Reproducibility of Self-Administered Questionnaire in Epidemiological Surveys

Nakako Fujiwara and Shinkan Tokudome

We evaluated the reproducibility of data on intake frequency of 33 food items, beverage intake frequency of 3 items, drinking and smoking habits, and past history of 10 diseases obtained from a self-administered questionnaire. The survey subjects consisted of 263 aged 39 to 79 years individuals in the general population. After about one year interval these subjects consecutively participated in two surveys and were unintentionally requested to answer to the same questionnaire. The means of percent exact agreements for 33 food items were 48.4% in all, 42.7% in males and 49.4% in females. The means of their percent agreements within one category difference were 85.4% in all, 83.3% in males and 85.8% in females. The reproducibility of beverage intake frequency was the highest in green tea, followed in order by black tea and coffee. The percent exact agreements were 85.4% in males and 81.8% in females for drinking habit, 87.5% in males and 99.0% in females for smoking habit, and 93.7% in males and 97.2% in females for past history. The reproducibility of the self-administered questionnaire was the highest in past history, followed in order by smoking habit, drinking habit and dietary habit. The values of reproducibility were higher in females than in males except for drinking habit. Although reduction of categories was needed to improve the reproducibility, the obtained values from the self-administered questionnaire were sufficiently high for epidemiological studies.

J Epidemiol, 1997; 7:61-69.

self-administered questionnaire, reproducibility, epidemiological survey, percent agreement, kappa statistic

From the epidemiological standpoint, various lifestyle surveys have been carried out to clarify the association with diseases. Evaluation of the precision and validity of obtained data is essential; however, such studies have not fully been performed due to limitations of time and cost. Reproducibility, or repeatability is essentially the degree of agreement between repeated measurements. Validity is the range to which a method of measurement provides a true assessment. There are many studies on the reproducibility and/or validity of data in epidemiological surveys; whereas a little consideration has been given to the reproducibility of the epidemiological data obtained from the cohort study. The aim of this paper is to study the reproducibility of the self-administered questionnaire for the cohort study which we are carrying out. In this paper, we especially evaluated the reproducibility of the epidemiological data by sex, assuming that there is no change in life style within one year.

MATERIALS AND METHODS

Study population

The subjects of whole survey consisted of 5,279 individuals in the general population cohort from 3 cities and 7 towns in Saga Prefecture in Japan. Of them, 263 individuals (41 males and 222 females) from one city and three towns, who accidentally participated in the same epidemiological survey after about one year interval, were analyzed.

The age of males ranged from 40 to 79 years of age (mean ± standard deviation, 54.8 ± 9.8), and that of females 39 to 73 years of age (53.6 ± 6.8).
Methods of survey

Epidemiological survey, which includes questionnaire study, was carried out twice according to the same format and procedure in 1989 and 1990.

A questionnaire consists of self-administered questions on intake frequency of 33 food items, beverage intake frequency of 3 items, drinking and smoking habits, and past history of 10 diseases.

The choice for each question on food and beverage intake frequency included five categories: consumes (1) rarely, (2) 1 - 2 times per month, (3) 1 - 2 times per week, (4) 3 - 4 times per week, and (5) nearly daily. The choice on habits of drinking and smoking included three categories: (1) drinks/smokes, (2) stopped drinking/smoking, (3) does not drink/never smoke. For the past history of diseases the choice included two categories: (1) no and (2) yes.

Statistical analyses

For assessing the reproducibility of categorical items, the following parameters were used: (1) percent agreement (exact and one category difference), (2) kappa statistic, (3) weighted kappa statistic and (4) Spearman's rank correlation coefficient.

The percent exact agreement is the percentage of the subjects who chose the identical category in the first and second surveys to all subjects who replied to the questionnaire in both surveys (number of effective repliers) 18. The percent agree-

Table 1. Reproducibility of data from a food intake questionnaire in males.

| Food item                      | Effective answers (pairs) | With exact category | Within limits of one category | Kappa   | Weighted kappa | 95% Confidence interval of weighted kappa | Spearman's rank correlation coefficient |
|--------------------------------|---------------------------|--------------------|--------------------------------|---------|----------------|------------------------------------------|----------------------------------------|
| Beef                           | 41                        | 65.9               | 100.0                          | 0.381*** | 0.639**        | 0.486~0.793                             | 0.638***                               |
| Pork                           | 41                        | 58.5               | 90.2                           | 0.331*** | 0.420          | 0.040~0.800                             | 0.415**                                |
| Ham                            | 40                        | 27.5               | 77.5                           | 0.009   | 0.292          | 0.027~0.557                             | 0.223                                  |
| Chicken                        | 40                        | 32.5               | 95.0                           | -0.017  | 0.455*         | 0.278~0.631                             | 0.416**                                |
| Liver                          | 38                        | 55.3               | 97.4                           | 0.292*  | 0.533*         | 0.321~0.745                             | 0.557***                               |
| Egg                            | 41                        | 46.3               | 92.7                           | 0.216*  | 0.557**        | 0.343~0.730                             | 0.588***                               |
| Milk                           | 41                        | 56.1               | 82.9                           | 0.380***| 0.685***       | 0.486~0.889                             | 0.638***                               |
| Yogurt                         | 40                        | 57.5               | 85.0                           | 0.250** | 0.625**        | 0.363~0.887                             | 0.630***                               |
| Cheese                         | 38                        | 73.7               | 89.5                           | 0.487***| 0.597*         | 0.282~0.912                             | 0.618***                               |
| Butter                         | 39                        | 46.2               | 84.6                           | 0.033** | -0.034         | -0.705~0.636                            | 0.069                                  |
| Margarine                      | 40                        | 42.5               | 77.5                           | 0.167   | 0.399          | 0.139~0.659                             | 0.284                                  |
| Fried food                     | 39                        | 38.5               | 84.6                           | 0.017   | -0.085         | -0.586~0.415                            | 0.059                                  |
| Fried vegetables               | 38                        | 39.5               | 81.6                           | 0.037   | -0.143         | -0.819~0.533                            | 0.102                                  |
| Fish                           | 41                        | 51.2               | 82.9                           | 0.278** | 0.226          | -0.162~0.614                            | 0.269                                  |
| Boiled fish paste              | 41                        | 39.0               | 75.6                           | 0.141   | 0.134          | -0.310~0.577                            | 0.173                                  |
| Fried fish                     | 41                        | 43.9               | 90.2                           | 0.178   | 0.133          | -0.454~0.719                            | 0.197                                  |
| Green leaf vegetables (eg. spinach) | 40                | 37.5               | 77.5                           | 0.124   | 0.068          | -0.425~0.560                            | 0.120                                  |
| Carrot, Pumpkin                | 41                        | 37.1               | 95.1                           | -0.037  | 0.458*         | 0.287~0.629                             | 0.514***                               |
| Tomato                         | 41                        | 39.0               | 80.5                           | 0.222** | 0.450*         | 0.147~0.753                             | 0.430*                                 |
| Cabbage, Lettuce               | 41                        | 43.9               | 80.5                           | 0.207*  | 0.231          | -0.210~0.672                            | 0.297                                  |
| Chinese cabbage                | 41                        | 29.3               | 78.0                           | 0.097   | 0.488*         | 0.295~0.682                             | 0.518***                               |
| Edible wild plants            | 39                        | 64.1               | 92.3                           | 0.337** | 0.371          | -0.206~0.947                            | 0.440***                               |
| Mushroom                       | 40                        | 35.0               | 87.5                           | 0.125   | 0.489*         | 0.257~0.720                             | 0.472**                                |
| Potato                         | 40                        | 40.0               | 90.0                           | 0.114   | 0.342          | 0.017~0.667                             | 0.271                                  |
| Seaweed                        | 40                        | 32.5               | 82.5                           | 0.065   | 0.282          | -0.017~0.581                            | 0.323**                                |
| Pickles                        | 41                        | 43.9               | 78.0                           | 0.227*  | 0.519*         | 0.310~0.729                             | 0.514***                               |
| Food boiled down in soy        | 41                        | 31.7               | 75.6                           | 0.109   | 0.449*         | 0.241~0.656                             | 0.444**                                |
| Boiled beans                   | 39                        | 43.6               | 82.1                           | 0.226*  | 0.370          | 0.042~0.697                             | 0.422**                                |
| Tofu                           | 41                        | 46.3               | 85.4                           | 0.229*  | 0.426          | 0.144~0.708                             | 0.456**                                |
| Orange                         | 40                        | 22.5               | 67.5                           | -0.023  | 0.175          | -0.120~0.470                            | 0.181                                  |
| Juice                          | 38                        | 26.3               | 57.9                           | 0.059   | 0.088          | -0.268~0.443                            | 0.109                                  |
| Fruit                          | 38                        | 36.8               | 76.3                           | 0.179*  | 0.232          | -0.177~0.642                            | 0.271                                  |
| Confectionery                  | 40                        | 32.5               | 75.0                           | 0.140   | 0.528**        | 0.369~0.687                             | 0.537***                               |

Mean of 33 items                |                           | 42.7               | 83.3                           |         |                |                                         |                                        |
Standard deviation              |                           | 11.9               | 8.6                            |         |                |                                         |                                        |

*p<0.05, **p<0.01, ***p<0.001

#1: Answers to the items in both surveys.
ment within one category difference is the percentage of total subjects who chose the same or one category difference in both surveys.

The kappa statistic is a parameter which measures the degree surpassing the agreement expected by coincidence alone \(^{13,16}\).

In addition, the weighted kappa statistic was calculated using the square of the distance between categories as weight \(^{17}\). To evaluate the values of kappa and weighted kappa statistics, we adopted the same criteria recommended by Landis and Koch \(^{10}\), which are as follows: kappa values greater than 0.75 or so may be taken to represent "excellent" agreement, values between 0.40 and 0.75 may be taken to represent "fair to good" agreement, values below 0.40 or so may be taken to represent "poor" agreement. These kappa values were statistically tested by comparing the values \(z = \frac{\text{kappa value}}{\text{its standard error}}\) with the limits of the normal distribution \(^{14}\).

Spearman's rank correlation coefficient was calculated after processing the same rank to correct for the number of observations that are tied for the same category \(^{18}\).

**RESULTS**

The means of exact percent agreements for the 33 food items were 48.4% in all, 42.7% in males (Table 1) and 49.4% in females (Table 2). The number of food items showing a percent exact agreement of 50% or more were 13 (39.4%) in all,

| Food item               | Effective answers (pairs) | With exact category | Within limits of one category | Kappa  | Weighted kappa | 95% Confidence interval of weighted kappa | Spearman's rank correlation coefficient |
|-------------------------|---------------------------|---------------------|-------------------------------|--------|----------------|----------------------------------------|----------------------------------------|
| Beef                    | 220                       | 57.3               | 39.8                          | 0.334*** | 0.458***       | 0.237~0.589                           | 0.471***                                |
| Pork                    | 208                       | 47.1               | 91.3                          | 0.217*** | 0.481***       | 0.359~0.592                           | 0.451***                                |
| Ham                     | 200                       | 52.0               | 84.5                          | 0.337*** | 0.439***       | 0.309~0.583                           | 0.477***                                |
| Chicken                 | 208                       | 50.0               | 93.8                          | 0.223*** | 0.498***       | 0.401~0.596                           | 0.461***                                |
| Liver                   | 209                       | 58.4               | 92.8                          | 0.381*** | 0.557***       | 0.425~0.689                           | 0.624***                                |
| Egg                     | 220                       | 48.6               | 97.3                          | 0.247*** | 0.464***       | 0.335~0.594                           | 0.456***                                |
| Milk                    | 211                       | 66.5               | 87.3                          | 0.444*** | 0.710***       | 0.614~0.805                           | 0.714***                                |
| Yogurt                  | 209                       | 49.8               | 85.2                          | 0.292*** | 0.525***       | 0.389~0.662                           | 0.548***                                |
| Cheese                  | 212                       | 62.7               | 90.6                          | 0.392*** | 0.484***       | 0.334~0.635                           | 0.484***                                |
| Butter                  | 211                       | 52.1               | 82.5                          | 0.277*** | 0.416***       | 0.373~0.559                           | 0.405***                                |
| Margarine               | 210                       | 44.3               | 82.9                          | 0.284*** | 0.616***       | 0.254~0.708                           | 0.592***                                |
| Fried food              | 214                       | 50.9               | 93.9                          | 0.260*** | 0.454***       | 0.329~0.578                           | 0.464***                                |
| Fried vegetables        | 218                       | 46.8               | 91.3                          | 0.212*** | 0.393***       | 0.257~0.530                           | 0.390***                                |
| Fish                    | 219                       | 60.3               | 89.5                          | 0.394*** | 0.495***       | 0.364~0.622                           | 0.482***                                |
| Boiled fish paste       | 213                       | 44.6               | 85.9                          | 0.277*** | 0.413***       | 0.286~0.540                           | 0.390***                                |
| Dried fish fish         | 213                       | 44.1               | 83.1                          | 0.234*** | 0.410***       | 0.258~0.533                           | 0.429***                                |
| Green leaf vegetables   | 220                       | 49.5               | 85.0                          | 0.250*** | 0.370***       | 0.231~0.509                           | 0.374***                                |
| Carrot, Pumpkin         | 221                       | 46.2               | 88.7                          | 0.208*** | 0.315***       | 0.153~0.476                           | 0.355***                                |
| Tomato                  | 219                       | 51.6               | 87.7                          | 0.275*** | 0.428***       | 0.273~0.583                           | 0.383***                                |
| Cabbage, Lettuce        | 221                       | 51.1               | 87.3                          | 0.260*** | 0.398***       | 0.275~0.520                           | 0.374***                                |
| Chinese cabbage         | 203                       | 36.5               | 70.9                          | 0.180*** | 0.283***       | 0.141~0.424                           | 0.288***                                |
| Edible wild plants      | 210                       | 54.3               | 88.6                          | 0.209*** | 0.224           | 0.010~0.459                           | 0.315***                                |
| Mushroom                | 216                       | 39.4               | 86.1                          | 0.155*** | 0.383***       | 0.255~0.511                           | 0.353***                                |
| Potato                  | 216                       | 51.9               | 85.6                          | 0.291*** | 0.275***       | 0.095~0.455                           | 0.294***                                |
| Seaweed                 | 215                       | 56.7               | 88.4                          | 0.339*** | 0.456***       | 0.316~0.596                           | 0.448***                                |
| Pickles                 | 217                       | 53.0               | 82.5                          | 0.284*** | 0.546***       | 0.428~0.665                           | 0.535***                                |
| Food boiled down in soy | 208                       | 42.8               | 77.4                          | 0.213*** | 0.388***       | 0.258~0.518                           | 0.410***                                |
| Boiled beans            | 213                       | 48.8               | 85.0                          | 0.220*** | 0.309***       | 0.133~0.486                           | 0.355***                                |
| Tofu                    | 222                       | 52.7               | 91.9                          | 0.274*** | 0.427***       | 0.299~0.556                           | 0.442***                                |
| Orange                  | 217                       | 39.2               | 77.0                          | 0.210*** | 0.382***       | 0.256~0.508                           | 0.389***                                |
| Juice                   | 208                       | 27.4               | 64.9                          | 0.086*  | 0.216*          | 0.071~0.361                           | 0.217*                                  |
| Fruit                   | 202                       | 51.0               | 87.1                          | 0.284*** | 0.471***       | 0.333~0.609                           | 0.514***                                |
| Confectionery           | 222                       | 42.8               | 85.1                          | 0.242*** | 0.522***       | 0.416~0.627                           | 0.526***                                |

Mean of 33 items 49.4 85.8
Standard deviation 7.6 6.1

\(* p<0.05, \quad ** p<0.01, \quad *** p<0.001

\#1: Answers to the items in both surveys.
Table 3. Reproducibility of beverages.

| Food item | Effective answers (pairs) | Percent agreement(%) | 95% Confidence interval of weighted kappa | Spearman's rank correlation coefficient |
|-----------|---------------------------|----------------------|------------------------------------------|----------------------------------------|
|           | Males                     |                      |                                          |                                        |
| Coffee    | 41                        | 39.0                 | 63.4                                     | 0.098 ~ 0.666                         | 0.0318***                                |
| Black tea | 39                        | 76.9                 | 97.4                                     | 0.421 ~ 1.012                         | 0.345**                                  |
| Green tea | 41                        | 80.5                 | 85.4                                     | -0.221 ~ 0.849                        | 0.669***                                 |
|           | Females                   |                      |                                          |                                        |
| Coffee    | 221                       | 57.0                 | 86.0                                     | 0.691 ~ 0.825                         | 0.739***                                 |
| Black tea | 222                       | 76.1                 | 91.9                                     | 0.250 ~ 0.693                         | 0.498***                                 |
| Green tea | 216                       | 85.2                 | 90.7                                     | 0.355 ~ 0.743                         | 0.594***                                 |

* p<0.05, ** p<0.01, *** p<0.001

#1: Answers to the items in both surveys.

8 (24.2%) in males and 17 (51.5%) in females. Percent exact agreements for 28 out of 33 food items showed higher in females than in males. The means of percent agreements within one category difference were 85.4% in all, 83.3% in males and 85.8% in females.

The food items giving a percent exact agreement of 30% or less were ham, Chinese cabbage, orange, and juice in males but only juice in females. For those food items that showed 30% or less of percent exact agreement, the reproducibility of these food items was improved from 57.9% to 78.0%, when used the percent agreement within one category difference.

The calculated values of kappa and weighted kappa according to the food item are also summarized in Tables 1 and 2. Beef, pork, chicken, liver, egg, milk, yogurt, cheese, tomato, pickles, tofu, and confectionery showed weighted kappa values exceeding 0.4 in males and females, which suggested fair to good agreement. In contrast, the weighted kappa values were less than 0.4 for fried vegetables, green leaf vegetables, cabbage or lettuce, edible wild plants, potato, boiled beans, orange, and juice in both sexes, suggesting poor agreement. For the reproducibility assessments using the Spearman’s rank correlation coefficient, almost the similar evaluations when using the percent agreements or kappa statistic, were also shown in Tables 1 and 2.

The reproducibility of three beverage intake frequency was the highest in green tea (percent exact agreements were 84.4% in all, 80.5% in males and 85.2% in females), followed in order by black tea and coffee (Table 3).

Because gross misclassification is very serious issue in epidemiological surveys, we examined the opposite reply between the two surveys. Table 4 shows the number of the subjects who chose an extreme category [(1) eat rarely or (5) nearly daily] on 33 food items and 3 beverages in the first survey, along with the number of the subjects who selected the exact opposite category in the second survey. The means for these 36 items revealed that about 40% among effective answers in both sexes chose category 1 or 5 in the first survey and very few among effective answers, 0.82% in all, 1.11% for males and 0.77% for females, chose the exact opposite category in the second survey. The items for which the rate of the exact opposite answer exceeded 3% among effective answers, were juice in females, coffee in males, and green tea in both sexes.

Figure 1 shows the relationship between percent exact agreement (y-axis) and estimated mean food intake per week (x-axis) for each items. The percent exact agreements for the food consumed rarely or frequently were higher in both sexes than for the moderate food intake. The r-square values for the secondary regression equation having food intake as an independent variable and percent exact agreement as a dependent variable, were more than 0.5 in both sexes.

Table 5 shows the percent exact agreement and distribution of reply to drinking and smoking habits. The percent agreements for drinking habit were 82.4% in all, 85.4% in males and 81.8% in females. Although the percent exact agreement showed more than 80%, among 159 females who answered "does not drink" in the first survey 16 females (10.1%) answered "drinks" in the second survey.

The percent agreements for the smoking habit were 97.2% in all, 87.5% in males and 99.0% in females.

The percent exact agreements for past history of diseases were more than 90% except for gallbladder operation and appendectomy in males (Table 6).

DISCUSSION

Factors affecting reproducibility of the food frequency questionnaire should include survey season, survey interval, educational effect, food intake frequency, sex and so on.
Reproducibility of Questionnaire Measurements

Table 4. The ratio of subjects who selected the exact opposite category between two surveys.

| Food item                  | Males             | Females            |
|----------------------------|-------------------|--------------------|
|                            | Extreme answers   | Exact opposite     | Extreme answers   | Exact opposite     |
|                            | Effective answers | answers (pairs)    | answers (pairs)   | answers (pairs)    |
|                            | #1 1(F) + 5(S)    | n(%)               | a              | b              | n              | a  | b  | a+b |
| Beef                       | 41 2(4.88)        | 1(2.44)            | 0              | 0              | 0              | 0  | 0  | 0   |
| Pork                       | 41 7(17.07)       | 4(9.76)            | 0              | 0              | 0              | 0  | 0  | 0   |
| Ham                        | 40 16(40.00)      | 10(25.00)          | 0              | 0              | 0              | 0  | 0  | 0   |
| Chicken                    | 40 5(12.50)       | 4(10.00)           | 0              | 0              | 0              | 0  | 0  | 0   |
| Liver                      | 38 16(42.11)      | 14(36.84)          | 0              | 0              | 0              | 0  | 0  | 0   |
| Egg                        | 41 19(46.34)      | 16(39.02)          | 0              | 0              | 0              | 0  | 0  | 0   |
| Milk                       | 41 28(68.29)      | 25(60.98)          | 1              | 0              | 1(2.44)         | 0  | 0  | 0   |
| Yogurt                     | 40 25(62.50)      | 33(82.50)          | 1              | 0              | 1(2.50)         | 0  | 0  | 0   |
| Cheese                     | 38 27(71.05)      | 25(65.79)          | 0              | 0              | 0              | 0  | 0  | 0   |
| Butter                     | 39 26(66.67)      | 22(56.41)          | 0              | 1              | 1(2.56)         | 0  | 0  | 0   |
| Margarine                  | 40 23(57.50)      | 16(40.00)          | 0              | 0              | 0              | 0  | 0  | 0   |
| Fried food                 | 39 1(2.56)        | 1(2.56)            | 0              | 0              | 0              | 0  | 0  | 0   |
| Fried vegetables           | 38 5(13.16)       | 2(5.26)            | 0              | 1              | 1(2.63)         | 0  | 0  | 0   |
| Fish                       | 41 18(43.90)      | 14(34.15)          | 0              | 0              | 0              | 0  | 0  | 0   |
| Boiled fish paste          | 41 12(29.27)      | 9(21.95)           | 0              | 1              | 1(2.44)         | 0  | 0  | 0   |
| Dried fish                 | 41 5(12.20)       | 4(9.76)            | 0              | 1              | 1(2.44)         | 0  | 0  | 0   |
| Green leaf vegetables      | 40 12(30.00)      | 11(27.50)          | 0              | 1              | 1(2.50)         | 0  | 0  | 0   |
| Carrot, Pumpkin            | 41 5(12.20)       | 6(14.63)           | 0              | 0              | 0              | 0  | 0  | 0   |
| Tomato                     | 41 16(39.02)      | 10(24.39)          | 0              | 1              | 1(2.44)         | 0  | 0  | 0   |
| Cabbage, Lettuce           | 41 12(29.27)      | 11(26.83)          | 0              | 1              | 1(2.44)         | 0  | 0  | 0   |
| Chinese cabbage            | 41 18(43.90)      | 8(19.51)           | 0              | 0              | 0              | 0  | 0  | 0   |
| Edible wild plants        | 39 26(66.67)      | 24(61.54)          | 1              | 0              | 1(2.56)         | 0  | 0  | 0   |
| Mushroom                   | 40 6(15.00)       | 8(20.00)           | 0              | 0              | 0              | 0  | 0  | 0   |
| Potato                     | 40 3(7.50)        | 5(12.50)           | 0              | 0              | 0              | 0  | 0  | 0   |
| Seaweed                    | 40 13(32.50)      | 12(30.00)          | 0              | 0              | 0              | 0  | 0  | 0   |
| Pickles                    | 41 23(56.10)      | 15(36.59)          | 0              | 0              | 0              | 0  | 0  | 0   |
| Food boiled down in soy    | 41 18(43.90)      | 11(26.83)          | 0              | 0              | 0              | 0  | 0  | 0   |
| Boiled beans               | 39 15(38.46)      | 11(28.21)          | 0              | 0              | 0              | 0  | 0  | 0   |
| Tofu                       | 41 20(48.78)      | 13(31.71)          | 0              | 0              | 0              | 0  | 0  | 0   |
| Orange                     | 40 8(20.00)       | 10(25.00)          | 0              | 0              | 0              | 0  | 0  | 0   |
| Juice                      | 38 15(39.47)      | 11(28.95)          | 1              | 0              | 1(2.63)         | 0  | 0  | 0   |
| Fruit                      | 38 13(34.21)      | 4(10.53)           | 1              | 0              | 1(2.63)         | 0  | 0  | 0   |
| Confectionery              | 40 15(37.50)      | 8(20.00)           | 0              | 0              | 0              | 0  | 0  | 0   |
| Coffee                     | 41 20(48.78)      | 20(48.78)          | 2              | 0              | 2(4.88)         | 0  | 0  | 0   |
| Black tea                  | 39 34(87.18)      | 31(79.49)          | 0              | 2              | 2(4.88)         | 0  | 0  | 0   |
| Green tea                  | 41 39(95.12)      | 37(90.24)          | 2              | 0              | 2(4.88)         | 0  | 0  | 0   |

Mean of 36 items: (39.32) (32.38) (1.11) (38.74) (35.21) (0.77)

#1: Answers to the items in both surveys.
#2 1(F): The number of answers chosen category 1 (rarely) in the first survey.
#3 1(S): The number of answers chosen category 1 (rarely) in the second survey.
#3 5(S): The number of answers chosen category 5 (nearly daily) in the second survey.

As to the interval and time of surveys carried out, one must recall that, better reproducibility is attainable if the interval between the two surveys is short, whereas reproducibility tends to be poorer when the seasonal variation would cause a real change in life style, especially in the intake frequency of food items. We conducted surveys from July to September in 1989 and 1990. The mean intervals of the examination was approximately one year: 354 days (265 to 417 days) for males and 357 days (259 to 417 days) for females. Thus, we can ignore factors such as subject's previous reply recall or seasonal variation that might affect reproducibility. Also, since subjects were queried on their average intake frequency of usual amount over the past one year, their physical conditions on the response day could not affect reproducibility.
We evaluated the reproducibility in epidemiological survey using self-administered questionnaire in the two surveys, assuming that they didn't change their lifestyle within one year. Although we don't have any chances of special intervention or education, if any, to the participants in the first survey, to clarify this assumption, we compared the average food intake of each of the 33 food and 3 beverage items in the first and second surveys. The items in the two surveys found to be significantly different by the Mann-Whitney U test were as follows: the items more frequent in the second survey were Chinese cabbage (female, p<0.001), edible wild plants (females, p<0.05) and juice (females, p<0.01), whereas tofu (males, p<0.05) was the only item less frequent in the second survey. These food intake differences in the two surveys seemed not affected by education.

Ozasa found good reproducibility for food items taken habitually such as boiled rice, bread and milk with Spearman's correlation coefficients of higher than 0.6 \[10\]. High reproducibility of data on the intake frequency has been reported for food items that are frequently or rarely consumed \[4\]. In our data, the relationships between the frequency of food intake and percent agreement are shown in Figure 1. Food items with more or less frequent intake showed better agreements.

Food items showing a percent exact agreement of 50% or more were 8 among 33 food items in males and 17 in females. Food frequency reproducibility was generally lower in males than in females. This is understandable because females primarily cook. Although the number of males was small, the results obtained here may reflect the sex difference because the age distribution for males was almost the same as for females.

Fukao et al. studied the reproducibility of 35 items almost the same as ours. The reproducibility values were obtained from surveys in five-month intervals and covered five categories. The average exact agreements for the 35 items were 46.2% \(41.3-45.1\%\) for males, 46.2-49.3% for females \[5\]. In our study, with the average 36 items, including 33 foods and 3 beverages, the exact agreements were 50.3% \(44.6\%\) in males and 51.4% in females). Fukao et al. also pointed out items of low agreements for Chinese cabbage \(24.6\%\), tomato \(29.9\%\), and orange \(30.6\%\), citing the seasonal variation fac-
Reproducibility of Questionnaire Measurements

Table 5. Distribution of reply to drinking and smoking habits.

**Drinking habit**

|                  | Males Reply of 1990 |                |                |                | Percent exact agreement |
|------------------|----------------------|----------------|----------------|------------------|-------------------------|
|                  | Drinks | Drank | Does not drink | Total            |                         |
| Drinks #1        | 31     | 2     | 1              | 34               |                         |
| Drank #2         | 2      | 1     | 1              | 4                |                         |
| Does not drink #3| 0      | 0     | 3              | 3                |                         |
| Total            | 33     | 3     | 5              | 41               |                         |

**Females Reply of 1990**

|                  | Males Reply of 1990 |                |                |                | Percent exact agreement |
|------------------|----------------------|----------------|----------------|------------------|-------------------------|
|                  | Drinks | Drank | Does not drink | Total            |                         |
| Drinks           | 33     | 0     | 22             | 55               |                         |
| Drank            | 0      | 0     | 0              | 0                |                         |
| Does not drink   | 16     | 1     | 142            | 159              |                         |
| Total            | 49     | 1     | 164            | 214              |                         |

**Smoking habit**

**Males Reply of 1990**

|                  | Males Reply of 1990 |                |                |                | Percent exact agreement |
|------------------|----------------------|----------------|----------------|------------------|-------------------------|
|                  | Smokers | Ex-smokers | Never-smoked | Total            |                         |
| Smokers          | 17      | 0           | 1             | 18               |                         |
| Ex-smokers       | 1       | 4           | 1             | 6                |                         |
| Never-smoked     | 0       | 2           | 14            | 16               |                         |
| Total            | 18      | 6           | 16            | 40               |                         |

**Females Reply of 1990**

|                  | Males Reply of 1990 |                |                |                | Percent exact agreement |
|------------------|----------------------|----------------|----------------|------------------|-------------------------|
|                  | Smokers | Ex-smokers | Never-smoked | Total            |                         |
| Smokers          | 5       | 0           | 0             | 5                |                         |
| Ex-smokers       | 0       | 2           | 1             | 3                |                         |
| Never-smoked     | 0       | 1           | 198           | 199              |                         |
| Total            | 5       | 3           | 199           | 207              |                         |

*#1: He/she drinks habitually. #2: He/she drank habitually but quits. #3: He/she does not drink.*

In our data, in which we need not consider the seasonal variation factor, the values for Chinese cabbage were 29.3% in males and 36.5% in females; for tomatoes, 39.0% in males and 51.6% in females, and for oranges, 22.5% in males and 39.2% in females, showing higher agreements than their data.

Ozasa et al. reported good agreements for smoking (85% for males, 99% for females) and drinking (76% for males, 78% for females) \(^9\). Spitz et al. also showed good reproducibility values for smoking and drinking, which indicated more than 80% agreement \(^10\). In our data also, the percent exact agreements of drinking and smoking habits showed more than 80% in both sexes. However, the reproducibility of the drinking habits in females was the lowest: that is, 16 among 159 females (10.1%) who answered "does not drink" in the first survey oppositely responded "drinks" in the second survey; while 22 among 55 females (40.0%) who answered "drinks" in the first survey responded "does not drink" in the second survey. Although Colditz et al. reported that reproducibility was
slightly reduced among heavy drinkers\(^{22}\), our data may be explained by that females who occasionally drink only a little in a month or so would inconsistently answer “drinks” or “does not drink” according to the chance.

Fukao et al. reported that in the questionnaire consisting of five categories, 1.56% answered the exact opposite category\(^5\). In our data, 1.11% for males and 0.77% for females answered the exact opposite category (Table 4). These figures show that in the food frequency questionnaire consisting of five chosen categories, the possibility of obtaining the exact opposite answer was around 1%. For green tea, more than 4% in both sexes chose the exact opposite answers, whereas the percent exact agreements were good. It is important to note that Fukao’s data imply the same results\(^5\). Ozasa et al. reported that there was a large Spearman’s correlation coefficient and more than two category-difference answers in dairy food\(^10\). With items in connection with which the opposite answers were often given, careful consideration must be taken as to conclusions. However, one can scarcely misconstrue the conclusions obtained from the epidemiological data, where the population consists of large subjects.

Although the means of exact agreements of food intake, where the choices were five categories, were 42.7% in males and 49.4% in females, the percent agreements within one category difference were improved to 83.3% in males and 85.8% in females. These results suggest the necessity of reducing categories to three to improve the reproducibility of the data obtained. The less selected category showed more agreement, of course, but information was missing. Deciding the number of categories for analysis is extremely difficult. One needs to discuss how to round out the categories while considering the frequency distribution of each item. We suggest that the reproducibility of self-administered questionnaire used here was sufficiently acceptable for epidemiological cohort studies.

**ACKNOWLEDGMENTS**

This study was supported in part by a Grant-in-Aid for Scientific Research on Priority from the Ministry of Education, Science and Culture; (Comprehensive Cancer Research Promotion Area), as part of a collaborative study on Research on Evaluation of Risk Factors for Cancer by Large-scaled Cohort Study (Chairman: K. Aoki). The authors would like to thank public health nurses at Saga City (Ms. Y. Okabe, Ms. M. Tateishi et al.), Chiyoda Town (Ms. M. Sanematsu et al.), Fukudomi Town (Ms. M. Tanaka et al.), and Kawasoe Town (Ms. M. Ohtsubo et al.) for their help.

**REFERENCES**

1. Everitt BS. Observer bias, reproducibility and validity. In; Statistical Methods for Medical Investigations, 2nd ed. John Wiley & Sons, New York, 1994; 25-35.
2. Willett WC, Sampson L, Stampfer MJ, et al. Reproducibility and validity of a semiquantitative food frequency questionnaire. Am J Epidemiol, 1985; 122: 51-65.
3. McKeown-Eyssen GE, Yeung KS, Bright-See E. Assessment of past diet in epidemiologic studies. Am J Epidemiol, 1986; 124: 94-103.
4. Pietinen P, Hartman AM, Haapa E, et al. Reproducibility and validity of dietary assessment instruments. II. A qualitative food frequency questionnaire. Am J Epidemiol, 1988; 128: 667-676.
5. Fukao A, Shimizu H, Maezawa M, Hisamichi S. The reproducibility of answer to the questionnaire for food intake frequency in a survey. Nihon Kōshū Eisei Zasshi, 1990; 37: 347-352 (in Japanese).

6. Jacobsen BK, Bønaa KH. The reproducibility of dietary data from a self-administered questionnaire. The Tromsø study. Int J Epidemiol, 1990; 19: 349-353.

7. Block G, Woods M, Potosky A, Clifford C. Validation of a self-administered diet history questionnaire using multiple diet records. J Clin Epidemiol, 1990; 43:1327-1335.

8. Kubo N, Ohno Y. Reproducibility of life-style information obtained from a self-administered questionnaire. Jpn J Cancer Clin, 1991; 37: 217-223 (in Japanese).

9. Martin-Moreno JM, Boyle P, Gorgojo L, et al. Development and validation of a food frequency questionnaire in Spain. Int J Epidemiol, 1993; 22: 512-519.

10. Ozasa K, Watanabe Y, Higashi A, Liang H, et al. Reproducibility of a self-administered questionnaire for dietary habits, smoking, and drinking. Jpn J Hyg, 1994; 48: 1048-1057 (in Japanese).

11. Haraldsdóttir J, Tjønneland A, Overvad K. Validity of individual portion size estimates in a food frequency questionnaire. Int J Epidemiol, 1994; 23: 787-796.

12. Lee J. Alternate approaches for quantifying aggregate and individual agreements between two methods for assessing dietary intakes. Am J Clin Nutrition, 1980; 33: 956-958.

13. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics, 1977; 33: 159-174.

14. Fleiss JL. The measurement of interrater agreement. In; Statistical Methods for Rates and Proportions, 2nd ed. John Wiley & Sons, New York, 1981: 212-236.

15. Thompson WD, Walter SD. A reappraisal of the kappa coefficient. J Clin Epidemiol, 1988; 41: 949-958.

16. Kraemer HC, Bloch DA. Kappa coefficients in epidemiology: an appraisal of a reappraisal. J Clin Epidemiol, 1988; 41: 959-968.

17. Cohen J. Weighted kappa: nominal scale agreement with provision for scaled disagreement or partial credit. Psychol Bull, 1968; 70: 213-220.

18. Daniel WW. The Spearman rank correlation coefficient. In; Biostatistics: A Foundation for Analysis in the Health Sciences, 4th ed. John Wiley & Sons, New York, 1987: 621-629.

19. Järvinen R, Seppänen R, Knekt P. Short-term and long-term reproducibility of dietary history interview data. Int J Epidemiol, 1993; 22:520-527.

20. Tsubono Y, Nishino Y, Fukao A, Hisamichi S, Tsugane S. Temporal change in the reproducibility of a self-administered food frequency questionnaire. Am J Epidemiol, 1995; 142: 1231-1235.

21. Spitz MR, Fueger JJ, Newell GR. The development of a comprehensive, institution-based patient risk evaluation program: II. Validity and reliability of questionnaire data. Am J Prev Med, 1988; 4: 188-193.

22. Colditz GA, Willett WC, Stampfer MJ, et al. The influence of age, relative weight, smoking, and alcohol intake on the reproducibility of a dietary questionnaire. Int J Epidemiol, 1987; 16: 392-398.