Report

Evaluation of the Probabilistic Distribution of Dietary Biotin Intake in Japan Using Monte Carlo Simulation

Taro MURAKAMI, Kouichi TAKAKURA and Tetsuo YAMANO

Osaka City Institute of Public Health and Environmental Sciences, 8–34 Tojo-cho, Tennoji-ku, Osaka, Osaka 543–0026, Japan
(Received April 13, 2010)

Summary

Biotin is a widely distributed water soluble vitamin. Adequate intake of biotin was set at 50 µg/d in Japan 2010. Recently, the importance of the application of probabilistic techniques to estimate the share of the population at risk of deficient and excessive nutrient intake has been increasingly emphasized for assessing nutrient adequacy. Monte Carlo simulation, a computer-based method of analysis that uses statistical sampling techniques yielding a probabilistic approximation to the solution of a mathematical model, has been used to estimate the probabilistic distribution of the dietary intake of food chemicals. For this study, we used two preliminary models to estimate the dietary biotin intake with food consumption data based on the National Health and Nutrition Survey in Japan. One is evaluated by biotin concentration data from the total diet study; the other is a dataset of biotin concentration in individual foods. After removing outliers from the individual foods dataset, probability density distributions from two models showed analogous mean, median, 5th percentile, and 95th percentile values. The daily biotin intakes from these probabilistic methods showed that more than 80% of the Japanese population had higher than the adequate intake of biotin. However, the contribution of each food group to the total daily biotin intake was somewhat different. Improvement of these methods necessitates the collection of more actual data associated with sample compositional variability and evaluation of uncertainty associated with the food group classification of biotin.

Key Words
dietary biotin intake, probabilistic distribution, nutrient adequacy, Monte Carlo simulation

Biotin is an important water-soluble vitamin that is a cofactor of enzymes involved in carboxylation and decarboxylation reactions. Human cells have four biotin-dependent carboxylases that catalyze key reactions in gluconeogenesis, amino acid catabolism, and fatty acid synthesis (1). Biotin deficiency has been reported in various circumstances such as those following prolonged feeding of undenatured egg whites (2) and in cases of genetic biotinidase deficiency (3). Symptoms of biotin deficiency include hair loss, skin rash, and neurological symptoms. In biotin-deficient infants, hypotonia, lethargy, and developmental delay are all characteristics of a neurological disorder resulting from a lack of biotin (4). The recommended dietary allowance (RDA) of biotin that is sufficient for the nutritional requirement of nearly all healthy individuals has not been determined for Japan. Although data about dietary intake are necessary to estimate the RDA, insufficient data are available to estimate the dietary biotin intake. Because of the lack of a standard table of food composition for biotin in Japan, the dietary biotin intake could not be estimated by a dietary assessment such as a diet record method, food frequency method, or diet recall method. Several Total Diet Studies (TDSs), also known as Market Basket Studies, used a method to estimate the dietary intake of food chemicals (5–7). Recently, TDSs have been applied to estimate the mean daily biotin intake in Japan (8–10). In a previous TDS, we estimated the mean daily biotin intake and its measurement uncertainty as 70.1±11.2 µg/d (mean±expanded uncertainty; k=2, 95% confidence interval) (9). In other TDSs conducted in Japan, Saitoh and Ushio estimated a daily biotin intake of 45.1 µg/d (8), and Watanabe and Taniguchi estimated a daily biotin intake of 50.7 µg/d (10). In Japan, adequate intake (AI) of biotin was set at 50 µg/d in 2010 (11), this value is from mean daily biotin intake data estimated from these TDSs.

However, the mean intake data are insufficient to estimate the population at risk of deficient and excessive nutrient intake. At least for some nutrients such as biotin, the importance of variation and uncertainty has emphasized the need for assessing nutrient adequacy for a population. Recently, Monte Carlo simulation has been used to estimate the probabilistic distribution of the dietary intake of food chemicals (12, 13). Monte Carlo simulation is a computer-based analytical method that uses statistical sampling techniques obtaining a

E-mail: tar-murakami@city.osaka.lg.jp
probabilistic approximation to the solution of a mathematical model. As described in this paper, we conducted two preliminary Monte Carlo simulation models to estimate the dietary biotin intake. In the first model, biotin concentration data obtained from TDSs 2007–2009 and food consumption data based on the National Health and Nutrition Survey (NHNS) in Japan 2005 (14) were used. In the second model, a dataset of biotin concentration in 930 individual foods was used in place of TDS data. Using these two models, the mean daily biotin intake and its probabilistic distribution among a population were investigated.

**Methods**

**Food consumption data.** The National Health and Nutrition Survey (NHNS) 2005 (14) was used for this study. For adaptation to the food groups employed in our TDS, data obtained from 2,759 males and 3,106 females, aged 18–69 y were classified into 13 food groups based on TDS in Japan (7). The mean daily intake and its standard deviation in each group are presented in Table 1.

**Total diet study.** About 200 food samples for TDS were purchased in Osaka City each year during 2007–2009. Each food sample, cooked when necessary, was combined into 13 food groups according to the proportions determined by the NHNS. Biotin concentrations in TDS samples were analyzed after hydrolyzed extraction with microbiological assay according to a method used in a previous study (9). Mean biotin concentration and the number of food samples in each group are shown in Table 2.

**Dataset of biotin concentration.** We used biotin concentration data in 930 individual foods from Japan (15), Denmark (16), Germany (17), Canada (18–21), and the USA (22). Individual foods were classified into one of 13 food groups based on TDS in Japan. A summary of the biotin concentration and the number of food samples in each group are shown in Table 3.

**Monte Carlo simulation.** The basic calculation that is iterated within the Monte Carlo simulation is a simple calculation:

\[ T = \sum_{i=1}^{n} C(i) \times I(i), \]

where \( T \) signifies total biotin intake in 1 d by a person, \( i \) represents the number of food groups, \( C \) is the biotin concentration in each food group, and \( I \) denotes the amount of food consumption in each food group. The daily biotin intake was obtained as the sum of biotin

| No. | Food group \(^1\) | Daily intake (g/d), \( I(i) \) |
|-----|-----------------|------------------|
| 1 | Rice/processed rice | 2007: 0.014 (4) | 2008: 0.003 (4) | 2009: 0.007 (4) |
| 2 | Cereal/potato | 2007: 0.021 (22) | 2008: 0.021 (22) | 2009: 0.046 (21) |
| 3 | Sugar/confectionery | 2007: 0.040 (19) | 2008: 0.034 (19) | 2009: 0.048 (18) |
| 4 | Fat/Oil | 2007: 0.000 (8) | 2008: 0.000 (6) | 2009: 0.000 (8) |
| 5 | Pulse | 2007: 0.072 (10) | 2008: 0.070 (10) | 2009: 0.060 (9) |
| 6 | Fruit | 2007: 0.013 (13) | 2008: 0.013 (13) | 2009: 0.016 (17) |
| 7 | Green, yellow vegetables | 2007: 0.030 (14) | 2008: 0.047 (12) | 2009: 0.028 (18) |
| 8 | Other vegetables | 2007: 0.016 (23) | 2008: 0.015 (26) | 2009: 0.025 (24) |
| 9 | Beverage | 2007: 0.006 (14) | 2008: 0.005 (16) | 2009: 0.004 (15) |
| 10 | Fish/shellfish | 2007: 0.072 (24) | 2008: 0.111 (27) | 2009: 0.103 (18) |
| 11 | Meat/egg | 2007: 0.238 (8) | 2008: 0.198 (15) | 2009: 0.215 (7) |
| 12 | Milk/dairy products | 2007: 0.029 (11) | 2008: 0.023 (10) | 2009: 0.062 (11) |
| 13 | Seasoning | 2007: 0.067 (20) | 2008: 0.034 (20) | 2009: 0.100 (25) |

| Total n | 190 | 200 | 195 |

\(^1\) Numbers in parentheses show numbers of food samples in respective groups.
intakes through the 13 food groups. The calculation equation was performed 10,000 times using the R statistical package (ver. 2.10.1; The Comprehensive R Archive Network) (23, 24).

**Monte Carlo simulation model 1**

A subject’s daily biotin intake (T) was calculated as

$$T = \sum_{i=1}^{13} C_{INF}(i) \times I(i) \, (\mu g/d),$$

where $C_{INF}(i)$ stands for the biotin concentration in each food group, which was one randomly sampled value from 3 y of mean biotin concentration data in TDS, and $I(i)$ is a parametric sampling from the distribution of the amount of food consumption. The mean and standard deviations are those in NHNS. Its distribution was assumed to be log-normal (range from zero to five times the mean intake based on NHNS).

**Monte Carlo simulation model 2-1**

For this model, it was assumed that a person randomly chose and consumed three individual foods from each food group in a day. A subject’s daily biotin intake (T) was calculated as

$$T = \sum_{i=1}^{13} \frac{C_{INF}(i) + 2 \times C_{INF}(i)}{3} \times I(i) \, (\mu g/d),$$

where $C_{INF}(i)$ is biotin concentration in individual foods that were three randomly sampled values from a dataset in each group, $I(i)$ is the amount of food consumption, as in model 1.

**Monte Carlo simulation model 2-2**

In the model described above, all data of reported values were used in random sampling. However, a broad range of biotin concentrations was found even within the same food group in the dataset (Table 3). For example, in the beverage group, the smallest value was 0.001 $\mu g/g$ (rice wine, distilled spirits, and cola); the largest value was 0.603 $\mu g/g$ (instant coffee). These extreme values might affect the estimation in the model described above. We removed outliers located higher than the 1.5-fold interquartile range above the third quartile from each food group. A subject’s daily biotin intake (T) was calculated using the same formula as that for model 2-1.

**Results**

The probability density distribution and summary of estimated daily biotin intake data are presented in Fig. 1. The mean, 5th percentile, and 95th percentile of daily biotin intake estimated from biotin concentration data from TDS samples were, respectively, 67.8, 40.1, and 108.2 $\mu g/d$ (Fig. 1, model 1). The values estimated from a dataset of individual foods were, respectively, 189.5, 62.2, and 368.7 $\mu g/d$ (Fig. 1, model 2-1). The mean value of model 1-1 was 2.8 times higher than that of model 1. After removing outliers, the probability density distributions from the two models were analogous in mean, median, 5th percentile, and 95th percentile estimated between model 1 and model 2-2.

Although the probability density distributions were similar between model 1 and model 2-2, the contribution of each food group to the total daily biotin intake was somewhat different (Fig. 2). For example, in model 1, the contribution was highest in the meat/egg group, followed by the fish/shellfish and seasoning groups. These three groups accounted for more than half of the daily biotin intake. In contrast, in model 2, the contribution was highest in rice/processed rice, followed by seasoning and cereal/potato groups. These three groups accounted for more than half of the daily biotin intake.

### Table 3. Biotin concentrations of respective food groups evaluated from individual foods.

| No. | Food group               | All data | Without outliers² |
|-----|--------------------------|----------|-------------------|
|     |                          | Mean     | Min.   | Max. (n) | Mean     | Min.   | Max. (n) |
| 1   | Rice/processed rice      | 0.116    | 0.004  | 0.600 (18) | 0.038    | 0.004  | 0.120 (15) |
| 2   | Cereal/potato            | 0.115    | 0.001  | 0.911 (121) | 0.056    | 0.001  | 0.240 (104) |
| 3   | Sugar/confectionery      | 0.034    | 0.000  | 0.320 (30)  | 0.017    | 0.000  | 0.053 (27)  |
| 4   | Fat/oil                 | 0.006    | 0.000  | 0.034 (8)   | 0.002    | 0.000  | 0.007 (7)   |
| 5   | Pulse                   | 0.178    | 0.005  | 1.140 (38)  | 0.094    | 0.005  | 0.340 (33)  |
| 6   | Fruit                   | 0.016    | 0.001  | 0.100 (109) | 0.012    | 0.001  | 0.040 (100) |
| 7   | Green, yellow vegetables | 0.040    | 0.002  | 0.840 (70)  | 0.016    | 0.002  | 0.053 (61)  |
| 8   | Other vegetables         | 0.039    | 0.000  | 0.414 (118) | 0.022    | 0.000  | 0.083 (106) |
| 9   | Beverage                | 0.063    | 0.001  | 0.603 (26)  | 0.005    | 0.001  | 0.014 (22)  |
| 10  | Fish/shellfish           | 0.062    | 0.002  | 0.410 (110) | 0.052    | 0.002  | 0.163 (104) |
| 11  | Meat/egg                | 0.197    | 0.000  | 2.274 (131) | 0.069    | 0.000  | 0.330 (108) |
| 12  | Milk/dairy products      | 0.034    | 0.002  | 0.430 (123) | 0.024    | 0.002  | 0.062 (114) |
| 13  | Seasoning               | 0.312    | 0.003  | 2.000 (28)  | 0.126    | 0.003  | 0.600 (24)  |

Total n = 930

|                  | 825 |
|------------------|-----|
| ¹Numbers in parentheses show numbers of food samples in respective groups. |
| ²Outliers are located higher than 1.5 fold interquartile range above the third quartile. |
Discussion

The importance of adequately characterizing the probability distribution has been emphasized for assessing the nutrient adequacy of a population. For this study, we applied Monte Carlo simulation models to estimate the probabilistic distribution of the daily biotin intake. Two models were used. One was evaluated using biotin concentration data from TDS 2007–2009 in Osaka (model 1). The other includes biotin concentration data from individual foods (model 2-1, 2-2).

In the TDS, each group consisted of 4–27 food samples (Table 2). Each sample was classified under one group to produce homogenized mixtures according to the proportions determined by NHNS. In addition, food samples were cooked. Therefore, the processing factor was included in the TDSs. At this time, model 1 is apparently the most appropriate to estimate the probabilistic distribution of daily biotin intake. To improve model 1 we need to collect more diversified TDS data associated with sample compositional variability (e.g. seasonal or geographical variation). However, collecting many TDS data is time-consuming and cost-intensive.

Therefore, we applied model 2 to estimate the probabilistic distribution from the available dataset of biotin concentration in individual foods in the place of TDS data. The daily biotin intake estimated using model 2-1 was 2.8 times higher than that estimated using TDS samples. In the dataset, each group contained outliers mainly consisting of dry or raw foods that were inedible without cooking. When these outliers were removed, the probability density distribution was analogous to that obtained using model 1. However, the food-group contribution to daily biotin intake was different between the two models (Fig. 2). In model 2, it was assumed that a person randomly chose three individual foods from each food group in a day and consumed them in equal amounts. Therefore, the evaluated

![Fig. 1. Probability density distribution of daily biotin intake estimated using Monte Carlo simulation with two models. Model 1: Model estimated from biotin concentration data by TDS 2007–2009 in Osaka. Model 2: Model estimated from biotin concentration data from individual foods. 2-1: Estimated with all biotin concentration data in Table 3. 2-2: Estimated with biotin concentration data without outliers in Table 3.](image1)

![Fig. 2. Contribution of each food group to total daily biotin intake with Monte Carlo simulation using two models. Models are identical to those presented for Fig. 1.](image2)
amount of intake of each food sample did not necessarily reflect their ingestion rates accurately. In another study conducted in Japan, Watanabe et al. estimated the daily biotin intake from the biotin concentration dataset for individual foods and NHNS. They estimated the daily biotin intake as 107.8 $\mu$g/d for males and 91.6 $\mu$g/d for females from 101 biotin-concentration data classified into 18 food groups (25). In a subsequent study, they also estimated the daily biotin intake of 54.5 $\mu$g/d from 330 biotin-concentration data classified into 98 food groups (10). The difference between these two earlier studies suggests that not only the number of datasets but also food group classifications are important factors to estimate the daily biotin intake. When we subdivided 13 food groups into 18 food groups in model 2, the reestimated value was 188.0 ±96.4 (mean ±SD), similar to the original value (189.5 ±96.2). To improve model 2, we need to collect more biotin-concentration data and divide foods into small groups; however, at this time there is insufficient biotin concentration data to estimate with.

Even though probabilistic models in this study present the limitations described above, from the probabilistic distributions of daily biotin intake, we were able to examine the full range of intakes, select any level of intake, read the probability of its occurrence, and determine the proportion of a population below or above any selected level of intake. The daily biotin intakes from these probabilistic methods showed that more than 80% of the population was above the adequate intake of biotin in Japan.

### Appendix Program for Monte Carlo simulation

#### Monte Carlo simulation model 1

The following R program calculates the daily biotin intake from 10,000 iterations using randomly sampled values from $C_{TDS}$ ($i$) and $I(i)$ in each food group.

```r
# Input data mc07–mc09=biotin concentration data evaluated by TDS($C_{mc}$), mw=mean of food consumption data based on data from NHNS 2005, sdw=standard deviations of food consumption data based on data from NHNS 2005.

mw<-c(351.1,163.4,30.1,11.1,60.7,115.8,96.3,198.8, 686.4,88.3,121.5,97.0,103.0)
sdw<-c(194.8,134.9,48.4,9.8,75.6,154.2,84.7,129.1, 524.2,81.6,83.6,136.2,95.3)
mc07<-c(0.014,0.021,0.040,0.000,0.072,0.013,0.030, 0.016,0.006,0.072,0.238,0.029,0.067)
mc08<-c(0.003,0.021,0.034,0.000,0.070,0.013,0.047, 0.015,0.005,0.111,0.198,0.023,0.034)
mc09<-c(0.007,0.046,0.048,0.000,0.060,0.016,0.028, 0.025,0.004,0.103,0.215,0.062,0.100)

# Generate lognormal distribution

rm2<-function(n, lm, lsd, min, max){
  if ((lm == 0) res <- rep(0, n)
  else{
    res<-NULL
    v <- log((lsd/lm)^2+1)
    m <- log(lm)-v/2
    s <- sqrt(v)
    for (i in 1:n){
      x <- exp(rnorm(1, mean=m, sd=s))
      if(x<min || x > max)[x <- rlm2(1, lm, lsd, min, max)]
      res<-c(res,x)
    }
    res
  }
}

# Sample random values from biotin concentration data estimated using TDS 2007–2009.

getmc<-function(i){
  rnum<-runif(1)
  if(rnum<1/3) x <- mc07[i]
  else if(rnum<2/3) x <- mc08[i]
  else x <- mc09[i]
  x
}

# Calculate the daily biotin intake for 10,000 iterations

model1 <- function(repl){
  res <- NULL
  for (i in 1:repl){
    x <- 0
    for (j in 1:13){
      x <- x + getmc(j) * rlm2(1, mw[j], sdw[j], 0, 5*sdw[j])
    }
    res <- c(res, x)
  }
  res
}

data1 <- model1(10000)

#### Monte Carlo simulation model 2

The following R program calculates the daily biotin intake from 10,000 iterations of randomly sampled three values $C_{INF}$ and $I(i)$ in each food group.

```r
# Input data g1–g13=biotin concentration data in individual food, mw=mean of food consumption data based on values from NHNS 2005, sdw=standard deviation of food consumption data based on values from NHNS 2005.

g1 <-c(CINF1, CINF2, ·)
g2 <-c(CINF5, ·)
g3 <-c(CINF9, CINF10)
g <- list(g1, g2, g3, g4, g5, g6, g7, g8, g9, g10, g11, g12, g13)
mw<-c(351.1,163.4,30.1,11.1,60.7,115.8,96.3,198.8, 686.4,88.3,121.5,97.0,103.0)
sdw<-c(194.8,134.9,48.4,9.8,75.6,154.2,84.7,129.1, 524.2,81.6,83.6,136.2,95.3)
```
model2 <- function(repl, num) {
  res <- NULL
  for (i in 1:repl) {
    x <- 0
    for (j in 1:13) {
      tmpwt <- rlm2(1, mw[j], sdw[j], 0, 5*mw[j])
      while (tmpwt > mw[j]/num) {
        tmpmc <- sample(g[[j]], 1)
        x <- x + tmpmc * mw[j]/num
        tmpwt <- tmpwt - mw[j]/num
      }
      tmpmc <- sample(g[[j]], 1)
      x <- x + tmpmc * tmpwt
    }
    res <- c(res, x)
  }
  res
}
data2 <- model2(10000, 3)

REFERENCES
1) Zempleni J, Mock DM. 1999. Biotin biochemistry and human requirements. J Nutr Biochem 10: 128–138.
2) Mock NI, Malik MI, Stumbo PJ, Bishop WP, Mock DM. 1997. Increased urinary excretion of 3-hydroxyisovaleric acid and decreased urinary excretion of biotin are sensitive early indicators of decreased biotin status in experimental biotin deficiency. Am J Clin Nutr 65: 951–958.
3) Pacheco-Alvarez D, Solorzano-Vargas RS, Rio ALD. 2002. Biotin in metabolism and its relationship to human disease. Arch Med Res 33: 439–447.
4) Food and Nutrition Board, Institute of Medicine. 1998. Dietary Reference Intakes for Thiamin, Riboflavin, Niacin Vitamin B6, Folate, Vitamin B12, Pantothenic Acid, Biotin, and Choline. p 377–378. National Research Council, National Academy Press, Washington DC.
5) Tsuda T, Inoue T, Kojiima M, Aoki S. 1995. Market basket and duplicate portion estimation of dietary intakes of cadmium, mercury, arsenic, copper, manganese, and zinc by Japanese adults. J AOAC Int 78: 1363–1368.
6) The Ministry of Health, Labour and Welfare of Japan. 2004. Reports on the Results of Dioxins from Foods in Japan (FY2002) (in Japanese).
7) The Ministry of Health, Labour and Welfare of Japan. 2008. Reports on the Results of Pesticides from Foods in Japan (FY2004) (in Japanese).
8) Saitoh Y, UsBio F. 2004. Estimate of the daily dietary intake of biotin, vitamin B6 and niacin from the 1999 total diet study. Eiyougaku zassi (Jpn J Nutr Diet) 62: 165–169 (in Japanese).
9) Murakami T, Yamano T, Nakama A, Mori Y. 2008. Estimation of the dietary intake of biotin and its measurement uncertainty using total diet samples in Osaka, Japan. J AOAC Int 91: 1402–1408.
10) Watanabe T, Taniguchi A. 2009. Estimation of dietary intake of biotin from the Japanese diet. Vitamins 3: 461–468 (in Japanese).
11) The Ministry of Health, Labour and Welfare of Japan. 2009. Dietary Reference Intakes for Japanese, 2010 (in Japanese).
12) Gibney MJ, van der Voet H. 2003. Introduction to the Monte Carlo project and the approach to the validation of probabilistic models of dietary exposure to selected food chemicals. Food Additives Contam 20 (Suppl 1): S1–S7.
13) Mato Y, Suzuki N, Katatani N, Kadokami K, Nakano T, Nakayama S, Sekii H, Komoto S, Miyake S, Morita M. 2007. Human intake of PCDDs, PCDFs, and dioxin like PCBs in Japan, 2001 and 2002. Chemosphere 67: S247–255.
14) Ministry of Health, Labour and Welfare of Japan. 2005. National Health and Nutrition Survey Japan 2005 (in Japanese).
15) Taniguchi A, Takeshi R, Fukushima A, Watanabe T. 2008. Biotin content of typical foods in Japan. Nippon Eigo Shokuryo Gakkaishi (J Jpn Nutr Food Sci) 61: 27–37 (in Japanese).
16) Technical University of Denmark. 2009. The official Danish food composition database Version 7.01. Department of Nutrition, National Food Institute, Technical University of Denmark. [Online]. Available: http://www.foodcomp.dk/v7 [accessed March 31, 2010].
17) Scherz H, Senser F. 2000. Food Composition and Nutrition Tables, 6th ed. CRC Press, Stuttgart.
18) Hoppenr K, Smith DC. 1978. An appraisal of the daily intakes of vitamin B12, pantothenic acid and biotin from composite Canadian diet. Can Inst Food Sci Technol 11: 71–74.
19) Hoppenr K, Lampi B. 1989. Total folate, pantothenic acid and biotin in some fish products. Can Inst Food Sci Technol 22: 170–172.
20) Hoppenr K, Lampi B. 1992. Biotin content of cheese products. Food Res Inst 25: 41–43.
21) Hoppenr K, Lampi B, O’Grady E. 1994. Biotin content in vegetables and nuts available on the Canadian market. Can Inst Food Sci Technol 27: 495–497.
22) Hardinge MG, Crooks H. 1961. Lesser known vitamins in foods. J Am Diet Assoc 38: 240–245.
23) Ihaka R, Gentleman R. 1996. A language for data analysis and graphics. J Comput Graph Stat 5: 299–314.
24) WU Vienna University of Economics and Business. 2010. The R project for statistical computing. The comprehensive R archive network. [Online]. Available: http://www.r-project.org/ [accessed March 31, 2010].
25) Watanabe T, Taniguchi A. 2006. Study on the estimate of daily intake of biotin by total diet study. Nippon Rinsho Eiyo Gakkaiishi (Journal of Japanese Society of Clinical Nutrition) 27: 304–312 (in Japanese).