The familial aggregation of stomach cancer has long been argued. As early as 1959, Macklin found that gastric cancer was unduly frequent in the families of gastric-cancer probands, whereas husbands and wives were not affected with gastric cancer more often than would be expected on the basis of random distribution. His conclusion stressed the importance of the genetic similarity, rather than the environmental one, between the members of the affected pairs as an explanation for the increased frequency. Studies suggesting the familial aggregation of stomach cancer, however, seem to have been of little interest in Western countries, probably because of its low morbidity rate relative to cancers at other sites in Western populations. In contrast, stomach cancer remains high in both mortality and morbidity in Japan, accounting for 17% of all deaths from malignancy. Recently, consistent results from a number of studies conducted in Japan have indicated that stomach cancer tended to aggregate among parents and offspring in comparison with other sites.

Takaaki Kondo, Hideaki Toyoshima, Yoshie Tsuzuki, Yoko Hori, Hiroshi Yatsuya, Koji Tamakoshi, Akiko Tamakoshi, Yoshiyuki Ohno, Shogo Kikuchi, Kiyoshi Sakata, Yoshijharu Hoshiyama, Norihiko Hayakawa, Noritaki Tokui, Tetsuya Mizoue, Takesumi Yoshimura for the JACC Study Group

Accepted 13 February 2003

Background The aim of this study is to evaluate the magnitude of the aggregation of a stomach cancer history in parents and their offspring in comparison with that of a history at other sites.

Methods We used the baseline data from the Japan Collaborative Cohort Study for Evaluation of Cancer Risk (JACC Study), which was initiated during 1988–1990 in Japan. Association of the cancer history of the subjects’ parents with that of the subjects themselves and any of the subjects’ siblings was evaluated with odds ratios (OR) by the crude and generalized estimating equations (GEE) technique for four sites: stomach, colorectum, liver, and lung/bronchus.

Results The aggregation of a history of stomach cancer between parents and their offspring was evident with significant OR > 2.5. The magnitude of the parent–offspring association of a disease history of the colorectum and liver was found to be greater than that for stomach cancer. Conversely, lung and bronchus cancer failed to demonstrate a significant aggregation.

Conclusions The hereditary and environmental influences shared by parents and offspring are likely to play a strong aetiological role in colorectal or liver cancer versus a weaker but still significant role in stomach cancer. In contrast, the aetiological role of familial predisposition to lung cancer was indeterminate, which suggests a predominant role of non-familial factors in the development of lung cancer.

Keywords Familial aggregation, generalized estimating equation, parents and offspring, odds ratio

The familial aggregation of stomach cancer has long been argued. As early as 1959, Macklin found that gastric cancer was unduly frequent in the families of gastric-cancer probands, whereas husbands and wives were not affected with gastric cancer more often than would be expected on the basis of random distribution. His conclusion stressed the importance of the genetic similarity, rather than the environmental one, between the members of the affected pairs as an explanation for the increased frequency. Studies suggesting the familial aggregation of stomach cancer, however, seem to have been of little interest in Western countries, probably because of its low morbidity rate relative to cancers at other sites in Western populations. In contrast, stomach cancer remains high in both mortality and morbidity in Japan, accounting for 17% of all deaths from malignancy. Recently, consistent results from a number of studies conducted in Japan have indicated that stomach cancer tended to aggregate among
members of the patient’s family.\textsuperscript{2,3} Lately, we have found that a positive history of stomach cancer in one or more first-degree relatives was associated with a significantly increased risk of death from the disease in both men and women.\textsuperscript{4}

In addition to stomach cancer, familial aggregation of cancer at other concordant sites has also been elucidated.\textsuperscript{5} Based on the family history information from a large-scale cancer registry database, elevated rates of site-specific cancer among family members of affected patients have been revealed for cancers of the breast, colon and rectum, and stomach. Moreover, a clear correspondence in cancer history between husbands and wives was demonstrated,\textsuperscript{6} which suggests the contribution of the long-term environmental influences shared by family members to familial aggregation of cancer.

Notably, the magnitude of the familial aggregation is known to vary from site to affected site. According to the study of Ogawa \textit{et al.},\textsuperscript{2} stomach cancer is less likely to aggregate among family members than cancers of the colon and rectum, or breast. In order to verify the presence of a site-specific aggregation of cancer history between parents and their offspring, with special reference to the magnitude of the association between stomach and other cancers, we attempted to analyse the dataset from a cohort study to explore the possible relationship of the cancer history of parents with that of their offspring. Our approach can be compared with that of reported studies in that we used the baseline data from a large-scale, multi-centre, population-based study which had been developed to evaluate factors related to cancer mortality.\textsuperscript{7}

\section*{Subjects and Methods}

\subsection*{Study population}

The cohort of the Japan Collaborative Cohort Study for Evaluation of Cancer Risk sponsored by the Ministry of Education, Culture, Sports, Science and Technology (JACC Study) was established from 1988 through 1990, when 110,792 Japanese subjects (46,465 men and 64,327 women), aged 40–79 years, completed a questionnaire on lifestyle and medical history. They were enrolled from 45 study areas throughout Japan, mostly in the course of undergoing medical check-ups supported by the municipalities, and they were followed up to the end of 1997. This investigation was approved by the Ethical Board of the Nagoya University School of Medicine. For more details of the JACC Study design, the reader is referred to Ohno and Tamakoshi.\textsuperscript{8}

A standardized questionnaire was used to obtain basic demographic characteristics in addition to details on smoking, drinking and dietary habits, history of previous illness, and history of disease in parents and siblings. If any history of cancer was identified, the affected site was coded according to the Ninth Revision of the International Classification of Diseases (ICD-9). Cancer history in the parents was defined as present if the questionnaire stated that either the subject’s father or mother (blood-related) had been affected, whereas a history of cancer in the offspring was confirmed when the subject, or any of subject’s brothers or sisters (blood-related), had a history of the disease. Since our study was community-based, multiple counting of the same parents was likely to occur when any one of a subject’s brothers or sisters was also a subject of this study, and could eventually jeopardize the independence of measurement. To avoid this potential problem, we began by grouping the subjects into four subpopulations according to their birth order; eldest, second eldest, third eldest, and the remainder. By conducting the subsequent statistical analysis on a subpopulation basis, every history of a subject’s parents was given equal weight as a single count.

\subsection*{Statistical analyses}

Twofold statistical analyses were done in our analysis. In the first approach, we assessed familial aggregation of a single disease history in a usual case-control design. With cancer history in the parents and offspring dichotomized, two 2×2 contingency tables were developed for the calculation of crude odds ratios (OR) and a 95\% CI to determine whether the distribution of the offspring with a previous history of cancer was equivalent according to the presence or absence of parental cancer history. Since we were concerned with the familial aggregation of the cancer history with respect to a site-specific concordance, the OR were calculated separately for four major sites: stomach as ICD9-coded 151, colorectum as 153–154, liver as 155, and lung and bronchus as 162.

In the second approach, we employed a multivariate logistic model in which the disease history of the mother and father was treated as a function of the status of the disease history of the parents’ subjects and siblings. Let \(Y_f\), \(Y_m\), and \(Y_s\) denote the history status of the subject, subject’s brother (if present), and subject’s sister (if present), respectively, with \(Y_f = 1\) if case and \(Y_f = 0\) if control. Besides, let \(\text{Covar}\) denote the vector of covariates for the subject. The multivariate model we adopted can be expressed by two ordinal logistic regression equations with shared coefficients:

\[
\logit P (Y_f = 1 | Y_f, \text{Covar}) = \alpha_f + \beta Y_f + \theta \text{Covar}
\]

\[
\logit P (Y_m = 1 | Y_m, \text{Covar}) = \alpha_m + \beta Y_m + \theta \text{Covar}
\]

where \(Y_f\) and \(Y_m\) denotes the outcome status of the subject’s father and mother, respectively, in terms of the presence of the disease history. The regression parameters are estimated by solving a set of estimating equations. In this model, \(\beta\) is the log OR measuring the increase in the log odds of a disease history in the father or mother of a subject, the subject’s brother, or the subject’s sister, with the disease history compared with the father or mother with offspring without the disease history. For the covariate, we included terms for the subject’s age as a continuous variable. For this logistic method, all the observations are not assumed to be independent because observations are clustered within families. To account for this correlation, we made inference on \(\beta\) by using the generalized estimating equations (GEE) method. A detailed explanation of the model is available in a previous report.\textsuperscript{9}

In our analysis, some study areas were not included because of the total absence of data on the past or family history in each area, and individuals with unknown birth order were excluded in the present study as well, which reduced the number of eligible subjects to 79,540. To assure sufficient sample size for each subpopulation, we restricted the application of statistical procedures to three major categories of the birth order; eldest, second eldest, and third eldest. As a statistical software package, we ran the Statistical Analysis System (SAS) release 6.12 installed at the Nagoya University Computation Centre. The model fitting
was performed using the `freq` procedure for the estimation of univariate OR and the `genmod` procedure for GEE to adjust for the correlation of observations within families.10

**Results**

When the 79,540 subjects were divided into subgroups according to birth order, 20,509 of them were found to be eldest offspring, 16,369 second eldest, and 13,986 third eldest (25.8%, 20.6%, and 17.6%, respectively) (Table 1). Across the subpopulation in birth order, the mean age differed little, and the male/female ratio was somewhat constant with a moderate female dominance of 55%, 57%, and 59% for the eldest, second eldest, and third eldest subjects, respectively.

The number of eldest children whose father or mother had previously been affected with cancer at any site was 4,190. Of these, 389 subjects had a previous history of cancer themselves or among their siblings. In contrast, out of 16,319 eldest subjects without a parental cancer history, 849 subjects reported that they themselves or their siblings had a cancer history. This distribution difference yielded a crude OR of 1.9 (95% CI: 1.6, 2.1) with a significant familial aggregation of cancer observed when the previous history of parents and that of offspring were cross-classified (Table 2). In the GEE model, the age-adjusted OR remained at 1.9 (95% CI: 1.7, 2.1), which was still statistically significant (Table 2). The crude OR with respect to the familial aggregation of the cancer history for second eldest and third eldest subjects also showed a statistical significance at 1.8 (95% CI: 1.6, 2.0) and at 1.7 (95% CI: 1.5, 1.9), respectively. GEE-based age-adjustment appeared to have given only minimal changes in OR at the three birth-order levels.

There is a nearly 2.5-fold significant increase in the history of stomach cancer among parental history-positive offspring as compared with the parental history-negative offspring throughout all three categories of subject’s birth order, both before and after age adjustment by the GEE method. Cancers of the colorectum and liver also showed a significant aggregational relationship in the previous history between the parents and offspring. A comparison of the magnitude of the parent–offspring aggregation in terms of OR showed that colorectal and liver cancer substantially exceeded stomach cancer with OR of 3.9 and 4.9, respectively, for the subpopulation of eldest subjects. Moreover, the finding that stomach cancer had a smaller OR value than cancers of the colorectum and liver held for the remaining two birth order subpopulations. The multivariate-adjusted OR for the familial aggregation of these two disorders differed slightly from the corresponding crude OR, and remained statistically significant.

A previous history of lung cancer was unlikely to aggregate between parents and offspring as the OR failed to indicate statistical significance, although the number of offspring with a previous disease history was too small to definitively establish that conclusion.

**Table 1** Distribution of 79,540 subjects by birth order and sex

<table>
<thead>
<tr>
<th>Birth Order</th>
<th>Male n (%)</th>
<th>Female n (%)</th>
<th>Total n (%)</th>
<th>Age Mean ± SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eldest</td>
<td>9207 (27.9)</td>
<td>11,302 (24.3)</td>
<td>20,509 (25.8)</td>
<td>57.3 ± 10.5</td>
</tr>
<tr>
<td>Second</td>
<td>6999 (21.2)</td>
<td>9,370 (20.1)</td>
<td>16,369 (20.6)</td>
<td>57.6 ± 10.2</td>
</tr>
<tr>
<td>Third</td>
<td>5,723 (17.3)</td>
<td>8,263 (17.8)</td>
<td>13,986 (17.6)</td>
<td>58.3 ± 10.1</td>
</tr>
<tr>
<td>Other</td>
<td>11,066 (33.5)</td>
<td>17,610 (37.8)</td>
<td>61,665 (36.1)</td>
<td>57.9 ± 9.8</td>
</tr>
<tr>
<td>Total</td>
<td>32,995 (100)</td>
<td>46,545 (100)</td>
<td>79,540 (100)</td>
<td>57.7 ± 10.1</td>
</tr>
</tbody>
</table>

**Table 2** Odds ratio (OR) with respect to aggregation of cancer history between parents and offspring

<table>
<thead>
<tr>
<th>Sites of cancer</th>
<th>No. offspring with history/ no. parents with history</th>
<th>No. offspring without history</th>
<th>Crude OR (95% CI)</th>
<th>Age-adjusted OR$^a$ (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Eldest subject (n = 20,509)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All sites</td>
<td>389/4,190</td>
<td>849/16,319</td>
<td>1.9 (1.6, 2.1)</td>
<td>1.9 (1.7, 2.1)</td>
</tr>
<tr>
<td>Stomach</td>
<td>75/1,844</td>
<td>299/18,665</td>
<td>2.6 (2.0, 3.3)</td>
<td>2.5 (2.0, 3.4)</td>
</tr>
<tr>
<td>Colorectum</td>
<td>6357</td>
<td>89/2152</td>
<td>3.9 (1.8, 8.4)</td>
<td>3.9 (1.8, 8.5)</td>
</tr>
<tr>
<td>Liver</td>
<td>9/398</td>
<td>94/20,111</td>
<td>4.9 (2.6, 9.2)</td>
<td>5.0 (2.7, 9.4)</td>
</tr>
<tr>
<td>Lung and bronchus</td>
<td>5474</td>
<td>108/20,035</td>
<td>2.0 (0.8, 4.8)</td>
<td>2.1 (0.9, 5.1)</td>
</tr>
<tr>
<td><strong>Second eldest subject (n = 16,369)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All sites</td>
<td>398/3,479</td>
<td>859/12,890</td>
<td>1.8 (1.6, 2.0)</td>
<td>1.9 (1.7, 2.2)</td>
</tr>
<tr>
<td>Stomach</td>
<td>86/1,541</td>
<td>334/14,828</td>
<td>2.6 (2.0, 3.2)</td>
<td>2.5 (2.0, 3.2)</td>
</tr>
<tr>
<td>Colorectum</td>
<td>6289</td>
<td>94/16,080</td>
<td>3.6 (1.7, 7.9)</td>
<td>3.9 (1.7, 8.4)</td>
</tr>
<tr>
<td>Liver</td>
<td>11375</td>
<td>106/15,996</td>
<td>4.5 (2.6, 8.1)</td>
<td>4.8 (2.8, 8.6)</td>
</tr>
<tr>
<td>Lung and bronchus</td>
<td>3368</td>
<td>98/16,001</td>
<td>1.3 (0.4, 4.2)</td>
<td>1.5 (0.5, 5.0)</td>
</tr>
<tr>
<td><strong>Third eldest subject (n = 13,986)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All sites</td>
<td>430/2,993</td>
<td>1007/10,993</td>
<td>1.7 (1.5, 1.9)</td>
<td>1.8 (1.6, 2.1)</td>
</tr>
<tr>
<td>Stomach</td>
<td>107/1423</td>
<td>401/12,563</td>
<td>2.5 (2.0, 3.1)</td>
<td>2.6 (2.1, 3.2)</td>
</tr>
<tr>
<td>Colorectum</td>
<td>10222</td>
<td>115/13,764</td>
<td>5.6 (3.1, 10.1)</td>
<td>6.3 (3.5, 11.2)</td>
</tr>
<tr>
<td>Liver</td>
<td>7280</td>
<td>128/13,706</td>
<td>2.7 (1.3, 5.7)</td>
<td>3.1 (1.5, 6.5)</td>
</tr>
<tr>
<td>Lung and bronchus</td>
<td>2280</td>
<td>148/13,706</td>
<td>0.7 (0.2, 2.6)</td>
<td>0.8 (0.2, 3.3)</td>
</tr>
</tbody>
</table>

$^a$ Odds ratios based on generalized estimating equations.
Discussion

The familial aggregation of cancer has previously been studied mostly in a patient-based design where, first of all, the patient is identified as a proband case, and the tracking of family members of the proband follows for cancer morbidity. Large-sized cancer registry data are often used to ascertain the diagnosis of the disease, as in the studies done in Sweden by Hemminki et al. or in Aichi, Japan, by Kato. We initially designed a proband-based study in which we attempted to compare the proportion of parents with a cancer history for individual subjects with versus without a history of the same cancer. However, division of offspring into subjects and siblings for separate analyses diminished the cell frequency of those with a disease history in the contingency table for OR estimation to an extent that the statistical evaluation was impractical in all but stomach cancer. Hence, we built the current parent–offspring comparison design for familial aggregation of a cancer history. We previously created a study design where the cancer histories of the subjects and siblings were combined into offspring’s history, and its relationship with the parental disease history was examined for familial aggregation and coaggregation of a history for cancer, and some other diseases. The utilization of a large-scale cohort baseline database assists our study in that a sufficient size sample enabled us to divide the population into subpopulations according to the birth order of the subjects and thus to avoid the chance of the same person being counted multiple times. From the almost identical mean age across the birth-order categories (as shown in Table 1), we surmise that the actual multiple counting of parents occurred less frequently than originally expected. If the chance of a subject having a sib relationship with other subjects had been high, the mean age over the strata of the birth order would have differed accordingly.

In our first statistical approach for the estimation of crude OR, the usual case-control analysis was done to examine the association of disorder status of the parents with the presence of the disorder in their offspring. Overestimations of the familial aggregation possibly attend this measure because information was reduced by pooling both parents into a single observation unit. The subject, the subject’s brother, and subject’s sister were also clustered into a single unit. To overcome the bias toward overestimation, we fitted a multivariate logistic model in which the disorder status of each of the parents (father or mother) was predicted by the status of each of the offspring (subject, subject’s brother, or subject’s sister) with intrafamilial correlations accounted for by the GEE technique. Besides, the effect of the subject’s age was considered in the model as a covariate. We found that these two methods yielded explicitly similar results in estimation of OR. This finding suggests that the overestimation of familial aggregation resulting from the clustering of the relatives’ information on the disease history is not as substantial as anticipated.

The major finding in our study is that a familial aggregation between parents and offspring was observed for three concordant sites, i.e. stomach, colorectum, and liver, in addition to all-site cancer. The fact that this finding was consistently observed across all three categories of subject’s birth order serves to underscore our conjecture that a family history of cancer among parents significantly increases the risk of cancer morbidity at some concordant sites among their offspring.

The OR observed in the range 2.5–2.6 for the aggregation of a history of stomach cancer somewhat approximate the figure in our previous paper, where we found the familial aggregation of stomach cancer to be less prominent than that of stroke, hypertension, tuberculosis, or diabetes mellitus. While the OR for familial aggregation varied from site to site of each cancer, some similarities were evident between our result and the earlier one reported by Ogawa et al.; i.e. stomach cancer is lower than cancers of the colorectum or liver in the magnitude of OR for the aggregation of disease history. This consistency in the relative magnitude of site-specific aggregations could reflect a difference in the aetiological contribution of shared hereditary and environmental influence. These shared factors may be less determinant for the familial aggregation of stomach cancer than for colorectal and liver cancers.

Lung and bronchus cancer failed to demonstrate any significant evidence of a parent–offspring aggregation, though the limited number of observed patients is likely to compromise the validity of such a conclusion. The lack of clear evidence for familial aggregation of lung cancer is consistent with the previous literature, while a case-control study indicated an increase in the risk of lung cancer in young adults but not in an older group, suggesting differences in the risk factors between younger and older people. Although ages at which the history of illness occurred were not available in our data, taking into consideration the mean age of the subjects, it is likely that a great proportion of the parents with lung cancer were affected at advanced ages. Presumably, acquired susceptibility due to the effects of some risk factors not shared by family members may outweigh any shared effects in the development of lung cancer in later life. In particular, smoking is a predominantly non-familial risk factor for lung cancer, and we have previously shown that the effect of smoking on lung cancer mortality continues into later life even after cessation of the smoking habit.

A mechanism to explain the familial aggregations we observed remains to be elucidated. Mutations of some genes have been proven to confer a great deal of lifetime risk in familial settings. These include RB in retinoblastoma, WT1 in Wilms tumour, and p53 in Li–Fraumeni syndrome. However, this type of mutation is estimated to be quite uncommon, e.g. inherited mutations in p53 accounting for less than 1% of breast cancer patients, even at young ages. More frequent types of gene variants classified as a genetic polymorphism are thought to carry high population risks. By influencing metabolic activation, detoxification, or the elimination of carcinogens, the interaction between low-penetrant gene susceptibility and environmental factors shared both by the parents and offspring is likely to contribute to the occurrence of a familial aggregation of cancer. Some examples of the gene–carcinogen interaction in relation to gastrointestinal cancer are heterocyclic amines and CYP1A1, alcohol and the ADH phenotype, and so on.

Our study is subject to some limitations associated with information bias, including among others, Japan’s medical culture. This discourages physicians from informing cancer patients of their true diagnosis because of fear of the incurable nature of the disease, with the result that measurement errors related to under-reporting are likely to ensue. By classifying the subjects and their siblings into the single category of offspring, we attempted to reduce the consequence of this bias, though the eventual effects of such a countermeasure for the bias could not be tested.
Various recall biases in relation to the previous history of the subjects’ siblings or parents may well be involved. Since individuals with a parental cancer history became more conscious of the disease than those without one, they might be prone, at least subconsciously, to evoke the recollection of such a family history, which can result in the differential misclassification of the existence of the history. Another problem concerns the ICD-9 coding of the disease. As we could not ascertain a diagnosis in any cancer history self-reported by the subjects, the misclassification or omission of the ICD-9 codes could not be ruled out, which also renders our results less convincing. In this respect, a study to confirm the reliability of the particulars of a given previous and family history is underway.17

In summary, our analysis revealed that a history of stomach cancer among parents is associated with an increased risk for the same cancer among their offspring. When OR were calculated to quantify the risk association for some affected sites, stomach cancer had a lower OR than cancers of the colorectum and liver, in agreement with the previous literature. Our result suggests the importance of family-shared factors whose aetiological influence is likely to differ substantially from site to site of cancer history.

Acknowledgements
The authors wish to express their appreciation to Dr Kunio Aoki, Professor Emeritus, Nagoya University School of Medicine and the former chairman of the JACC Study Group, and to Dr Haruo Sugano, the former Director, Cancer Institute of Japanese Foundation for Cancer Research, who greatly contributed to the initiation of this study. This work was supported by a Grant-in-Aid for Scientific Research on Priority Areas (C)(2) (No.12218216) from the Ministry of Education, Culture, Sports, Science and Technology of Japan. The JACC Study was also supported by Grants-in-Aid for Scientific Research from the same Ministry (No. 63010074, 1010068, 2151065, 3151064, 4151063, 5151069, 6279102, and 11181101).

KEY MESSAGES

• The hereditary and environmental influences shared by parents and offspring are likely to play a strong aetiological role in colorectal or liver cancer, but the role seems to be weaker for stomach cancer.

• The aetiological role of familial predisposition to lung cancer was indeterminate, which suggests a predominant role of non-familial factors in the development of lung cancer.

References