A Cohort Study on the Relation of Lifestyle, Personality and Biologic Markers to Cancer in Miyagi, Japan: Study Design, Response Rate and Profiles of the Cohort Subjects

Akira Fukao¹, Yoshitaka Tsubono², Shoko Komatsu¹, Ichiro Tsuji¹, Yuko Minami¹, Shigeru Hisamichi¹, Toru Hosokawa³, Mieko Kawamura⁴, Akira Takano⁵, Nobuyuki Sugahara⁶, Takashi Ikeda⁶, and Mitsuaki Nishikori⁷

To assess the relation of lifestyle factors, personality and some biologic markers to cancer incidence, and to evaluate the effectiveness of cancer screening, a cohort study was conducted among all residents aged 40 to 64 years in 14 municipalities of Miyagi Prefecture, Japan. The baseline survey consisted of self-completed questionnaires concerning lifestyle and personality (Eysenck Personality Questionnaire-Revised), and serum samples collected from a part of the subjects. The proportion of the subjects who appropriately responded to the lifestyle questionnaire (response rate) was 91.7% (47,806 persons of the total 51,921 residents), indicating that representative data could be obtained. Of these, 41,442 persons (87.1%) also completed a personality questionnaire, and serum samples were collected from 8,896 persons (18.7%). Linkage of these baseline data with the cancer incidence data of a population-based cancer registry, and with the records of cancer screening at a cancer detection center will provide useful information for comprehensive cancer prevention. J Epidemiol, 1995; 5: 153-157.

In Japan, cancer has been the leading cause of death since the year 1980, and cancer control has been regarded as one of the most important and urgent public health policies. A large scale cohort study using the general population has widely been accepted as one of the most convincing methods to provide a basis for cancer prevention. A cohort study by Hirayama, which covered 270,000 Japanese, reported many meaningful results¹. However, as more than 30 years has passed since the baseline survey of that study, the results may have been outdated due to the recent changes in lifestyle of the Japanese people. Moreover, since the endpoint of that study was mortality, the possibility that the estimation of risks was influenced by the recent improvement of prognosis for cancer patients could not be ruled out.

We, therefore, carried out a cohort study to include the recent lifestyle of people in Miyagi Prefecture, where a population-based cancer registry had been established and the reliable data on cancer incidence could be obtained. The aims of this study are to demonstrate the cancer risks relating to lifestyle factors such as diet, smoking, alcohol drinking, familial history, past medical history, occupation and education, and to evaluate the effectiveness of the secondary cancer prevention program for gastric, cervical, lung, breast and colorectal cancers. Serum samples were collected from a part of the cohort subjects to measure several serum markers possibly associated with cancer. In addition, psychological state, which has been recognized as

Received March 27, 1995; accepted May 25, 1995.
¹Department of Public Health, Tohoku University School of Medicine, Sendai, Japan.
²Epidemiology and Biostatistics Division, National Cancer Center Research Institute, East, Kashiwa, Japan.
³Department of Educational Psychology, Faculty of Education, Tohoku University, Sendai, Japan.
⁴Laboratory of Nutrition, Department of Applied Biological Chemistry, Faculty of Agriculture, Tohoku University, Sendai, Japan.
⁵Miyagi Prefectural Cancer Registry, Sendai, Japan.
⁶Cancer Detection Center, Miyagi Prefectural Cancer Society, Sendai, Japan.
⁷Health and Welfare Department, Miyagi Prefectural Government, Sendai, Japan.
Address for correspondence: Akira Fukao, Department of Public Health, Tohoku University School of Medicine, 2-1 Seiryo-machi Aobaku Sendai, 980 Japan.

153
an important factor determining lifestyle and health behavior, was examined by using an authorized personality test.

This paper describes the epidemiological background and the study design, and provides the response rates and the profile of the cohort subjects.

SUBJECTS AND METHODS

Study Setting

Miyagi Prefecture, with 71 municipalities and a population of 2,176,295 (1985 Census), is economic and cultural center in the Northeastern District of Japan. About 1,040,000 of its citizens are employed; 15% in primary, 27% in secondary and 58% in the tertiary industries.

The Miyagi Prefectural Cancer Registry, established in 1959, is one of the oldest and the most accurate regional cancer registries in Japan. Data of cancer cases include not only official reports from clinics and hospitals but also information actively collected through the medical records of hospitals and the documents of pathology or autopsy. The percentage registered by death certificates only (%DCO) was 8.8% in the year of 1986. Data on cancer incidence, including site of cancer, basis of diagnosis, clinical stage and histological type, can be obtained from the registry records.

The Cancer Detection Center of Miyagi Prefectural Cancer Society has conducted screening for gastric, cervical, lung, breast and colorectal cancers exclusively in this prefecture, and the records of examinees, i.e. name, date of birth, address and results of screening tests, have been stored. By using these records, data on participation in the cancer screening of the cohort subjects can be obtained systematically.

The information concerning emigrations and deaths, needed to follow the cohort subjects, can be obtained from the regional health centers and the municipal governments.

Sample Size of the Cohort and Follow-up Period

In the initial plan for the study, the sample size, the age range, and the follow up period of the cohort were decided 50,000, 40 to 64 years of age, and 10 years, respectively. According to the data of the Cancer Registry, the expected number of annual incidence of cancer was estimated to be 200 in this cohort, so a total of 2,000 cases was anticipated for 10 years (500,000 person-years).

Identification of Cohort Subjects

Considering the efficiency and accuracy of collection of baseline and follow-up data of the cohort subjects, we decided to include all residents aged 40 to 64 years in several municipalities randomly selected from the 62 ones. Nine cities were, a priori, excluded because of high migration rates. These 62 municipalities were stratified into three groups according to the number of people aged 40 to 64 years based on the Census of 1985, i.e. 3,500 or less, 3,501 to 6,000, and 6,001 or more, and random sampling was carried out under the condition that the corresponding population totaled about 50,000 and even numbers of municipalities were allocated in each stratum. We added the latter condition to make two comparable groups for a future intervention study. Finally 14 municipalities with a total of population of approximately 52,000 were selected as the study areas.

Data Collection

We requested these 14 municipal governments to compile the corresponding residents registered as of April 1, 1990, and found that they actually totaled 53,464 (26,460 males, 27,004 females).

After excluding persons hospitalized or away on business for extended length of time and those judged to be unable to respond for some reasons such as physical or mental disabilities, we included a total of 51,921 residents (25,279 males, 26,642 females) as initial subjects at the baseline survey.

All initial subjects received self-completed questionnaires on lifestyle and personality from the leaders of public health activities in each community between May and September, 1990. The questionnaires completed and sealed by the subjects were collected by the same personnel within a few weeks.

The questionnaire on lifestyle consisted of about 90 items concerning 12 factors (past medical histories, familial histories, recent physical condition, drinking habit, smoking habit, dietary habit, occupation, marital status, education, screening histories, health insurance, and reproductive histories for women). Its feasibility had been repeatedly assessed in small populations prior to the survey.

For personality test, we used the short-form Eysenck Personality Questionnaire-Revised (EPQ-R), translated into Japanese, consisting of 48 simple questions to assess four dimensions of personality (psychoticism, extraversion/introversion, neuroticism, and lie scales). The psychometric properties of the test had been evaluated by one of the authors (T.H.) using another study sample.

Collection of Serum Samples

The persons who had responded to the questionnaire survey were briefly informed of the purpose of the research and asked if they would be willing to offer their blood samples when they attended at the annual health check-up program under the Health and Medical Services Law for the Aged. In this program, residents aged 40 and over, who are not covered by the health-insurance policies related to employment, are offered physical measurement (height and weight), blood pressure, electrocardiogram, urine tests (protein and glucose), and blood tests (blood cell counts, serum lipids, plasma glucose and liver func-
The serum samples were collected from only those accepting and signing their names on a prescribed form in each municipal center between May and September, 1990. The serum samples of 2-3 ml per individual were taken in brown colored tubes to shade light and were stored at $-80^\circ$C. We also obtained the data of the health check-up program. For reasons relating to manpower and costs, collection of the serum samples was carried out in only six municipalities.

**Confirmation of cohort subjects**

All data were coded according to a prescribed format and were entered into a computer. The subjects supplying their name, address and date of birth, and completing over half of the lifestyle questionnaire were regarded as eligible, while the others were excluded from the study. Even when the questionnaire was incomplete or erroneous, we did not request the person to complete the unanswered questions or to correct the errors. The rationality of the data was evaluated from the various directions and the inconsistency was treated according to the predetermined rules.

**Calculation of standardized mortality ratios (SMRs)**

To show epidemiological background, SMRs for all causes and several main causes of deaths (cancer of all sites, gastric cancer, lung cancer, cerebrovascular diseases and heart diseases) in Miyagi Prefecture and the study areas were calculated. They were calculated using the number of deaths by cause between 1983 and 1987 (based on the annual reports from the Prefectural Government) and the age-specific population (based on the Census of 1985) in Miyagi Prefecture and the study areas, and the age-specific mortality rates by cause of 1985 in Japan (based on the Vital Statistics from the Ministry of Health and Welfare). Accordingly, the SMR represents the ratio of the observed number of deaths to the expected one based on the mortality rates in Japan.

**Profiles of the cohort subjects**

To show the profiles of the cohort subjects, the proportion of those who married, the mean number of children, the mean age at the last education, the distribution of occupations, and the proportion of National Health Insurance holders were calculated using the baseline data. The National Health Insurance holders are the target popula-

### Table 1. Selected 14 municipalities and their population aged 40 to 64 years.

| Population size | No. of municipalities | Total population* (aged 40-64 years) | No. of municipalities where serum samples were collected |
|-----------------|-----------------------|-------------------------------------|--------------------------------------------------------|
| -3,500          | 8                     | 19,881                              | 3                                                      |
| 3,501-6,000     | 4                     | 18,559                              | 2                                                      |
| 6,001-8,000     | 2                     | 13,683                              | 1                                                      |
| Total           | 14                    | 52,123                              | 6                                                      |

*: according to Census 1985

### Table 2. Standardized mortality ratios (SMRs) by cause in Miyagi Prefecture and study areas.

|                | All causes | Cancer (all sites) | Gastric cancer | Lung cancer | Cerebrovascular diseases | Heart diseases |
|----------------|------------|--------------------|----------------|-------------|--------------------------|---------------|
| **Males**      |            |                    |                |             |                          |               |
| Japan          | 1.0* (552.9)† | 1.0 (150.0)        | 1.0 (40.6)     | 1.0 (28.0)  | 1.0 (84.4)               | 1.0 (94.1)    |
| Miyagi Prefecture | 0.97* (0.99)‡ | 0.98               | 0.97           | 0.98        | 1.21                     | 0.94          |
| Study areas    | 0.99* (0.92-1.04) | (0.89-1.06)     | (0.88-1.09)    | (1.14-1.29) | (0.89-1.00)              |               |
| **Females**    |            |                    |                |             |                          |               |
| Japan          | 1.0 (324.3) | 1.0 (80.3)         | 1.0 (19.2)     | 1.0 (7.7)   | 1.0 (58.9)               | 1.0 (59.0)    |
| Miyagi Prefecture | 1.0 (0.98-1.03) | 0.99             | 0.93           | 1.01        | 1.27                     | 0.96          |
| Study areas    | 1.02 (0.94-1.11) | 0.91               | 0.94           | 0.92        | 1.30                     | 1.07          |

*: SMR (1983-1987)
†: Age-adjusted mortality rate in 1985 (per 100,000)
‡: 95% confidence interval for SMRs computed based on Poisson distribution
Table 3. Response rates for baseline survey and number of eligible cohort subjects.

|            | No. of residents aged 40-64 | No. of persons surveyed (A) | No. of persons responding to the survey (B) | Response rate (B/A) | No. of eligible cohort subjects (C) | C/A |
|------------|-----------------------------|-----------------------------|---------------------------------------------|---------------------|-------------------------------------|-----|
| Males      | 26,460                      | 25,279                      | 23,232                                      | 91.9%               | 22,836                              | 90.3%|
| Females    | 27,004                      | 26,642                      | 25,009                                      | 93.9                | 24,769                              | 93.0 |
| Total      | 53,464                      | 51,921                      | 48,241                                      | 92.9                | 47,605                              | 91.7 |

Table 4. Distribution of cohort subjects by sex and age.

| Age       | Males (%) | Females (%) | Total (%) |
|-----------|-----------|-------------|-----------|
| 40-44     | 5,689 (25) | 5,301 (21) | 10,990 (23) |
| 45-49     | 3,696 (16) | 3,935 (16) | 7,631 (16) |
| 50-54     | 3,928 (17) | 4,556 (18) | 8,484 (18) |
| 55-59     | 4,657 (20) | 5,376 (22) | 10,033 (21) |
| 60-64     | 4,866 (21) | 5,601 (23) | 10,467 (22) |
| Total     | 22,836 (100) | 24,769 (100) | 47,605 (100) |

Table 5. Profiles of the cohort subjects.

|                                   | Males | Females |
|-----------------------------------|-------|---------|
| Married (%)                       | 96.2  | 97.5    |
| No. of children (Mean ± SD)       | 2.6 ± 1.0 | 2.7 ± 1.0 |
| Age at the last education (Mean ± SD) | 16.7 ± 2.4 | 16.6 ± 2.1 |
| Occupation (%)                    |       |         |
| School teacher                    | 1.9   | 1.6     |
| Clerk                             | 7.3   | 8.1     |
| Blue-collar worker                | 25.5  | 16.3    |
| Manager or official               | 9.4   | 0.8     |
| Agriculture and forestry          | 24.8  | 30.4    |
| Fishery                           | 6.8   | 1.8     |
| Sales                             | 4.3   | 5.4     |
| Housewife                         | 0     | 16.7    |
| None                              | 0.3   | 1.0     |
| Others                            | 19.7  | 17.9    |
| National Health Insurance holders (%) | 53.1  | 51.6    |

RESULTS

Epidemiological background

Table 1 shows the census population of selected study areas (14 municipalities) according to population size.

The standardized mortality ratios (SMRs) for all causes and several main causes in the study areas are shown in Table 2. The SMR for cancer of all sites was slightly low for both sexes in the study areas, but the difference was not statistically significant. The SMRs for cerebrovascular diseases were significantly high for both sexes not only in the study areas but also in Miyagi Prefecture.

Response rates

The number of corresponding residents registered as of April 1, 1990, the number of initial subjects surveyed by questionnaires, the number of persons responding to the survey and the number of persons eligible as cohort subjects are shown in Table 3. After excluding 1,543 persons who were unable or unlikely to respond for some reasons, we regarded 51,921 persons as the initial subjects and distributed questionnaires to them. The overall response rate was 92.9% (48,241/51,921), and the rate by municipalities ranged from 86.4% to 98.5%. After exclusion of the ineligible subjects according to the criteria, 91.7% (47,605/48,241) of the initial subjects were confirmed as eligible in our study. The proportion of the eligible in the initial subjects by municipalities ranged from 85.0% to 98.2%. Of these 41,442 persons (87.1%) also completed the EPQ-R. Serum samples were collected from 8,896 subjects (18.7%) in six municipalities.

Profile of the cohort subjects

The distribution of cohort subjects according to sex and age at the baseline survey and their demographic profiles are shown in Table 4 and Table 5, respectively. The proportion of the subjects engaged in agriculture, forestry and fishery was the highest for both sexes (31.6% for males and 32.2% for females). No differences in marital status, education and number of children were observed between sexes. A little more than half of the subjects (53.1% for males and 51.6% for females) were the National Health Insurance holders.

DISCUSSION

The mortality rate for gastric cancer, which is still the leading cancer, has been decreasing during the last 20 years in Japan, whereas those for lung, colorectal and breast cancers have been increasing. These trends are explained by the changes of lifestyle, especially dietary habit. We, therefore, believe that a prospective study including the recent population is needed to work out the strategy adapted to these changes.

The advantages of our study are that: 1) since the
response rate of baseline survey is satisfactorily high (92.1%), representative data would be available, 2) since not only lifestyle factors but also personalities are assessed using a standardized questionnaire (EPQ-R), the relation of cancer incidence to psychological state, which is believed to modify lifestyle and health behavior, can be evaluated, 3) since serum samples are collected from a part of the cohort subjects, the relationship between some biologic markers and cancer incidence can be assessed, 4) information of incidence of cancer can be obtained from a reliable cancer registry, 5) secondary prevention can be evaluated by use of a cancer screening database, and 6) the two comparable groups can be made for future intervention trials.

We have already collected the deaths and emigrations data as of March, 31, 1992, and found that the number of subjects who died from all causes and cancer of all sites, and those who emigrated from the study areas were 221 (4.64/1,000), 99 (2.08/1,000) and 359 (7.54/1,000), respectively. The data of cancer incidence during this period will be available within a few years. Approximately 4,500 serum samples have been analyzed for β-carotene levels and the results of a cross-sectional study on the correlation between the levels of β-carotene and several factors will be described elsewhere.

Although several case-control studies were carried out to evaluate screening programs for gastric (7), cervical (8), lung (9) and colorectal (10) cancers in Japan, the possibility of self-selection bias could not be ruled out because the possible factors which may affect both cancer death and participation in the screening were not taken into account in these studies. We believe that more reliable evaluation will be possible by using our cohort since the information including these factors has already been obtained and the screening data can be systematically obtained based on the records in the Cancer Detection Center.

One weakness of our study is that validity and reproducibility of the questionnaire employed in the baseline survey have not yet been evaluated. Although several validation studies suggested the usefulness of a self-administered food-frequency questionnaire (11), these results cannot always be generalized to other populations. Since we consider that the evaluation of validity and reproducibility of the questionnaires is necessary to guarantee the quality and reliability of a study as a whole, a sampling study is now under consideration.

Several other large scale population-based cohort studies have been started in the last few years in Japan (12,13). In these studies, including ours, special attention is paid to the recent changes of lifestyle among Japanese people, and blood samples are collected to analyze several biomarkers. Comprehensive application of the results of these studies is expected to contribute to formulation of the framework for cancer prevention.

**ACKNOWLEDGMENTS**

This work was supported by the Special Research Project of Cancer Prevention funded by Miyagi Prefecture. We thank all staff of municipal governments, the Public Health Union and the Miyagi Prefectural Cancer Society for their collaboration.

**REFERENCES**

1. Hirayama T. Diet and Cancer. Nutr Cancer, 1979; 1: 67-81.
2. Parkin DM, Muir CS, Whelan SL et al. eds. Cancer Incidence in Five Continents vol VI. IARC scientific publications No. 120, Lyon, 1992.
3. Eysenck SBG, Eysenck HJ, Barrett P. A revised version of the psychoticism scale. Person Individ Diff, 1985; 6: 21-29.
4. Hosokawa T, Ohyama M. Reliability and validity of a Japanese version of the short-form Eysenck Personality Questionnaire-Revised. Psychol Rep, 1993; 72: 823-832.
5. Watanabe S. Long term trends in cancer mortality rates from 1955 to 1985 in Japan. Jpn J Clin Oncol, 1989; 19: 305-317.
6. Oshima A, Hirata N, Ubuaketa T, Umeda K, Fujimoto I. Evaluation of a mass screening program for stomach cancer with a case-control study design. Int J Cancer, 1986; 38: 829-833.
7. Fukao A, Tsubono Y, Tsuji I et al. The evaluation of screening for gastric cancer in Miyagi Prefecture, Japan: a population-based case-control study. Int J Cancer, 1994; 60: 1-4.
8. Sobue T, Hashimoto S, Yokoi N, Fujimoto I. A case-control study of the effectiveness of cervical-cancer screening in Osaka, Japan. Jpn J Cancer Res, 1988; 79: 1269-1275.
9. Sobue T, Suzuki T, Ngaru T and The Japanese Lung-Cancer Screening Research Group. A case-control study for evaluating lung-cancer screening in Japan. Int J Cancer, 1992; 50: 230-237.
10. Hiwataishi N, Morimoto T, Fukao A et al. An evaluation of mass screening using fecal occult blood test for colorectal cancer in Japan: a case-control study. Jpn J Cancer Res, 1993; 84: 1110-1112.
11. Willett W. Nutritional Epidemiology. Monographs in Epidemiology and Biostatistics vol 15. Oxford University Press, New York and Oxford, 1990.
12. Aoki K ed. The reports on baseline analysis for the national collaborative cohort study of cancer. Nagoya, Aichi Cancer Center, 1993 (in Japanese).
13. Watanabe S. Large population-based prospective studies in Japan. Eur J Cancer, 1993; 29A: 2305-2314.