

Segregation Analysis of Gastric Cancer in a Japanese Population

Masao Kanamori, Audrey H. Schnell, Manami Inoue, Yoshitaka Yamamura, Youjie Wang,
Mizue Suzuki, Hiroichi Takeuchi, Kazuya Shinmura, Jun Yokota, Kazuo Tajima,
Robert C. Elston and Haruhiko Sugimura

Affiliations of authors: M. Kanamori, M. Suzuki, H. Takeuchi (Department of Public Health), Y. Wang, K. Shinmura, H. Sugimura (Department of Pathology), Hamamatsu University School of Medicine, Hamamatsu, Shizuoka, Japan; M. Inoue, Y. Yamamura, K. Tajima, Division of Epidemiology and Prevention, Aichi Cancer Center Research Institute, Japan; J. Yokota, Biology Division, National Cancer Center Research Institute, Tokyo; A.H. Schnell, R.C. Elston, Department of Epidemiology and Biostatistics, Case Western Reserve University, Cleveland, Ohio, USA

KEY WORDS Mode of inheritance; multifactorial; age of onset

ABSTRACT Though its incidence and mortality have declined, gastric cancer is still the most common type of cancer in Japan. The cause of gastric cancer appears to be both genetic and environmental, with the possibility that the differentiated (internal) type is more environmentally determined than the non-differentiated (diffuse) type. Prior to this study, no formal segregation analysis of gastric cancer in Japan has been performed. Segregation analysis of gastric cancer that allowed for variable age of onset was performed on 851 two-generational pedigrees ascertained via gastric cancer probands collected from the Hospital based Epidemiological Research Program in the Aichi region of Japan. Families were classified based on the proband's histopathological classification of having either differentiated type or non-differentiated type of cancer. A random no major effect hypothesis was rejected, as was a purely recessive or dominant hypothesis. The most parsimonious model was one of purely multifactorial inheritance with males having a higher susceptibility than females. Under a model where genotype influences age of onset, a dominant or recessive mode of inheritance with multifactorial effects also fitted the data. In addition, the analyses were performed separately for the differentiated type and the non-differentiated type and homogeneity was not rejected.

INTRODUCTION

Gastric cancer is the most common type of cancer in Japan, though the incidence and mortality have markedly decreased during the past several decades. Many authors have indicated several etiologic factors in the development of gastric cancer (Kato et al. 1981; Young et al. 1988; Powell et al. 1990; Blot et al. 1991; Lauren 1965; Lauren and Nevalainen 1993; Tredaniel et al. 1997; Fleischauer et al. 2000). It

has been hypothesized that the so-called intestinal (differentiated) type of gastric cancer is more dependent on environmental factors and the diffuse type (non-differentiated) type more on host-related factors. Inoue et al. (1999) found a significantly increased odds ratio of gastric cancer for habitual smokers, and this was higher both for the differentiated type than for the non-differentiated type, and at younger than at older ages. These authors also pointed out that these risk factors were less pronounced in female than in male patients. The cause of gastric cancer appears to be multifactorial.

Familial aggregation of gastric cancer has been documented for many years (Creagan et al. 1973; Gar'kavtseva et al. 1987; Janerich et al. 1990; La Vecchia et al. 1992; Mecklin, et al. 1988; Triantafillidis et al. 1993; Zanghieri et al. 1990; Zhao et al. 1994; Munoz et al. 1997; Caldas et al. 1999) but there have been few systematic studies. Common dietary habits, such as intrafamilial exposure to food carcinogens, and *Helicobacter pylori* infection have been suspected as contributory factors. Genetic aspects have only been recently considered, especially in the case of hereditary non-polyposis colorectal cancer, with which gastric cancer is one of the most frequently associated neoplasms second only to uterine cancer. Several studies have been performed examining the microsatellite instability of familial gastric cancer, but the contribution of this phenotype does not seem to be as important as in hereditary non-polyposis colorectal cancer (Shinmura et al. 1997; Shinmura et al. 1998). There are still no genetic markers for gastric cancer predisposition except in a small subset of familial cases in Maori (Guilford et al. 1998).

The mode of inheritance for gastric cancer in Japan has not been previously studied. The purpose of this analysis is to investigate the mode of inheritance of gastric cancer in a Japanese

Correspondence to: Masao Kanamori, Department of Public Health Hamamatsu University School of Medicine, Hamamatsu, Japan 431-3192.
Tel: 81-53-435-2329 *Fax:* 81-53-435-2330
E-mail: kanamori@hama-med.ac.jp

population with particular attention to differences between men and women.

SUBJECTS AND METHODS

Study Subjects

The present study consisted of probands with gastric cancer and their families collected between January 1988 and June 1995 by the Hospital based Epidemiologic Research Program at Aichi Cancer Center (HERPACC) in Aichi prefecture, the central area of Japan. Information on life-time family history was routinely collected from patients on their first visit, using a self-administered questionnaire checked by a trained interviewer. Each patient was asked about the health of his or her first degree relatives. The data procedures used here are as described by Inoue et al. (1997, 1998, 1999). There was a total of 879 gastric cancer probands, newly diagnosed on the basis of histopathologic examination determined by the Japanese classification of gastric carcinomas (Kanehara, 1996), who responded to the HERPACC questionnaire during this time period. In accordance with this system, each tumor in this study was originally classified as one of the following: papillary adenocarcinoma; tubular adenocarcinoma; well-differentiated adenocarcinoma; moderately differentiated type; poorly differentiated adenocarcinoma; solid type; non-solid type; signet-ring cell carcinoma; mucinous adenocarcinoma; adenosquamous carcinoma; squamous-cell carcinoma or carcinoid tumor. These histological types were then grouped into two major categories according to the degree of structural differentiation: "differentiated type" or "non-differentiated type". We used a conversion table (Hanai and Fujimoto 1982) of the Lauren (1965) and Japanese classifications for the grouping. Those histological types that fitted neither of the 2 major categories, e.g., adenosquamous carcinoma, squamous-cell carcinoma, or carcinoid tumor, were excluded from the analyses.

There was one proband per family. For other family members (on average 5.65 per family), information on gastric cancer status, age of onset, age at examination (the last identified age at which a person is known to be affected or unaffected), and whether they were dead or alive was provided by the proband and any other relative(s) present at the time of the visit.

A total of 28 probands were excluded from the analysis; 8 probands had unclassified type and 20 probands had no information on histopathological type. Of the 851 probands whose families were analyzed, 387 had differentiated type and 464 had non-differentiated type (see Table 1). Each pedigree member was classified as one or the other type based on the proband's histopathological type. In other words, all affected family members were assigned the same type as the proband. For those affected family members whose actual histopathological type was determined, the type was usually consistent with the assigned type. Positive gastric cancer status for first degree relatives (the parents and siblings) were validated in a sample of the reported cases by medical record review and found to be accurate.

STATISTICAL METHODS

Segregation analysis was performed using the REGTL program that is part of the Statistical Analysis for Genetic Epidemiology (S.A.G.E.) computer package (S.A.G.E. 1998). REGTL performs maximum likelihood segregation analysis of a dichotomous trait with a variable age of onset that follows, possibly after transformation, a logistic distribution. REGTL uses a modified Class A regressive model (S.A.G.E., 1998) that allows for a residual sibling correlation (Karunaratne et al. 1998). The models in REGTL allow for up to 3 'types' of individuals (AA, AB and BB), where type refers to the presence of some factor (A or B) that can be transmitted from generation to generation. Type is defined in terms of transmission: two people are of the same type if and only if their offspring by a mate of given type have the same phenotypic distribution. The probability that this factor is transmitted from parent to offspring is a transmission probability which depends on the parent's type: the probability that a person of a given type transmits factor A to offspring. Mendelian inheritance, if it occurs, is assumed to be through a single autosomal locus with 2 alleles (A and B), where A is the allele associated with the disease. The word type is used generally, whatever the mode of transmission, Mendelian inheritance being a specific mode of transmission in which the types are genotypes and the transmission probabilities for the three genotypes (AA, AB and BB) are 1, .5 and 0, respectively. We assume that a proportion of people in the population are susceptible to the disease and a

proportion are not susceptible; the latter would never become affected, however long they lived, and so for them age of onset is irrelevant. Let susceptibility, γ , be the probability that a person is susceptible. The likelihood for an individual who is affected at age a is taken to be $\gamma f(a)$, that for an individual affected by age at examination a' (if age of onset is unknown) to be $\gamma F(a')$, and for an individual who is unaffected by age a' to be $1 - \gamma F(a')$, where f is a power transformed logistic density and F is the corresponding cumulative distribution. The density function is thus of the form

$$f(a) = \frac{\alpha \exp[\beta + \alpha a + \text{covariates} + \text{familial effects}]}{1 + \exp[\beta + \alpha a + \text{covariates} + \text{familial effects}]^2}$$

and the cumulative distribution is

$$F(a) = \frac{1}{1 + \exp\{-[\beta + \alpha a + \text{covariates} + \text{familial effects}]\}} \text{ where}$$

a denotes power transformed age, β is a baseline parameter, $\hat{\alpha}$ is the age coefficient, the mean of the distribution conditional on familial effects and covariates is $-(\beta + \text{familial effects})/\hat{\alpha}$, and the variance of the conditional distribution is $\frac{1}{\hat{\alpha}^2}$. The familial effects include spouse and parental effects quantified by regressive coefficients in order to model multifactorial transmission (Bonney 1986). REGTL also allows for the inclusion of a sibling covariate equal to the proportion of other siblings in the sibship who are affected. Major effect transmission is modeled by the three transmission probabilities τ_u ($u=AA, AB$ or BB), the probability that a person of type u transmits A to an offspring. If Hardy - Weinberg equilibrium is assumed, then the relative frequencies of the 3 types of individuals in the population are q^2 , $2q(1-q)$ and $(1-q)^2$ for AA , AB and BB , respectively, so that q is the frequency of allele A when there is Mendelian transmission.

Two different biological models for how types determine disease are available in REGTL. In the first method (Model 1), types (or genotypes in the case of Mendelian inheritance) influence the age of onset distribution. Here all three types of individuals have the same susceptibility to disease γ , and there can be a separate baseline parameter β associated with each type ($\beta_{AA}, \beta_{AB}, \beta_{BB}$). There may be a common age coefficient $\hat{\alpha}$, or $\hat{\alpha}$ can also be made type-dependent. Alternately, in the second method (Model 2), type influences susceptibility. There are therefore up to 3 different susceptibilities ($\gamma_{AA}, \gamma_{AB}, \gamma_{BB}$), but a single age of onset distribution for all individuals, so that the parameters β and $\hat{\alpha}$ are

common to all types. When there is only one type, Model 1 and Model 2 are necessarily equivalent. In both models, the baseline parameter(s) and susceptibility(ies) can be made sex-dependent.

Because families were ascertained via a proband, a correction for ascertainment was included. Single ascertainment was allowed for by conditioning the likelihood on the proband's actual age of onset. It is advantageous to condition on the probands being affected by their ages at examination (or, if age at examination is unavailable, their ages of onset) because this retains more information. However, in these families age at examination and age of onset were highly correlated and so it was more appropriate to condition on age of onset. In addition, because there was heterogeneity across generations, conditioning on the founders was also employed.

A random, no major effect (NME) hypothesis and hypotheses corresponding to either a dominant mode or recessive mode of inheritance were fitted to the data. Hypotheses were also tested that allowed for multifactorial effects of spouses, mothers, fathers, and siblings. When estimating the multifactorial effects, restrictions were placed on the parameters such that the effect of an unaffected parent or spouse could not be greater than 0, and that of an affected parent or spouse could not be less than 0. Similarly, the coefficient of the sibling covariate was restricted to be non-negative. Multifactorial effects were added to the dominant and the recessive hypothesis. A purely multifactorial hypothesis with no major effect, allowing for multifactorial effects of spouses, mothers, fathers, and siblings, was also fitted to the data.

The likelihoods of two hypotheses were tested, one against the other, when one could be considered as a special case of the other. Under certain conditions, twice the difference in $\log_e$ likelihoods is asymptotically distributed as chi-square when the more restricted hypothesis holds, with degrees of freedom (d.f.) equal to the difference in the number of functionally independent parameters being estimated between the two hypotheses. We use this distribution as an approximation in the upper tail for all cases (Atwood et al. 1995) to test general departure from the restricted hypothesis, although sometimes the asymptotic distribution is a mixture of chi-square distributions (Self and Liang 1987). All hypotheses were compared to a general model with multifactorial effects where the transmission probabilities are estimated, but with the restriction

of homogeneity of the trait distribution across generations (Demenais and Elston 1981). In addition, the hypothesis that included both a major gene and a multifactorial component was compared to the hypothesis with only a multifactorial component (no major gene component) and to the hypothesis with only a major gene component.

In addition to the significance tests indicated above, Akaike's (1974) information criterion (AIC), defined as $AIC = -2 \log \text{Likelihood} + 2 \times (\text{number of independent parameters estimated})$ was used to compare maximum likelihoods of different hypotheses. Of all the hypotheses tested, that with the smallest AIC value is considered the most likely to explain the data.

Hypotheses were fitted simultaneously estimating a power parameter (λ) (Box and Cox 1964) to transform for the age of onset distribution. Several sets of initial estimates were used for each hypothesis fitted in order to find the global maximum, rather than a local maximum, of the likelihood. Segregation analysis was performed both for all families together and separately for families grouped by histopathological type. Tests for homogeneity were performed using the best fitting hypothesis by comparing twice the sum of the maximum log likelihoods for the two separate subgroups of families to twice the maximum log likelihood for all families combined.

RESULTS

Table 1 shows the number of affected and unaffected subjects by histological type of gastric cancer and sex. For the differentiated type, there were 154 affected females (28.8%) and 380 (71.2%) affected males. There were 1065 (50.2%) unaffected females and 1057 (49.8%) unaffected males. For the undifferentiated type, there were 243 affected females (41%) and 350 (59%) affected males. There were 1222 (50.7%) unaffected females and 1188 (49.3%) unaffected males. For both types of cancers there were more affected males than females, but the numbers of unaffected males and females were similar (see Table 2).

Table 3 gives the mean age of onset and mean age at examination by histological type of gastric cancer and sex. For the differentiated type, the mean age of onset was 61.7 years and 60.81 years for females and males, respectively. For the undifferentiated type, the mean age of onset was 56 and 56.4 years for males and females, respectively. The age of onset for males and females was similar regardless of histological type. The mean age of onset for the differentiated type was lower than for the non-differentiated type. For unaffected individuals, the mean age at examination was similar for both types and slightly younger for males than females.

Table 4 shows the results under Model 1,

Table 1: Number of affected and unaffected subjects by type of gastric cancer and sex

Type		Female No. (%)	Male No. (%)	Total No.
Differentiated	Affected	154(28.8)	380(71.2)	534
	Unaffected	1065(50.2)	1057(49.8)	2122
	Total	1219(46.0)	1437(54.0)	2656
Non-differentiated	Affected	243(41.0)	350(59.0)	593
	Unaffected	1222(50.7)	1188(49.3)	2410
	Total	1465(49.0)	1538(51.0)	3003
Combined		2684(47.0)	2975(53.0)	5659

Table 2: Number of family members by relationship to proband, type of cancer, affection status and sex

Relationship	Differentiated				Non-differentiated			
	Affected		Unaffected		Affected		Unaffected	
	Male	Female	Male	Female	Male	Female	Male	Female
Probands	286	101	0	0	271	193	0	0
Parents	44	22	323	345	39	28	412	423
Children	15	3	29	34	10	2	12	16
Siblings	35	28	702	669	30	20	761	773
Spouses	0	0	3	17	0	0	3	10

Table 3: Mean Age of Onset (and Standard Deviation) in years for Affected Subjects, and Mean Age at Examination (and Standard Deviation) in years for Unaffected Subjects, by Sex Separately for Differentiated and Non-Differentiated Gastric Cancer

	Mean Age of Onset		Mean Age at Examination*	
	Females	Males	Females	Males
Differentiated	61.71 (11.95) N=154	60.81 (9.89) N=380	60.17 (17.59) N=1059	58.47 (17.61) N=1049
Non-differentiated	56.00 (11.88) N=243	56.37 (11.32) N=350	59.18 (17.11) N=1210	57.53 (17.40) N=1179
Combined	58.21 (12.21) N=397	58.68 (10.82) N=730	59.64 (17.34) N=2269	57.97 (17.50) N=2228

*35 subjects has missing information

Table 4: Parameter Estimates and -2 Log_e Likelihoods for Model 1 Hypotheses. Hypotheses are tested against the General Model

Parameter	General Model	NME	Dominant	Dominant with Multifactorial	Recessive with Multifactorial	Multifactorial ^a
τ_{AA}	.727	---	1	1	1	---
τ_{AB}	.361	---	.5	.5	.5	---
τ_{BB}	.361	---	0	0	0	---
q	.095	---	.001	.001	.031	---
β_{AA}	-3.757	-1.075	-9.220	-4.023	-2.884	-8.559 (2.103)
β_{AB}	$=\beta_{AA}$	$=\beta_{AA}$	$=\beta_{AA}$	$=\beta_{AA}$	$=\beta_{AA}$	$=\beta_{AA}$
β_{BB}	-7.528	$=\beta_{AA}$	-16.352	-11.029	-9.661	$=\beta_{AA}$
α	.142	.096	.103	.108	.108	.104 (.008)
λ	1.508	.778	.344	.670	.683	.988 (.358)
Mother Unaffected	-.889	---	---	-.663	-.662	-.631 (.348)
Mother Affected	0*	---	---	0*	0*	0*
Father Unaffected	-2.341	---	---	-.346	-1.597	0*
Father Affected	0*	---	---	1.324	.073	1.646 (.266)
Female γ	.375	.546	.548	.372	.377	.361(.105)
Male γ	.657	1.000	1	.625	.658	.633(.170)
-2 Log _e (likelihood)	1911.67	1968.09	1963.36	1914.29	1915.16	1918.64
χ^2	---	59.06	51.69	2.62	3.49	6.97
Degrees of freedom	---	8	6	2	2	4
P value	---	<.001	<.001	.270	.175	.137
AIC		1978.09	1977.36	1936.29	1937.16	1936.64

* Parameter estimate converged to a bound.

^a Numbers in parentheses are standard errors.

See text for definition of all symbols.

where (geno) type influences age of onset. In the initial analyses, the sibling covariate and spouse effects were estimated to be 0 and therefore fixed at that value for all subsequent analyses. Allowing the susceptibilities to be dependant on sex resulted in much larger log_e likelihoods and so sex dependent susceptibilities were included in all models. However, allowing the baseline parameter and/or the age coefficient to depend on sex did not significantly improve the likelihoods. The NME hypothesis was rejected when compared to the general model with multifactorial effects. A dominant hypothesis had a slightly larger log_e likelihood than the

recessive hypothesis (data not shown). Similarly, the dominant hypothesis with multifactorial effects had a slightly larger log_e likelihood than the recessive hypothesis with multifactorial effects. Neither was rejected when compared to the general model with multifactorial effects. The purely multifactorial hypothesis did not fit significantly worse than the dominant with multifactorial effects hypothesis ($\chi^2_2 = 4.35$, $P=.114$) but the purely dominant hypothesis did ($\chi^2_4 = 49.07$, $P<.001$). The most parsimonious hypothesis was one of purely multifactorial inheritance where females were less susceptible than males. However, the best fitting

hypothesis according to the AIC was one of dominance with multifactorial inheritance, but there was only a very small difference in AIC between this hypothesis and the purely multifactorial one.

Table 5 shows the results under Model 2, where the (geno)types influence susceptibility. As under Model 1, the sex dependant susceptibilities were again highly significant, while allowing the baseline parameter and/or age coefficient to depend on sex did not significantly improve the likelihood. Now the recessive hypothesis had a slightly larger $\log_e$ likelihood than the dominant hypothesis (data not shown) and the recessive with multifactorial effects hypothesis had a slightly larger $\log_e$ likelihood than the dominant with multifactorial effects hypothesis. The NME hypothesis was again rejected. When compared to the general model with multifactorial effects, the recessive with multifactorial effects hypothesis was not rejected, but the dominant with multifactorial effects hypothesis was (barely) rejected. The purely

multifactorial hypothesis did not fit significantly worse than the recessive hypothesis with multifactorial effects ($\chi^2 = 4.32$, $P = .229$), but the purely recessive hypothesis did ($\chi^2 = 39.69$, $P < .001$). The most parsimonious hypothesis was one of purely multifactorial inheritance where females were less susceptible than males. Using the AIC criterion, this was also the best hypothesis.

The families classified as differentiated or non-differentiated gastric cancer were analyzed separately (data not shown) and homogeneity between the two subtypes was not rejected ($P = .984$).

DISCUSSION

Our report is the first segregation analysis of Japanese gastric cancer families. Since gastric cancer is the most common cancer in Japan, multiple cases within a family could occur by chance alone. The most parsimonious hypothesis was one of multifactorial inheritance where females are less susceptible than males and this

Table 5: Parameter Estimates and $-2 \log_e$ likelihoods for Model 2 Hypotheses. Hypothesis were tested against the General Model

Parameter	General Model	NME	Recessive	Dominant with Multifactorial	Recessive with Multifactorial	Multifactorial ^a
τ_{AA}	0*	---	1	1	1	---
τ_{AB}	.914	---	.5	.5	.5	---
τ_{BB}	.914	---	0	0	0	---
q	.593	---	.441	.099	.475	---
β	-6.861	-10.075	-7.839	-8.238	-8.301	-8.559(2.103)
α	.103	.096	.106	.105	.105	.104 (.008)
λ	.963	.778	1.128	1.008	1.103	.988 (.358)
Mother Unaffected	-.895	---	---	-.427	-.425	-.631 (.348)
Mother Affected	0*	---	---	0*	0*	0*
Father Unaffected	-1.664	---	---	-.443	-.359	0*
Father Affected	0*	---	---	1.220	1.306	1.646 (.266)
Female γ_{AA}	.781	.546	.700	1	1	.361 (.105)
Female γ_{AB}	$=\gamma_{BB}$	$=\gamma_{AA}$	$=\gamma_{BB}$	$=\gamma_{AA}$	$=\gamma_{BB}$	$=\gamma_{AA}$
Female γ_{BB}	.051	$=\gamma_{AA}$	.058	.154	.1	$=\gamma_{AA}$
Male γ_{AA}	.607	1.000	1.000	.723	.743	.633 (.170)
Male γ_{AB}	$=\gamma_{BB}$	$=\gamma_{AA}$	$=\gamma_{BB}$	$=\gamma_{AA}$	$=\gamma_{BB}$	$=\gamma_{AA}$
Male γ_{BB}	.715	$=\gamma_{AA}$	.216	.593	.573	$=\gamma_{AA}$
$-2 \log_e$ (likelihood)	1909.03	1968.09	1958.33	1915.03	1914.32	1918.64
χ^2	---	59.06	49.30	6.00	5.29	9.61
Degrees of freedom	---	9	6	2	2	5
P value	---	<.001	<.001	.05	.07	.087
AIC		1978.09	1958.33	1939.03	1938.32	1936.64

* Parameter estimate converged to a bound.

^a Number in parentheses are standard errors.

See text for definition of all symbols.

model is necessarily equivalent under either biological model. Under both Model 1 and Model 2, sex dependent susceptibilities and the multifactorial effects were highly significant. Under Model 1, a dominant with multifactorial effects hypothesis and a recessive with multifactorial effects hypothesis also fitted the data. Under Model 2, a recessive with multifactorial effects hypothesis also fitted the data, but the dominant with multifactorial effects hypothesis was rejected. Except in those cases where the two models must be the same (the NME and multifactorial hypotheses) or where the hypothesis does not fit the data (the purely Mendelian hypotheses), the AIC criterion favors Model 1. When the effect of the mother was set equal to the effect of the father (i.e. equal parent effects) all modes of inheritance, including the purely multifactorial, were rejected when compared to the general model. Unfortunately, environmental covariates such as smoking were not available in this analysis as this information was collected only for the probands.

A different risk profile has been reported for the two subtypes of gastric cancer (Buiatti et al. 1991), and a strong familial history of early onset gastric cancer (mostly diffuse type) has also been reported (Mecklin et al. 1988). Carneiro reported cases of the diffuse type of gastric carcinoma that segregated in an autosomal dominant manner in a polyposis family (1993). Bonney et al. studied the segregation of gastric pH, nitrate, and nitrites in a random sample of Columbian families, where gastric cancer is common (1987). They found a bimodal distribution of nitrite concentration levels and suggested recessive Mendelian inheritance for this phenotype. Nitrite is implicated in the formation of carcinogenic N-nitroso compounds, suggesting that elevated gastric nitrite concentration could be a biological marker for persons predisposed to gastric cancer. Most cases of Japanese gastric cancer are accompanied by intestinal metaplasia and atrophic gastritis. The results of our analysis are compatible with autosomal recessive inheritance together with multifactorial effects for both types of gastric cancer families. Bonney et al. (1987) estimated the frequency of a recessive allele related to measurable gastric nitrite to be 0.57; our estimates of gene frequency for a recessive gene were 0.31 under Model 1 (Table 4) and 0.48 under Model 2 (Table 5). This pattern of mixed inheritance is reasonable when we consider the requirement of multi-step changes in the full sequence of gastric carcinogenesis. Combined

with recent documentation regarding a modest contribution of environmental risk to familial gastric cancer vs. sporadic gastric cancer, (Shimura et al. 1997), we can expect the familial aggregation of gastric cancer to be partially genetic in nature. Larger pedigree structures and the inclusion of environmental covariates in the analysis would make the results of any segregation analysis more discriminatory. Biological markers such as *Helicobacter pylori* infection, the details of pathological information about the background mucosa, genetic markers in tumors and germlines, in addition to information on the histological subtypes, would further clarify the genetic and environmental etiology of human gastric carcinogenesis (Oda et al. 1994; Iida et al. 1999).

ACKNOWLEDGEMENT

This work is supported by Grant-in-Aid for Scientific Research on Priority Areas, Grant-in-Aid for Scientific Research 1998-B from the Ministry of Education, Culture, Science, and Sports, Grant in aid from the Ministry of the Health and Welfare, the Smoking Research, and the Uehara Memorial Foundation. The contributions of the participants of the Japanese gastric study group are deeply appreciated. This work was also supported in part by US Public Health Service research grant GM 28356 from the National Institute of General Medical Sciences, and by resource grant RR 03655 from the National Center for Research Resources.

REFERENCES

- Akaike H 1974. A new look at the statistical model identification. *IEEE Trans Automatic Control*, **19**: 716-23.
- Atwood LD, Wilson AF, Bailey-Wilson JE, Carruth JN, Elston RC 1995. On the distribution of the likelihood ratio test statistic for a mixture of two normal distributions. *Com in Stats*, **25(3)**: 733-40.
- Blot WJ, Devesa SS, Kneller RW, Fraumeni JF 1991. Rising incidence of adenocarcinoma of the esophagus and gastric cardia. *JAMA*, **265**: 1287-9.
- Bonney GE 1986. Regressive models for familial disease and other binary traits. *Biometrics*, **42**: 611-25.
- Bonney GE, Elston RC, Correa P, Tannenbaum SR, Zavala DE, Fonham E, Zarama G, Gordillo G, Cuellar C 1987. Genetic etiology of gastric carcinoma, II, Segregation analysis of gastric pH, nitrate, nitrite. *Genet Epid* **4**: 103-14.
- Box GEP, Cox DR 1964. An analysis of transformations (with discussion). *J R Statist Soc B*, **26**: 211-52.
- Buiatti E, Palli D, Bianchi S, Decarli A, Amadori D, Avellini C, Cipriani F, Cocco P, Giacosa A, Lorenzini L 1991. A case-control study of gastric

- cancer and diet in Italy. III, Risk patterns by histologic type. *Int J Cancer*, **48**: 369-74.
- Caldas C, Carneiro F, Lynch HT, Yokota J, Wiesner GL, Powell SM, Lewis FR, Huntsman DG, Pharoah PD, Jankowski JA, MacLeod P, Vogelsang H, Keller G, Park KG, Richards FM, Maher ER, Gayther SA, Oliveira C, Grehan N, Wight D, Seruca R, Roviello F, Ponder BA, Jackson CE 1999. Familial gastric cancer: overview and guidelines for management. *J Med Genet*, **36**: 873-80.
- Carneiro F, David L, Seruca R, Casted S, Nesland JM, Sobrinho-Simoes M 1993. Hyperplastic polyposis and diffuse carcinoma of stomach, A study of a family. *Cancer*, **72**: 323-9.
- Creagan ET, Fraumeni JF 1973. Familial gastric cancer and immunologic abnormalities. *Cancer*, **32**: 1325-31.
- Demerais FM, Elston RC 1981. A general transmission probability model for pedigree data. *Hum Hered*, **31**: 93-9.
- Fleischauer AT, Poole C, Arab L 2000. Garlic consumption and cancer prevention: meta-analyses of colorectal and stomach cancers. *Am J Clin Nutr*, **72**: 1047-52.
- Gar'kavtseva RF, Kazubskaja TP, Sel'chuk VI, Trubnikov VI, Gindilis VM 1987. Genetic analysis of predisposition to cancer of the stomach. *Vopr Onkol*, **33**: 17-23.
- Guilford P, Hopkins J, Harraway J, McLeod M, McLeod N, Harawira P, Taite H, Scoular R, Miller A, Reeve AE 1998. E-cadherin germline mutations in familial gastric cancer. *Nature*, **392**: 402-5.
- Hanai A, Fujimoto I 1982. Cancer incidence in Japan in 1975 and changes of epidemiological features for cancer in Osaka. *Nat Cancer Inst Monogr*, **62**: 3-7.
- Iida S, Akiyama Y, Ichikawa W, Yamashita T, Nomizu T, Nihei Z, Sugihara K, Yuasa Y 1999. Infrequent germ-line mutation of the E-cadherin in gene in Japanese familial gastric cancer kindreds. *Clin Cancer Res*, **5**: 1445-47.
- Inoue M, Tajima K, Hirose K, Hamajima N, Takezaki T, Kuroishi T, Tominaga S 1997. Epidemiological features of first-visit outpatients in Japan: comparison with general population and variation by sex, age, and season. *J Clin Epidemiol*, **50**: 69-77.
- Inoue M, Tajima K, Yamamura Y, Hirose K, Hamajima N, Kodera Y, Kito T, Tominaga S 1998. Family history and subsite of gastric cancer: data from a case-referent study in Japan. *Int J Cancer*, **76**: 801-5.
- Inoue M, Tajima K, Yamamura Y, Hamajima N, Hirose K, Nakamura S, Kodera Y, Kito T, Tominaga S 1999. Influence of habitual smoking on gastric cancer by histologic subtype. *Int J Cancer*, **81**: 39-43.
- Janerich DT, Mayne ST, Thompson, WD, Stark, AD, Fitzgerald EF, Jacobson HI 1990. Familial clustering of neural tube defects and gastric cancer. *Int J Epidemiol*, **19**: 516-21.
- Japanese Research Society for Gastric Cancer 1995. *Japanese classification of gastric carcinoma* (1st Engl. Ed.) Tokyo: Kanehara Karunaratne PM, Elston RC, Loman N, Olsson H, Ranstam 1998. Analysis of Swedish Male Breast Cancer Family Data: A Simple Way to Incorporate a Common Sibling Effect. *Genet Epidemiol*, **15**: 201-12.
- Kato Y, Kitagawa T, Nakamura K, Sugano H 1981. Changes in the histologic types of gastric carcinoma in Japan. *Cancer*, **48**: 2084-7.
- La Vecchia, C, Negri E, Franceschi S, Gentile A 1992. Family history and the risk of stomach and colorectal cancer. *Cancer*, **70**: 50-5.
- Lauren P 1965. The two histological main types of gastric carcinoma: diffuse and so-called intestinal-type carcinoma. *Acta Pathol Microbiol Immunol Scand*, **64**: 31-49.
- Lauren P, Nevalainen TJ 1993. Epidemiology of intestinal and diffuse types of gastric carcinoma. A time-trend study in Finland with comparison between studies from high- and low-risk areas. *Cancer*, **71**: 2926-33.
- Mecklin, JP, Nordling S, Saario I 1988. Carcinoma of the stomach and its heredity in young patients. *Scand J Gastroentero*, **23**: 307-11.
- Munoz SE, Ferraroni M, La Vecchia C, Decarli A 1997. Gastric cancer risk factors in subjects with family history. *Cancer Epidemiol Biomarkers*, **6**: 137-40.
- Oda T, Kanai Y, Oyama T, Yoshiura K, Shimoyama Y, Birchmeier W, Sugiura T, Hirohashi S 1994. E-cadherin gene mutations in human gastric carcinoma cell lines. *Proc Natl Acad Sci*, **91**: 1858-62.
- Powell J, McConkey CC 1990. Increasing incidence of adenocarcinoma of the gastric cardia and adjacent sites. *Brit J Cancer*, **62**: 440-3.
- S.A.G.E.: Statistical Analysis for Genetic Epidemiology, Release 3.1 1998. Computer program package available from the Department of Epidemiology and Biostatistics, Case Western Reserve University School of Medicine, Cleveland, Ohio.
- Self SG, Liang KY 1987. Asymptotic properties of maximum likelihood estimators and likelihood ratio test under non-standard conditions. *J Am Stat Assoc*, **82**: 605-10.
- Shinmura K, Yin W, Isogaki J, Saitoh K, Kanazawa K, Koda K, Yokota J, Kino I, Arai T, Sugimura H 1997. Stage-dependent evaluation of microsatellite instability in gastric carcinoma with familial clustering. *Cancer Epidemiol Biomarkers*, **6**: 693-7.
- Shinmura K, Tani M, Isogaki J, Wang Y, Sugimura H, Yokota J 1998. RER phenotype and its associated mutations in familial gastric cancer. *Carcinogenesis*, **19**: 247-51.
- Tredaniel J, Boffetta P, Buiatti E, Saracci R, Hirsch A 1997. Tobacco smoking and gastric cancer: review and meta-analysis. *Int J Cancer*, **72**: 565-73.
- Triantafyllidis JK, Kosmidis P, Kottaridis S 1993. Familial stomach cancer. *Am J Gastroenterol*, **88**: 1789-90.
- Young PC, Davis S 1988. Epidemiological characteristics of adenocarcinoma of the gastric cardia and distal stomach in the United States, 1973-1982. *Int J Epidemiol*, **17**: 293-7.
- Zanghieri G, Di Gregorio C, Sacchetti C, Fante R, Sassatelli R, Cannizzo G, Carriero A, Ponz de Leon M 1990. Familial occurrence of gastric cancer in the 2-year experience of a population based registry. *Cancer*, **66**: 2047-51.
- Zhao L, Blot WJ, Liu WD, Chang YS, Zhang JS, Hu YR, You WC, Xu GW, Fraumeni JF 1994. Familial predisposition to precancerous gastric lesions in a high-risk area of China. *Cancer Epidemiol Biomarkers*, **3**: 461-4.