Finnish Food Safety Authority Evira

Evira

Bayesian estimation of the total intake of chemical contaminants from multiple food products

Mikkelä A., Ranta J., Pasonen P., Suomi J., Hirvonen T., Suominen K., and Tuominen P.
Finnish Food Safety Authority, Risk Assessment Research Unit, Helsinki, FINLAND

Introduction

The Bayesian model was built to estimate the acute as well as the chronic intake of chemical contaminants due to food consumption. The model was applied to estimate the cadmium intake from grains, meat, fish and milk. The food consumption data were based on the 48-h dietary recall of 2038 adults, and cadmium concentration data were drawn from the control program.

Modeling

Concentration

The cadmium concentration in each food product was modeled using a log-normal distribution. The measurements below the limit of detection (LOD) were treated as censored data in order to achieve optimum estimates. In this model, there is no need to use artificial lower bound, middle bound or upper bound for small concentrations.[1]

Consumption

The long-term mean consumption of each food product was obtained as a combination of the long-term mean daily serving size (days when used) and consumption frequency. The long-term mean daily serving of each food product was modeled as a hierarchical log-normal model in order to take into account the within-individual as well as between-individual variation. Consumption frequency was estimated by binomial-logit model with random effects for individuals[2]. Multinormal prior distributions were assigned for the individual parameters (mean daily serving and frequency) to take into account the correlation in consumption between different food products.

Full likelihood function of the concentration-consumption data

$$\overbrace{\underbrace{\prod_{f=1}^{F}\prod_{k=1}^{K_f} P(C_{f,k}\mid \mu_f,\tau_f^c)^{(1-\delta_{f,k})}\cdot P(C_{f,k}<\mathrm{LOD}\mid \mu_f,\tau_f^c)^{\delta_{f,k}}}_{\substack{\text{Log-normal}\\ \text{(Concentration)}}}}^{\text{Likelihood}}\cdot$$

$$\overbrace{\underbrace{\prod_{i=1}^{I}\prod_{f=1}^{F}\prod_{t=1}^{T} P(D_{i,f,t}\mid m_{i,f},\tau_f^d)\cdot P(\{m_i\}\mid\{m_0\},\{T^d\})}_{\substack{\text{Hierarchical Log-normal}\\ \text{(Long-term mean daily serving)}}}}^{\text{Likelihood}}\cdot$$

$$\overbrace{\underbrace{\prod_{i=1}^{I}\prod_{f=1}^{F} P(X_{i,f}\mid p_{i,f},N_i)\cdot P(\{p_i\}\mid\{p_0\},\{T^f\})}_{\substack{\text{Binomial-logit}\\ \text{(Consumption frequency)}}}}^{\text{Likelihood}}$$

Binary variable $\delta_{f,k}$ denotes whether a concentration was exact or below LOD.

Intake

The joint posterior distribution of all unknown model parameters is the product of the likelihood function and the joint prior distribution of all unknown model parameters. Once the posterior has been obtained the predictions for quantities of interest can be made. The chronic total intake (CTI) of a individual (*i'*) was obtained by the combination of the estimated mean concentration and long-term mean consumption:

$$\mathrm{CTI}_{i'}=\sum_{f=1}^{F}E(D_{i',f})\cdot p_{i',f}\cdot E(C_f)$$

$E(D_{i',f})$ = long-term mean daily serving, $p_{i',f}$ = consumption frequency,

$E(C_f)$ = mean concentration, i' = individual, f = food product, F = number of food products

Technical implementation of the model:

1. Data in Excel-format
2. Loading the Excel file into R software
3. Data modification in R
4. Model written in OpenBUGS[3] software
5. Computation of the OpenBUGS model via R
6. Results can be handled in R

Results

The proposed model provides a proper way to assess the intake of chemical contaminants from a single food product as well as total diet taking into account individual variation. The model was made to be flexible for different requirements of research and data sets. The concentration model performs well even under limited data set and heavy censoring. The mean chronic cadmium intake from all studied food products was predicted to be around 50 ng/kg bw/day in Finnish adult population (see Figure 1). The estimate is relatively small compared to a tolerable daily intake (357 ng/kg bw/ day)[4]. However, all possible sources of cadmium were not included in the estimation. The majority of total cadmium intake was estimated to be due to wheat consumption (see Figure 2).

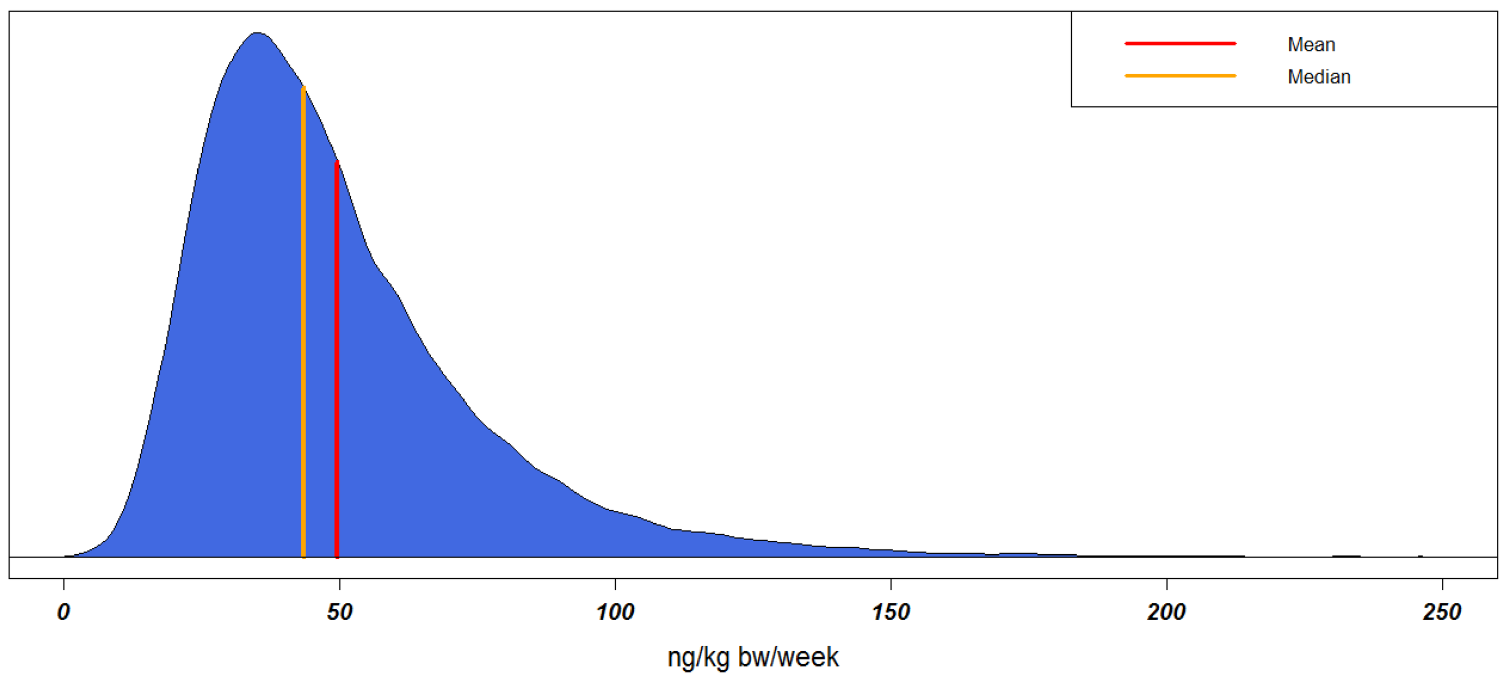

Figure 1. Posterior distribution for the chronic cadmium intake from all studied food products. Distribution desribes the estimated variation between different individuals in the chronic intake, accounting parameter uncertainty.

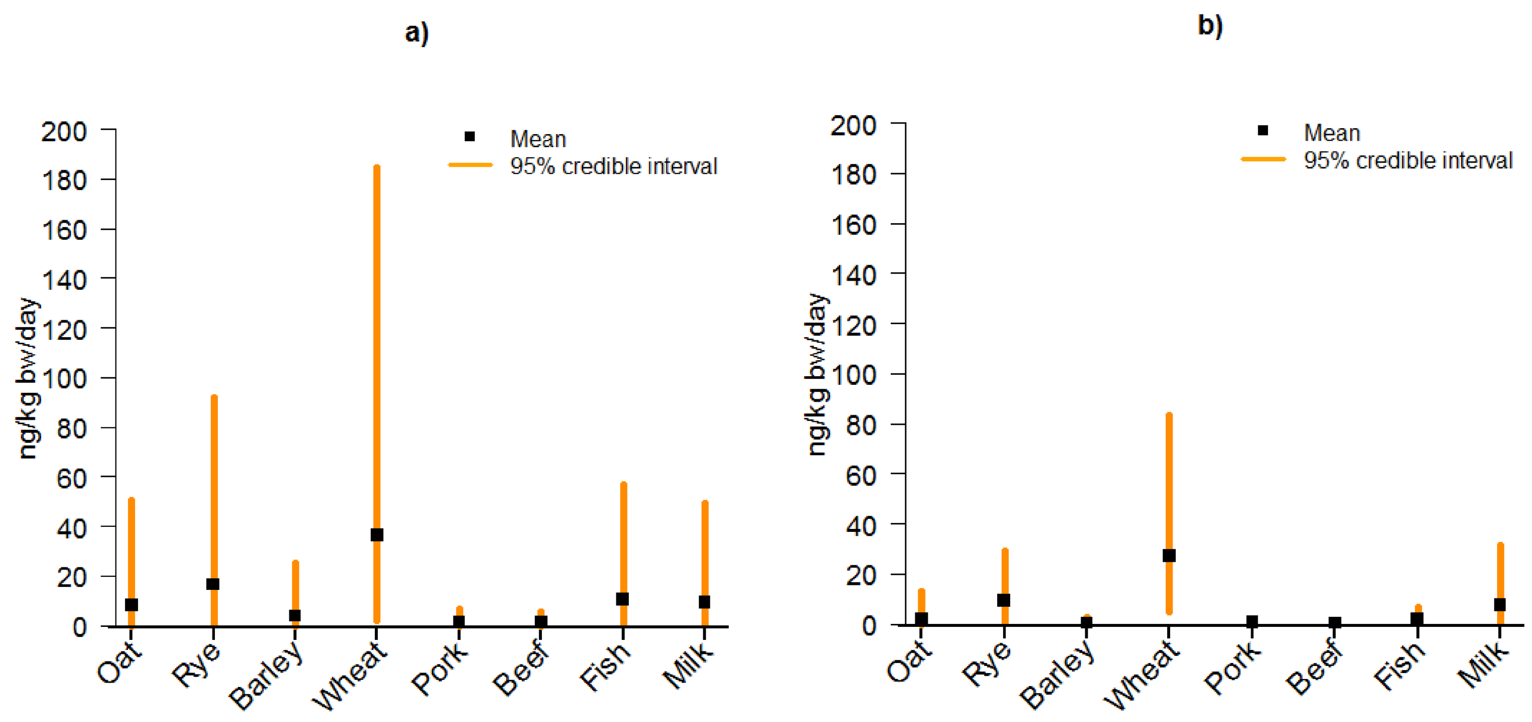

Figure 2. Mean and 95% credible interval for the acute (a) and chronic (b) cadmium intake from each food product. Credible intervals describe the estimated variation between different individuals in intakes, accounting parameter uncertainty.

Discussion

Bayesian methods fit well for this type of modeling because uncertainty about all unknown model parameters can be handled at once, and the resulting posterior distribution presents the total uncertainty and variability in the unknown quantities. In future, the proposed model could be expanded to estimate the joint intake of several contaminants at once. The modeling of the possible correlation between the consumption amount and consumption frequency is another issue that needs to be further studied. The model could also be converted into a tool that anyone can use.

References

1. Mikkelä A, Ranta J, González M, Hakkinen M, Tuominen P. Campylobacter QMRA: a Bayesian Estimation of Prevalence and Concentration in Retail Foods Under Clustering and Heavy Censoring. Risk Analysis. 2016; 36(11): 2065–2080.
2. Slob W. Probabilistic dietary exposure assessment taking into account variability in both amount and frequency of consumption. Food and Chemical Toxicology. 2006; 44: 933–951.
3. Lunn DJ, Jackson C, Best N, Thomas A, Spiegelhalter D. The BUGS Book: A Practical Introduction to Bayesian Analysis. Boca Raton: CRC Press, 2013.
4. Alexander J. et al. Scientific opinion. Cadmium in food. Scientific opinion of the Panel on Contaminants in the Food Chain. EFSA Journal. 2009; 980: 1-139.