

REPRINTED FROM:

MEDINFO 80

Proceedings of the Third World Conference
on Medical Informatics
Tokyo, September 29 – October 4, 1980

Edited by

DONALD A. B. LINDBERG

*Information Science Group,
School of Medicine,
University of Missouri,
Columbia, Missouri*

and

SHIGEKOTO KAIHARA

*Hospital Computer Centre,
University of Tokyo Hospital, Tokyo*

Part 1



1980

NORTH-HOLLAND PUBLISHING COMPANY – AMSTERDAM • NEW YORK • OXFORD

A STATISTICAL ANALYSIS OF DETECTING RISK FACTORS FOR STOMACH CANCER PATIENTS

Haruo YANAI⁽¹⁾, Yasuhiko YOSHIMOTO⁽²⁾, Hirofumi TAKAGI⁽³⁾,

Kazuho MAEDA⁽³⁾, Hideo KURITA⁽⁴⁾

(1) Department of Psychology, Chiba University, Chiba, Japan

(2) Radiation Effect Research Center, Hiroshima, Japan

(3) Department of Epidemiology, University of Tokyo, Tokyo, Japan

(4) Department of Public Health, Nagoya Municipal University, Nagoya, Japan

In order to elucidate agents relating to stomach cancer, we analyzed the data of 14 epidemiological items such as educational level, occupation, habits of foods and drinking and so on, using the data of 729 patients of stomach cancer and 671 controls. As results, it was suggested that occupational hazard, eating dose of boiled rice, speed of eating meal, habits of drinking as well as of smoking were exclusively the factors associated with agents developing stomach cancer. In addition, it was also suggested that constitutional or hereditary factors are suspected to be risk factors for stomach cancer especially in the younger age group.

INTRODUCTION

In Japan, it is well known that high incidence of stomach cancer has been observed in the last decade, and therefore various types of retrospective studies as well as of prospective studies are needed in order to elucidate agents relating to stomach cancer.

It should be noted, however, that it is usually the case in studies that have been done before that many items are analyzed separately so as to calculate the relative risk for each item without paying attention to correlations among the items used in the study. From the statistical point of view, analyses carried out separately for each item sometimes lead to the incorrect interpretation of the result, provided that correlations among the items are not negligible. Hence, it is generally admitted to apply some methods of multivariate analysis, which have been characterized to uncover the uncorrelated structures of the items.

In the meantime, we often encounter epidemiological data measured in terms of a nominal scale such as job classification. In such a case, a method of multivariate analysis called the Quantification II developed by C. Hayashi of a Japanese statistician (Hayashi, 1952) may duly be applied in order to discriminate a case group from the control group. In what follows, we try to give a brief sketch of the result which was obtained by applying the Quantification II to the epidemiological items of stomach cancer patients and controls.

MATERIALS USED IN THE ANALYSIS

Data used in the study are those based on the relatives of 2,051 patients with gastric carcinoma (male 1,366 and female 685), and those based on 1,619 patients with atrophic gastromucosa and 470 patients with

normal gastromucosa, most of whom visited the Research Institute of Aichi Cancer Center during 1967 through 1972. One of the authors Kurita interviewed with each of the patients and collected retrospectively all the data of epidemiological and clinical items from them.

Furthermore, obtained data of the case and control groups are classified into the four age groups in male case; i.e., 30-39, 40-49, 50-59 and 60-69. The method of Quantification II was performed to each of the four groups in male case. In addition, the same method was applied to the data of both sexes, in which all the age groups were combined.

We showed in Table 1 number of subjects used in the analysis.

Table 1. Number of Subjects used in the Analysis

Sex & age	Male				Female
	30-39	40-49	50-59	60-69	30-69
Patient group with stomach cancer	129	200	200	200	640
Healthy control group	200	200	179	92	750
Including Normal gastromucosa	100	66	18	7	280
Atrophic gastromucosa	100	134	161	85	470

Table 2. Assigned Weight in Each of Categories Calculated by Method of Quantification II

Item	Category	Sex		Male					Female
		Age		30-39	40-49	50-59	60-69	30-69	30-69
(1) Educational level	(a) Compulsory education			0.39	-0.28	-0.36	0.46	0.25	1.46
	(b) Highly educated			-0.30	0.28	0.78	-1.38	-0.36	-2.96
(2) Occupation	(a) Manager, Specialist			-0.76	-1.07	-0.50	-5.41	-1.25	-1.70
	(b) Shopmaster, Canvasser			0.48	0.41	0.89	-0.77	0.56	-1.73
	(c) Driver			0.36	0.59	-0.16	-0.29	0.41	
	(d) Farmer, Fisherman			0.24	-3.27	-3.83	-1.71	-2.34	-0.98
	(e) Craftman, Labourer			-3.24	2.34	2.80	1.07	1.54	3.48
	(f) House wife, others			-0.03	-0.98	-0.80	1.32	-0.65	-0.45
(3) Occupational hazard	(a) High temperature or dust exposure			2.29	2.87	2.16	0.31	2.13	-0.26
	(b) Somatic tiredness			1.17	-0.14	1.08	0.92	0.71	2.85
	(c) Psychogenic tension			-2.76	-2.24	-3.41	-2.65	-2.94	-4.40
	(d) Well condition			-4.12	-1.79	-0.37	-0.36	-1.35	-1.76
(4) Income per month	(a) Till 50,000 Yen			0.24	3.58	-1.56	-0.45	0.77	-0.57
	(b) Till 60,000 Yen			3.42	2.37	1.30	-2.23	0.88	2.11
	(c) Till 80,000 Yen			1.00	2.65	2.05	1.13	1.27	1.33
	(d) Till 100,000 Yen			-3.59	-2.16	-1.47	0.09	-1.96	-1.80
	(e) Till 120,000 Yen			-3.66	-0.22	-2.05	2.80	-0.85	-1.38
	(f) Till 150,000 Yen			1.67	-0.81	-2.28	-0.20	-0.94	-2.26
	(g) Over 150,000 Yen			2.12	-2.13	0.73	-0.96	0.24	-0.35
(5) Occupational duty hours	(a) Till 8 hours			1.89	-0.81	0.83	1.59	0.46	-3.60
	(b) Over 8 hours			-0.59	0.79	-0.21	-0.40	0.03	5.60
	(c) Irregular			-0.61	-2.59	-0.29	-1.48	-0.92	1.05
(6) Eating-dose of boiled rice	(a) Eaten bread often			2.99	0.87	-0.63	-3.18	-0.67	-2.34
	(b) Till 7 cups			-1.81	-0.53	-0.56	-1.05	-0.67	0.08
	(c) Over 7 cups			1.21	0.44	0.67	1.03	0.84	1.11
(7) Speed of eating meal	(a) Fast			-1.70	-0.94	0.77	-1.49	-0.97	-1.70
	(b) Moderate			1.35	0.03	-1.41	1.85	0.39	0.94
	(c) Slow			0.13	0.37	0.24	-0.25	0.24	0.73
(8) Eating hot meal	(a) Rarely			-0.98	-4.64	-0.19	-0.50	-2.31	1.16
	(b) Sometimes			0.74	1.48	0.08	0.56	0.37	-1.80
	(c) Very frequently			-0.20	-0.02	-0.04	-0.31	0.01	1.03
(9) Drinking coffee	(a) Rarely			0.68	1.73	-0.29	-0.95	0.37	-0.46
	(b) Sometimes			-1.00	-1.46	0.31	0.90	-0.73	1.28
	(c) Every day			-0.19	-1.50	0.45	2.17	-0.06	0.43
(10) Taking milk	(a) Rarely			0.48	-1.17	0.39	-0.42	-0.17	0.62
	(b) Sometimes			-1.02	-0.18	-0.10	0.46	-0.41	-0.85
	(c) Every day			0.81	1.60	-0.55	0.24	0.80	-0.27

Item	Sex		Male					Female
	Category	Age	30-39	40-49	50-59	60-69	30-69	30-69
(11) Eating highly salted foods	(a) Rarely		-0.28	-0.79	0.63	0.02	0.09	0.22
	(b) Sometimes		-1.07	-0.49	-2.50	-2.60	-1.21	-2.42
	(c) Every day		0.26	0.36	0.61	0.90	0.33	0.88
(12) Habit of drinking alcohol (Japanese Sake)	(a) No habit		0.09	0.54	0.50	0.77	0.46	-0.14
	(b) Sometimes		-1.14	-1.56	-1.76	-0.07	-1.27	1.05
	(c) Bier daily		0.18	-1.94	1.11	-1.45	-0.85	0.95
	(d) 180 ml a day		-1.53	0.16	-1.73	-0.88	-0.50	2.13
	(e) 360 ml a day		2.43	-1.18	0.13	0.94	0.17	1.86
	(f) 540 ml a day		-0.32	3.61	3.01	0.73	2.02	1.79
	(g) Over 540 ml a day		5.25	5.42	3.05	-5.77	0.60	0.89
(13) Habit of smoking (No. of cigarettes smoked a day × Smoking-year)	(a) None		-2.00	-2.24	-4.65	-2.46	-2.99	-0.32
	(b) 100		0.87	-1.87	-2.80	-1.81	-1.64	1.63
	(c) 200		-1.73	0.78	0.41	-3.96	-1.22	1.04
	(d) 400		1.04	-1.95	-1.37	-0.60	-0.70	2.36
	(e) 600		-0.43	0.78	-1.65	0.33	-0.24	1.31
	(f) 800		4.47	2.64	3.12	-0.96	2.15	2.32
	(g) Over		7.04	3.23	3.34	2.24	3.84	1.09
(14) History of cancer among parents and siblings	(a) Two or over		4.03	-0.09	-1.91	2.50	-0.14	-0.44
	(b) One stomach cancer and one another cancer		1.78	4.95	-3.17	5.62	1.86	-1.55
	(c) One cancer except stomach		-0.21	-0.14	-0.01	0.00	-0.12	0.46
	(d) One stomach cancer		-0.53	0.29	0.71	-1.78	0.09	-1.46
	(e) None		1.28	-0.26	0.60	-0.40	0.30	-0.63

RESULT

We showed in Table 2 weights assigned to each category of all the 14 items. High positive weight of a category implies that the category is suspected to be a risk factor for developing stomach cancer, taking the intercorrelations between any two pairs of categories into consideration. Furthermore, we showed in Table 3 the results of classification, in which rates of correct diagnosis between the stomach cancer patients and the controls are given for each of the four age groups and the totals of the four age groups in each sex together with the theoretical rate of correct diagnosis determined as a function of Mahalanobis generalized distance. The average rates of correct diagnosis when the both groups are combined amount to 74.2%, 75.0%, 74.5% and 77.4% in the respective age groups. It should be noted, however, that the rate of correct diagnosis decrease when all the age groups are combined.

From Table 2, it is seen that occupational hazard, eating dose of boiled rice, speed of eating meal, habits of drinking as well as of smoking are suggested exclusively to be the factors associated with becoming high risk groups tending to develop stomach cancer.

Table 3. Accurately Diagnosed Rate (%)
by the Method of Quantification II

Sex	Age	Stomach cancer group	Control group	On the whole	When computed theoretically
Male	30-39	76.0%	73.5%	74.2%	74.1%
	40-49	73.0%	77.0%	75.0%	75.0%
	50-59	75.0%	73.3%	74.1%	72.4%
	60-69	82.5%	66.3%	77.4%	71.9%
	Total	71.1%	68.3%	69.6%	68.6%
Female	Total	61.9%	69.4%	65.7%	67.9%

On scrutinizing the weights for different age groups, we see that severe intake of habitual drinks together with smoking prove to be high risk factors especially in the younger age group, while severe intake of boiled rice as well as of coffee will prove to be high risk factors in the older age group.

As for the sex difference, remarkable difference was not found except for the fact that occupational duty is proved to be a high risk factor in the female case.

Next, we showed in Fig. 2 incidence rates for the following items such as No. 6, 7, 11, 12 and 13 for the groups of cancer patients and controls together with the D group of stomach cancer, which is characterized as the group with high discriminant scores (see to Fig. 1). On scrutinizing Fig. 2, it is readily suggested that excessive eating of boiled rice, eating meal too fast, eating highly saled foods, and excessive drinking and smoking are closely connected with high risk factors of stomach cancer.

From the clinical data approached from differences of lesion into stomach, it was suggested that stomach cancer patients in the D group are in a worse condition as compared with patients other than the D group.

Next, we showed in Fig. 3, the average discriminant scores of the two groups, which are obtained by applying the Quantification II to the groups of both sexes, in each of which all the age groups are combined.

From this Figure, we see that the average score of discriminant scores increases in each sex as the calendar year proceeds on.

In view of this result and considering the fact that mortality rates of stomach cancer are almost the same below the age of thirty-five and they differ remarkably between the sexes over the age of thirty-five in

Fig. 2. Frequency Distribution of Each Category in Each Item (Male, 50-59 years old)
See to Table 2 for the name of each category

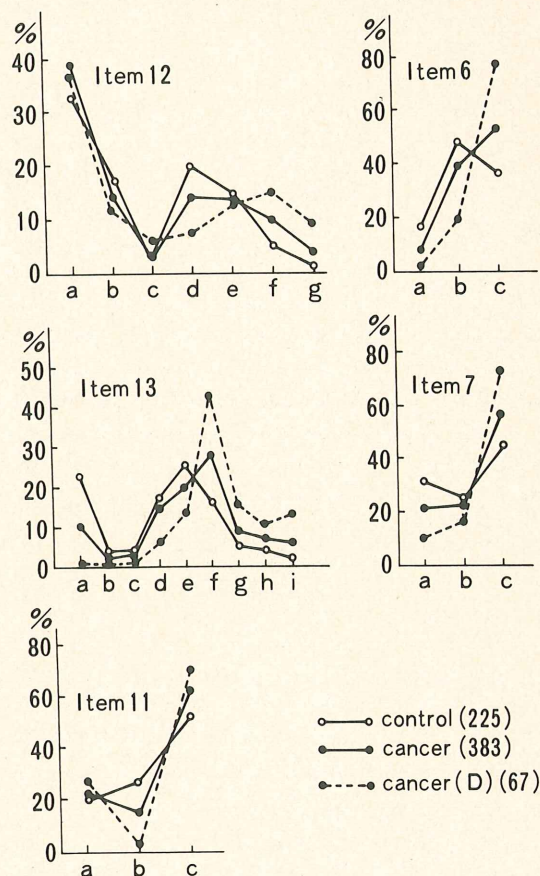


Fig. 3. Mean Discriminant Score in Sex and Age Differences

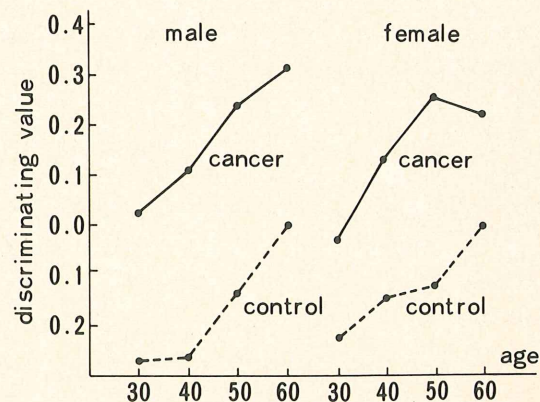
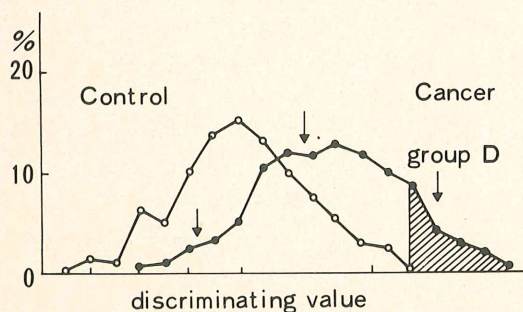


Fig. 1. Distribution of Discriminant Scores in the Two Groups of Stomach Cancer and Control



the world, it was suggested that environmental factors will be main risk factors in the high age groups, while constitutional or hereditary factors are also risk factors together with some of environmental factors as described above. Kurita (1976) pointed out that Blood type A might be a risk factor in the younger age group especially below the age thirty.

From the result described above, clinico-epidemiological factors of stomach cancer may be summarized in the Table 4. (Table 4 is mainly due to Kurita.)

DISCUSSION

Beside our study, a number of studies have been done in order to elucidate unknown etiology of stomach cancer. One of them is the case control study performed to the Japanese living in Hawaii Island, (Haenszel, 1972) in which eating salted fish and drinking coffee everyday are suggested to be risk factors for stomach cancer and taking milk everyday is suspected to function as an inhibitive factor for stomach cancer, as opposed to the negative weights assigned to the category (c) of the item No. 10.

From the statistical point of view, we should keep it in mind that risk factors should be evaluated not separately for each item but that they should be evaluated as a set of items with mutual high correlation. Let us consider the weights assigned to the item of "taking milk, in which positive weights are assigned to the category everyday" in all the age groups of male. It should be noted, however, that this item is closely related to the item "eating-dose of boiled rice" in that people who take milk everyday prefer bread to rice. In view of this, it does not necessarily follow from Table 2 that taking milk everyday is not a risk factor by itself.

However, it should be noted that our study has at least one weak point in that time trend has not been taken into consideration in the analysis. Employing the methods of exponential Weibul distribution as well as Cox's method may duly develop the problems as mentioned above. (Dryer, 1975, Cox, 1972)

Finally, we should mention here that a possibility can be suggested in making a checklist aimed at the general use for prevention of stomach cancer.

Table 4. Summarized Table about Clinico-Epidemiological Findings of Stomach Cancer, approached from Differences of Lesion into Stomach

Cancer group, divided from site of lesion	Age group, suffering higher	Sex difference	Predominant morphological pattern	Hypothetical incubation period	Most important aetiology assumed
Cancer of cardia	All age groups through young to old	High in male	Semi-protruded type with invasive tendency	Relatively short	Severe intake of irritative or habitual drinks and foods, together with smoking (constitutional factor in female)
Cancer of gastric body	Young under 49 age of year	Dominant in female	Depressed type (with excavation)	Short	Constitutional or hereditary factor (especially, anatomical structure of stomach in male, sex hormonal balance in female). Intake of irritative drinks and foods.
Cancer of gastric angle		Nearly same	Excavated type		
Cancer of pyloric antrum	Old over 50 age of year	High in male	Protruded type with crater	Long	Weak but ceaseless irritative outer-factors for gastromucosa throughout life such as great intake of rice with salted foods. Physical aging phenomenon (especially, intestinal metaplasia of gastromucosa)