Low intake	High intake[a]
1–15	27.4 ± 0.3	264 ± 2
4–15	28.5 ± 0.3	274 ± 2
7–15	30.1 ± 0.5	291 ± 3
≥16	33.3 ± 0.2	307 ± 1

[a] Median value of the peak for higher food consumption.

between LOD and LOQ, rather than direct quantitative measures, may offer an alternative approach.

4. Conclusions

We performed an exposure assessment of hidden egg protein in processed foods under the Japanese food labeling system. The findings of this study show that processed food products in Japan, including confectionery items, are generally appropriately labeled. The risk of triggering allergic reactions due to egg protein contamination from foods without ingredient labels in the Japanese food

Table 4
Estimated exposure of egg protein under two scenarios in different target populations.

Scenario	Age group	Mean (mg)	SD (mg)	2.5 % (mg)	50 % (mg)	97.5 % (mg)
Low intake	1–15	0.0018	0.0043	2.1×10^{-13}	5.0×10^{-4}	1.1×10^{-2}
	4–15	0.0019	0.0045	2.2×10^{-13}	5.2×10^{-4}	1.1×10^{-2}
	7–15	0.0020	0.0048	2.3×10^{-13}	5.5×10^{-4}	1.2×10^{-2}
	≥16	0.0022	0.0053	2.5×10^{-13}	6.1×10^{-4}	1.3×10^{-2}
High intake	1–15	0.0164	0.0359	9.3×10^{-6}	6.3×10^{-3}	9.0×10^{-2}
	4–15	0.0171	0.0373	9.6×10^{-6}	6.6×10^{-3}	9.3×10^{-2}
	7–15	0.0181	0.0396	10.0×10^{-6}	7.0×10^{-3}	9.9×10^{-2}
	≥16	0.0188	0.0410	10.5×10^{-6}	7.2×10^{-3}	10.2×10^{-2}

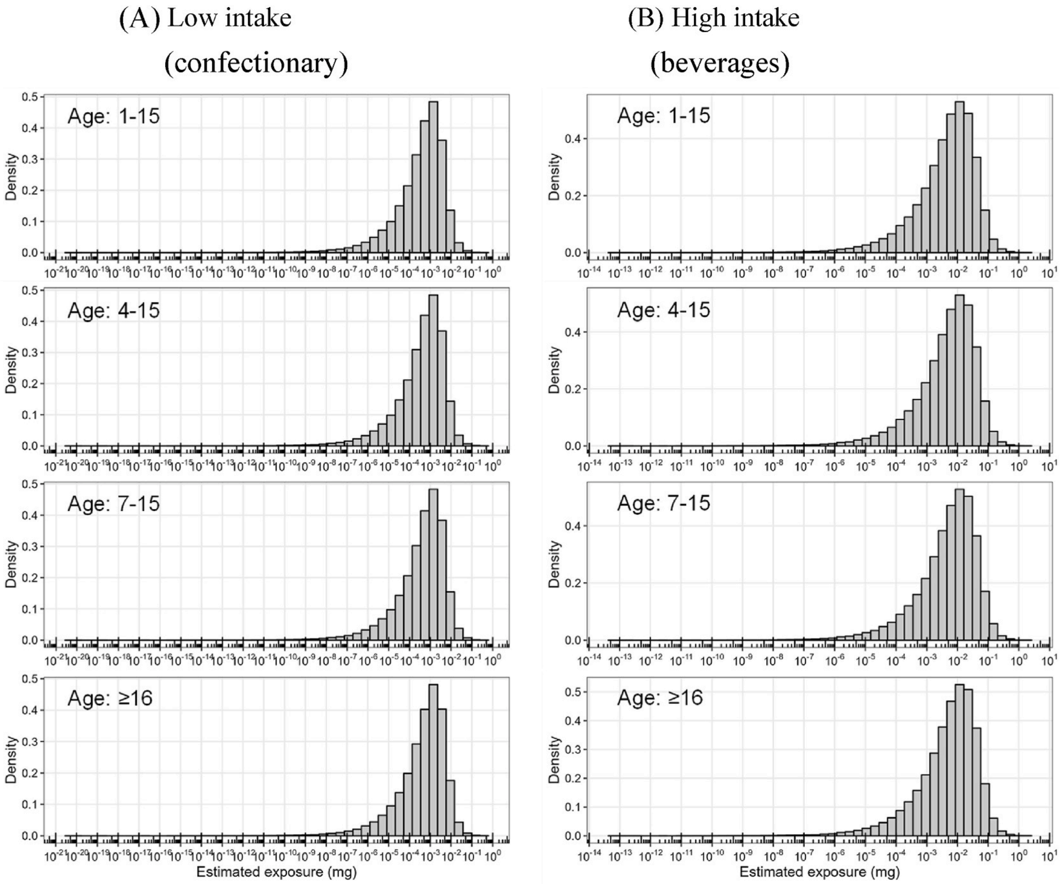

Fig. 5. Estimated distribution of egg protein exposure under (A) the low intake scenario and (B) the high intake scenario for different target populations (1–15, 4–15, 7–15, ≥16 years old).

allergen labeling system is also considered to be extremely low and is considerably less than international RfDs. The findings also show that the Japanese food industry might be able to decrease the usage of cautionary descriptions outside the labeling margin by using the quantitative risk assessment approach recommended by FAO/WHO, by using the VITAL program in Australia and New Zealand.

Data availability statement

The data are included in article and supplementary materials in this article.

CRediT authorship contribution statement

Hiroshi Akiyama: Writing – original draft, Validation, Project administration, Methodology, Formal analysis, Data curation, Conceptualization. **Yoshinari Suzuki:** Writing – original draft, Methodology, Formal analysis, Data curation, Conceptualization. **Reiko Adachi:** Writing – review & editing, Validation, Data curation. **Momoka Kadokura:** Formal analysis, Data curation. **Asaka**

Takei: Formal analysis, Data curation. **Masayoshi Tomiki:** Writing – review & editing, Validation. **Kosuke Nakamura:** Writing – review & editing, Validation. **Rie Ito:** Writing – review & editing, Validation. **Yusuke Iwasaki:** Writing – review & editing, Validation. **Clare Mills:** Writing – review & editing. **Yukihiro Ohya:** Writing – review & editing, Project administration, Conceptualization. **Tatsuki Fukuie:** Writing – review & editing, Project administration, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors acknowledge the assistance and advice offered by Dr. Hirohisa Saito, National Center for Child Health and Development Division of General Allergy, and Dr. Fumi Irie, Health Care Administration and Management, Kyushu University. This study was supported in part by grants from the Food Safety Commission, Cabinet Office (Research Program for Risk Assessment Study on Food Safety, Grant No. JPCAFSC20202008), the Ministry of Health, Labour and Welfare of Japan (Grant No. 23KA1007), and a grant from The Japan Food Chemical Research Foundation.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2024.e33545.

References

[1] C.M. Warren, J. Jiang, R. Gupta, Epidemiology and burden of food allergy, Curr. Allergy Asthma Rep. 20 (2020) 1–9, 2020, https://doi:10.1007/s11882-020-0898-7.
[2] M. Ebisawa, K. Ito, T. Fujisawa, Japanese guidelines for food allergy 2020, Allergol, Int. 69 (2020) 370–386, https://doi.org/10.1016/j.alit.2020.03.004.
[3] C.B. Madsen, et al., Can we define a level of protection for allergic consumers that everyone can accept? Regul. Toxicol. Pharmacol. 117 (2020) 104751 https://doi.org/10.1016/j.yrtph.2020.104751.
[4] H. Akiyama, R. Adachi, Japanese food allergy-labeling system and comparison with the international experience; detection and thresholds, Food Safety 7 (2021) 101–116, https://doi.org/10.14252/foodsafetyfscj.d-21-00008.
[5] K.J. Allen, et al., Allergen reference doses for precautionary labeling (VITAL 2.0): clinical implications, J. Allergy Clin. Immunol. 133 (2014) 156–164, https://doi.org/10.1016/j.jaci.2013.06.042.
[6] C. Bindslev-Jensen, M. Briggs, D. Osterballe, Can we determine a threshold level for allergenic foods by statistical analysis of published data in the literature? Allergy 57 (2002) 741–746, https://doi.org/10.1034/j.1398-9995.2002.23797.x.
[7] J.O. Hourihane, et al., Peanut Allergen Threshold Study (PATS): novel single-dose oral food challenge study to validate eliciting doses in children with peanut allergy, J. Allergy Clin. Immunol. 139 (2017) 1583–1590, https://doi.org/10.1016/j.jaci.2017.01.030.
[8] S.L. Taylor, et al., Establishment of reference doses for residues of allergenic foods: report of the VITAL Expert panel, Food Chem. Toxicol. 63 (2014) 9–17, https://doi.org/10.1016/j.fct.2013.10.032.
[9] B.C. Remington, et al., Updated population minimal eliciting dose distributions for use in risk assessment of 14 priority food allergens, Food Chem. Toxicol. 139 (2020) 111259. https://doi:10.1016/j.fct.2020.111259.
[10] G.F. Houben, et al., Full range of population Eliciting Dose values for 14 priority allergenic foods and recommendations for use in risk characterization, Food Chem. Toxicol. 146 (2020) 111831. https://doi:10.1016/j.fct.2020.111831.
[11] M.W. Wheeler, et al., Bayesian stacked parametric survival with frailty components and interval-censored failure times: an application to food allergy risk, Risk Anal. 41 (2021) 56–66, https://doi.org/10.1111/risa.13585.
[12] FAO and WHO, Summary report of the ad hoc Joint FAO/WHO Expert consultation on risk assessment of food allergens. Part 2: review and establish threshold levels in foods of the priority allergens. https://cdn.who.int/media/docs/default-source/food-safety/jemra/2nd-allergen-summary-report-20aug2021.pdf?sfvrsn=915a8417_8, 2021.
[13] B.C. Remington, et al., Quantitative risk assessment of foods containing peanut advisory labeling, Food Chem. Toxicol. 62 (2013) 179–187. https://doi:10.1016/j.fct.2013.08.030.
[14] L. Rimbaud, et al., Quantitative risk assessment relating to the inadvertent presence of peanut allergens in various food products, Int. Food Risk Anal. J. 3 (2013) 3. https://doi:10.5772/56646.
[15] A. Crépet, et al., Peanut traces in food: a probabilistic risk assessment based on the French MIRABEL survey, Food Control 131 (2022) 108403, https://doi.org/10.1016/j.foodcont.2021.108403.
[16] J. Touma, et al., Sesame as an allergen in Lebanese food products: occurrence, consumption and quantitative risk, Food Chem. Toxicol. 156 (2021) 112511, https://doi.org/10.1016/j.fct.2021.112511.
[17] National Institute of Health and Nutrition, Report of Ministry of Health, Labour and Welfare in Japan, 2010. https://www.mhlw.go.jp/file/06-Seisakujouhou-11130500-Shokuhinanzenbu/0000170577.pdf.
[18] H. Akiyama, et al., Summary of inspection of allergen-containing foods in Kanagawa prefecture accumulated in past 10 years (fiscal year 2006-2015), Bull. Kanagawa Ins. Of P. H. 46 (2016) 46–48.
[19] Y. Asada, et al., Survey of foods containing allergic substances in Okayama prefecture 2014~2015, Ann. Rep. Okayama Ins. For Environ. Sci. Pub. Health 40 (2016) 119–123.
[20] K. Hagino, et al., Examination of allergic substance in foods (april 2015-march 2017), Ann. Rep. Tokyo Metr. Inst. Pub. Health 68 (2017) 137–141.
[21] K. Kimoto, et al., Examination of allergic substance in foods (april 2017-march 2020), Ann. Rep. Tokyo Metr. Inst. Pub. Health 71 (2020) 165–171.
[22] Y. Sugano, K. Aotsuka, T. Suzuki, The monitoring test for allergenic materials in processed foods produced in Hokkaido in fiscal years 2014 and 2015, Rep. Hokkaido Inst. Public Health 66 (2016) 17–21.
[23] Y. Sugano, K. Aotsuka, T. Suzuki, The monitoring test for allergenic materials in processed foods produced in Hokkaido in fiscal years from 2016 to 2019, Rep. Hokkaido Inst. Public Health 70 (2020) 25–31.
[24] H. Akaie, Entropy and the likelihood of a model, Butsuri 35 (1980) 608–614.

[25] D.R. Helsel, Statistics for Censored Environmental Data Using Minitab and R, second ed., Wiley, 2012, 978-0-470-47988-9.
[26] A. Gelman, et al., A weakly informative default prior distribution for logistic and other regression models, Ann. Appl. Stat. 2 (2008) 1360–1383, https://doi.org/10.1214/08-aoas191.
[27] B. Carpenter, et al., Stan: a probabilistic programming language, J. Stat. Software 76 (2017) 1–32, https://doi.org/10.18637/jss.v076.i01.
[28] S. Watanabe, Asymptotic equivalence of Bayes cross validation and widely applicable information criterion in singular learning theory, J. Mach. Learn. Res. 11 (2010) 3571–3594, https://doi.org/10.48550/arXiv.1004.2316.
[29] A. Vehtari, A. Gelman, J. Gabry, Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC, Stat. Comput. 27 (2017) 1413–1432, https://doi.org/10.1007/s11222-016-9696-4.
[30] Food Safety Commission of Japan, Report of Research Program for Risk Assessment Study and Survey Program on Food Safety (2016). https://www.fsc.go.jp/fsciis/technicalResearch/show/cho99920151508.
[31] J. Walton, A. Wittekind, Soft drink intake in europe—a review of data from nationally representative food consumption surveys, Nutrients 15 (2023) 1368, https://doi.org/10.3390/nu15061368.
[32] European Food Safety Authority, Management of left-censored data in dietary exposure assessment of chemical substances, EFSA J. 8 (2010) 1557. https://doi:10.2903/j.efsa.2010.1557.
[33] EPA, ProUCL Version 5.1 Technical Guide, Office of Research and Development, Washington, DC, 2015, EPA/600/R-07/041. https://www.epa.gov/sites/production/files/2016-05/documents/proucl_5.1_tech-guide.pdf.
[34] T. Huynh, H. Quick, et al., A comparison of the β-Substitution method and a Bayesian method for analyzing left-censored data, Ann. Occup. Hyg. 60 (2016) 56–73. https://doi:10.1093/annhyg/mev049.
[35] L. Nie, H. Chu, S. Feng, Estimating variance parameters from multivariate normal variables subject to limit of detection: MLE, REML, or Bayesian approaches? Stat. Med. 28 (2009) 2605–2616. https://doi:10.1002/sim.3644.
[36] D.H.A. Hady, M.S. Rania, Bayesian and non Bayesian parameter estimation for bivariate pareto distribution based on censored samples, Int. J. Eng. Res. Afr. 5 (2015) 1–6.
[37] N. Feroze, M. Aslam, Comparison of Bayesian and maximum likelihood estimation for analysis of burr type V distribution under left censored samples, J. Stat. 22 (2015) 116–192.
[38] Y. Suzuki, N. Tanaka, H. Akiyama, Attempt of bayesian estimation from left-censored data using the Markov chain Monte Carlo method: exploring Cr(VI) concentrations in mineral water products, Food Safety 8 (2020) 67–89. https://doi:10.14252/foodsafetyfscj.D-20-00007.