



Journal of Epidemiology

Contents lists available at ScienceDirect

Journal of Epidemiology

journal homepage: <http://www.journals.elsevier.com/journal-of-epidemiology/>

Original Article

Characteristics and prognosis of Japanese male and female lung cancer patients: The BioBank Japan Project



Koshi Nakamura^a, Shigekazu Ukawa^a, Emiko Okada^a, Makoto Hirata^b, Akiko Nagai^c, Zentarō Yamagata^d, Toshiharu Ninomiya^e, Kaori Muto^c, Yutaka Kiyohara^f, Koichi Matsuda^{g,h}, Yoichiro Kamataniⁱ, Michiaki Kubo^j, Yusuke Nakamura^{g,k}, BioBank Japan Cooperative Hospital Group, Akiko Tamakoshi^{a,*}

^a Department of Public Health, Hokkaido University Graduate School of Medicine, Sapporo, Japan

^b Laboratory of Genome Technology, Institute of Medical Science, The University of Tokyo, Tokyo, Japan

^c Department of Public Policy, Institute of Medical Science, The University of Tokyo, Tokyo, Japan

^d Department of Health Sciences, University of Yamanashi, Yamanashi, Japan

^e Department of Epidemiology and Public Health, Graduate School of Medical Sciences, Kyushu University, Fukuoka, Japan

^f Hisayama Research Institute for Lifestyle Diseases, Fukuoka, Japan

^g Laboratory of Molecular Medicine, Institute of Medical Science, The University of Tokyo, Tokyo, Japan

^h Laboratory of Clinical Genome Sequencing, Graduate School of Frontier Sciences, The University of Tokyo, Tokyo, Japan

ⁱ Laboratory for Statistical Analysis, RIKEN Center for Integrative Medical Sciences, Yokohama, Japan

^j RIKEN Center for Integrative Medical Sciences, Yokohama, Japan

^k Section of Hematology/Oncology, Department of Medicine, The University of Chicago, Chicago, USA

ARTICLE INFO

Article history:

Received 18 October 2016

Accepted 11 December 2016

Available online 13 February 2017

Keywords:

Lung cancer

Smoking

Stage

Histological type

Mortality

ABSTRACT

Background: In Japanese males and females, lung cancer is currently the second and fourth most common type of cancer, and the first and second leading cause of cancer-related deaths, respectively.

Methods: Of all Japanese male and female lung cancer patients aged ≥ 20 years whom the BioBank Japan Project originally enrolled between 2003 and 2008, 764 males and 415 females were registered within 90 days after their diagnosis. We described the lifestyle and clinical characteristics of these patients at study entry. Furthermore, we examined the effect of these characteristics on all-cause mortality.

Results: In the lung cancer patients registered within 90 days, the frequencies of occult or stage 0, stage I, II, III and IV were 0.4%, 55.8%, 10.8%, 22.0% and 11.0% for males and 0.3%, 62.4%, 9.9%, 17.1% and 10.2% for females, respectively. The proportions of histological types in males and females were 56.3% and 82.4% for adenocarcinoma, 26.9% and 8.2% for squamous cell carcinoma, 4.5% and 1.5% for large cell carcinoma, 7.7% and 4.1% for small cell carcinoma and 4.6% and 3.8% for others, respectively. Among 1120 participants who registered within 90 days, 572 participants died during 5811 person-years of follow-up. Low body mass index, ever smoker, more advanced stage, squamous cell or small cell carcinoma and high serum carcinoembryonic antigen level at study entry were crudely associated with an increased risk of all-cause mortality after adjustment for age.

Conclusions: This study showed the association of several lifestyle and clinical characteristics with all-cause mortality in lung cancer patients.

© 2017 The Authors. Publishing services by Elsevier B.V. on behalf of The Japan Epidemiological Association. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Introduction

Lung cancer has increased in incidence over the past several decades, and is a major health problem throughout the world.^{1,2} In Japan, owing to a similar increasing trend,³ lung cancer is currently the second and fourth most common type of cancer in

* Corresponding author. Department of Public Health, Hokkaido University Graduate School of Medicine, North 15, West 7, Kita-ku, Sapporo City, Hokkaido 060-8638, Japan.

E-mail address: tamaa@med.hokudai.ac.jp (A. Tamakoshi).

Peer review under the responsibility of The Japan Epidemiological Association.

males and females, respectively.⁴ Smoking is a well-known strong risk factor for the development of lung cancer.⁵ However, the frequency of this unfavorable habit is still higher in Japanese males compared with males in the Western countries.⁶ Thus, the prevalence of lung cancer may have a possibility to remain high in Japanese males. Meanwhile, different cause-effect relationships exist between smoking and the onset of lung cancer when taking histological type into consideration; smoking has a greater effect on the incidence of squamous cell carcinoma and small cell carcinoma, as compared to other histological types.⁵

Lung cancer is a fatal malignancy.^{1,2} The latest Japanese National Vital Statistics reported a total of 73,396 annual deaths due to cancer of the lung, trachea and bronchus (52,505 deaths in males and 20,891 deaths in females).⁷ As a result, lung cancer is the leading cause of cancer-related death in Japanese males and the second leading cause of cancer-related death in Japanese females, respectively.⁷ This poor prognosis is mainly due to the unfavorable fact that lung cancer is often diagnosed at an advanced stage.^{8–11} The frequency of late diagnosis of lung cancer also varies among histological types.^{10,11}

The BioBank Japan (BBJ) Project is a large patient-based biobank designed to implement personalized medicine for common diseases such as cancer and cardiovascular disease. Since the morbidity and mortality of lung cancer are different between males and females,^{4,7} it is worthwhile to examine clinical characteristics and mortality of lung cancer patients by sex. In this study, we attempted to investigate lifestyle and clinical characteristics of Japanese lung cancer patients and their prognosis by sex, using the BBJ Project database.

Participants and methods

Study design and population

Details of the study design and protocol of the BBJ Project were described elsewhere.^{12,13} Briefly, the project enrolled patients with any of 47 common target diseases including lung cancer between June 2003 and March 2008, from 66 hospitals consisting of 12 cooperating medical institutions throughout Japan. This project collected clinical information and biological samples from participants annually until March 2013, regardless of whether the patients were newly diagnosed or treated cases. This project also followed up participants who had at least one of 32 diseases until 2014. The study protocol of the BBJ Project was approved by the research ethics committees of the Institute of Medical Science, the University of Tokyo, RIKEN Yokohama Institute and the 12 cooperating medical institutions. Written informed consent was obtained from all participants.

The BBJ Project originally enrolled 3779 lung cancer patients (2428 males and 1351 females). Of the 3779 participants, 197 were excluded due to missing data on the time from their diagnosis of lung cancer to study entry. Accordingly, we described the overview of lifestyle factors and clinical profiles of the remaining 3582 lung cancer patients (2299 males and 1283 females). To describe the characteristics of possible newly diagnosed lung cancer patients, we selected 1179 participants (764 males and 415 females) who were registered in the BBJ Project within 90 days after their diagnosis of lung cancer. Of these 1179 participants, 59 were excluded due to refusal to participate in the follow-up survey ($n = 57$) or loss of follow-up ($n = 2$). Thus, 1120 participants (728 male and 392 female) were included in the survival analyses to examine the prognosis and its risk factors in lung cancer patients.

Data collection

Clinical data were collected at study entry via interviews and medical records.¹⁴ The data included age at study entry, time from the diagnosis of lung cancer to study entry, height, weight, smoking habits, alcohol intake, past history, family history and clinical information at study entry including laboratory examination data (e.g., blood chemical markers and imaging data). Body mass index was calculated as weight in kilograms divided by height in meters squared. The blood chemical marker used in this report was serum carcinoembryonic antigen (CEA) level. The stage of lung cancer was classified according to the Japanese Classification of Lung Cancer, the 6th edition (2003). In this report, stage was primarily determined by the findings of pathologic stage, while missing data of pathologic stage were complemented with the findings of clinical stage. The histological type of lung cancer was determined primarily based on the findings in surgically resected tissues, while missing histological data were complemented with the findings from biopsy or cytological samples.

Follow-up survey

Follow-up survival survey was conducted until 2014 to determine whether the status of each participant was alive, relocated, unidentified, or dead by confirming their residence cards.¹⁵ The new addresses of participants who relocated were also recorded in the next survival survey. The date of death was also recorded for deceased participants. The outcome assessed in this study was death due to all causes. Details of the follow-up survey were described elsewhere.^{13,15}

Statistical analysis

We performed a descriptive analysis of lifestyle and clinical characteristics of lung cancer patients for males and females separately, because smoking habits influences the proportion of histological type and cancer stage at diagnosis of lung cancer.^{5,10,11} The characteristics of interest were age at study entry, body mass index, smoking and drinking habits, comorbidities of chronic obstructive pulmonary disease and interstitial pneumonia/pulmonary fibrosis, family history of lung cancer, stage, histological type, and serum CEA level. The frequency of each characteristic was calculated in male and female study participants after excluding those with missing data. This analysis was performed in all participants and participants who were registered in the BBJ Project within 90 days after their diagnosis of lung cancer. The same analysis was conducted after stratifying the participants by age at study entry (20–59 versus ≥ 60 years old) or histological type.

We evaluated the prognosis of lung cancer in participants registered within 90 days after their diagnosis for male and female separately and for males and females combined. We calculated the 5-year relative survival rate by dividing the 5-year cumulative survival rate by the sex- and age-adjusted expected survival rate in our eligible lung cancer patients. The 5-year cumulative survival rate was calculated, using the Kaplan–Meier method. The expected survival rate was calculated, using a survival-rate table of a reference Japanese cohort from the Cancer Registry and Statistics, Cancer Information Service, National Cancer Center, Japan,¹⁶ which was based on sex- and age-specific mortality rates and Gompertz–Makeham's law in abridged life tables, published annually by the Statistics and Information Department of the Ministry of Health, Labour and Welfare, Japan.¹⁷

We examined the impact of lifestyle and clinical characteristics on all-cause mortality in lung cancer patients registered within 90 days after their diagnosis, with males and females combined. We used a Cox proportional hazards model to estimate the hazard ratios and 95% confidence intervals (CIs) for mortality for each characteristic category, with one category set as the reference. The model was stratified by sex and institution to account for variability in baseline hazards between sexes and among the institutions. The model incorporated age at study entry (years as a continuous variable) and the entry year (2003, 2004, 2005, 2006, 2007 and 2008, setting year 2003 as the reference) in addition to each characteristic. We examined the association between BMI and mortality using a continuous variable, as well as the categorical variable. In this analyses, we excluded the participants with missing data or the group of small number of participants. Similar analyses were conducted only for participants with adenocarcinoma, the most common histological type of lung cancer.^{8–11}

All statistical analyses were performed using SAS version 9.4 (SAS Institute Inc., Cary, NC, USA). All p values were two-tailed, and the significance level was set at $p < 0.05$.

Results

Characteristics of all participants with lung cancer

Among 2299 male and 1283 female lung cancer patients aged ≥ 20 years at the enrollment, the mean (standard deviation) age at study entry was 67.5 (9.6) years for males and 65.9 (9.8) years for females. The distribution of the participants according to age at study entry is illustrated in Fig. 1; age distribution was slightly negatively skewed in both sexes, with the peak in the 70–74 year age group for males and the 65–69 year age group for females. The distribution of all the male and female participants, according to the time from the diagnosis to study entry was as follows: 33.2% and 32.3% for ≤ 90 days, 24.0% and 19.9% for 91–365 days, 29.6% and 33.0% for 1–4 years, 10.2% and 11.2% for 5–9 years, 2.3% and 2.8% for 10–14 years, 0.5% and 0.7% for 15–19 years, 0.2% and 0.1% for ≥ 20 years. Tables 1 and 2 summarize other characteristics of the male and female lung cancer patients, respectively. Male and female participants who had a history of smoking at study entry (current or former smoker) accounted for nearly 90% and 30% of known cases, respectively. Over 10% of male and female participants had a family history of lung cancer. Among the participants with stage information, over half of males and approximately two-thirds of females were classified as stage I at study entry, while nearly one-tenth were diagnosed as stage IV. In both sexes, adenocarcinoma was the most common histological type followed by squamous cell carcinoma, small cell carcinoma, and large cell carcinoma, although the proportion of adenocarcinoma was higher in females compared with males. Over one-fourth of males and one-fifth of females had a serum CEA of 5 ng/mL or higher, respectively.

Characteristics of participants registered within 90 days year after diagnosis

Of all lung cancer patients aged ≥ 20 years, 764 male and 415 female participants were registered within 90 days after the diagnosis of lung cancer. The mean age (standard deviation) at study entry was 66.0 (9.6) years for males and 64.7 (9.4) years for females. Age distribution of these participants were shifted to younger as compared to the overall participants for both sexes (Fig. 1). In both sexes, the frequency of stage IV was slightly higher and the proportion of serum CEA ≥ 5 ng/mL was higher in those registered within 90 days compared with the overall participants (Tables 1 and 2). In females, the frequency of ever smoker was

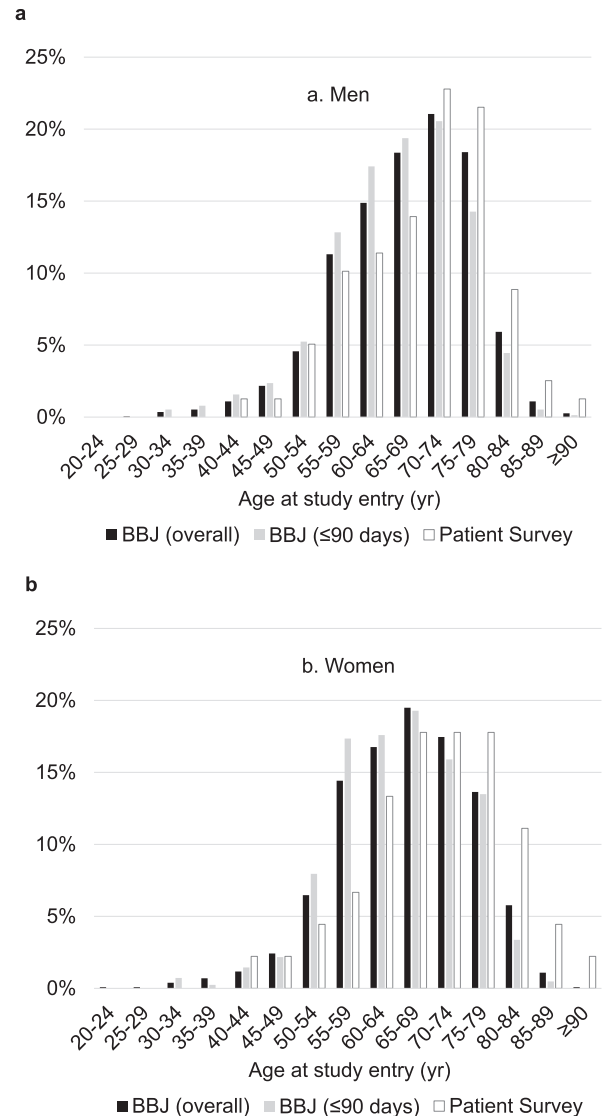


Fig. 1. Distribution of the (a) male and (b) female lung cancer patients according to age at study entry. Bars colored in black, gray and white represent patients in the BioBank Japan (BBJ) Project: (■) the overall participants and (▒) those registered within 90 days after their diagnosis and (□) patients in the Patient Survey in Japan, 2005,¹⁸ respectively.

higher in those registered within 90 days compared with the overall participants (Table 2). Other characteristics were similar for those registered within 90 days and the overall participants in both sexes.

When these male and female participants were stratified by age at study entry, there were several differences between the younger (aged 59 years or less) and older (aged 60 years or more) groups (Tables 1 and 2). In males, adenocarcinoma was more predominant in the younger group compared with the older group, although adenocarcinoma was the most common histological type in both age groups. In females, the frequencies of ever smoker, alcohol drinker and a positive family history of lung cancer was higher in the younger group than in the older group, while the frequencies of obesity was lower in the younger, than older, age group. In both sexes, the frequency of participants with serum CEA ≥ 5 ng/mL was higher in the older group than in the younger group.

When these participants were stratified by histological type, we found the differences in several characteristics (Tables 3 and 4). The frequency of ever smoker was predominant in all histological types

Table 1

Characteristics of the 2299 overall male lung cancer patients in the BioBank Japan Project. Data are also presented specifically for the 764 male lung cancer patients registered within 90 days after their diagnosis and those grouped according to age at study entry.

	Overall participants		Those registered within 90 days after diagnosis					
	Total (n = 2299)		Total (n = 764)		Age at study entry			
	n	%	n	%	20–59 yr (n = 178)		≥60 yr (n = 586)	
Body mass index								
<18.5 kg/m ²	243	11.0	79	10.5	18	10.2	61	10.6
18.5–24.9 kg/m ²	1605	72.3	545	72.4	129	73.3	416	72.1
≥25.0 kg/m ²	371	16.7	129	17.1	29	16.5	100	17.3
Unknown	80		11		2		9	
Smoking habit								
Never	267	11.7	77	10.1	16	9.0	61	10.4
Ever (current/former)	2017	88.3	686	89.9	162	91.0	524	89.6
Unknown	15		1		0		1	
Alcohol drinking habit								
Never	603	26.5	198	26.0	42	23.7	156	26.7
Ever (current/former)	1669	73.5	563	74.0	135	76.3	428	73.3
Unknown	27		3		1		2	
History of COPD								
Absence	2214	96.3	731	95.7	173	97.2	558	95.2
Presence	85	3.7	33	4.3	5	2.8	28	4.8
History of IP/PF								
Absence	2246	97.7	745	97.5	175	98.3	570	97.3
Presence	53	2.3	19	2.5	3	1.7	16	2.7
Family history of lung cancer								
Absence	1990	86.6	654	85.6	154	86.5	500	85.3
Presence	309	13.4	110	14.4	24	13.5	86	14.7
Stage								
Occult	10	0.6	1	0.2	0	0.0	1	0.2
0	8	0.5	1	0.2	1	0.8	0	0.0
IA	565	34.4	178	33.2	44	33.1	134	33.3
IB	349	21.2	121	22.6	20	15.0	101	25.1
IIA	47	2.9	17	3.2	5	3.8	12	3.0
IIB	143	8.7	41	7.6	12	9.0	29	7.2
IIIA	202	12.3	57	10.6	16	12.0	41	10.2
IIIB	172	10.5	61	11.4	19	14.3	42	10.4
IV	147	8.9	59	11.0	16	12.0	43	10.7
Unknown	656		228		45		183	
Histological type								
Adenocarcinoma	1141	55.5	388	56.3	99	60.4	289	55.0
Squamous cell carcinoma	587	28.6	185	26.9	38	23.2	147	28.0
Large cell carcinoma	80	3.9	31	4.5	12	7.3	19	3.6
Small cell carcinoma	148	7.2	53	7.7	10	6.1	43	8.2
Others	99	4.8	32	4.6	5	3.0	27	5.1
Unknown	244		75		14		61	
Serum CEA level								
<5 ng/mL	1547	74.3	479	70.3	116	72.5	363	69.7
≥5 ng/mL	535	25.7	202	29.7	44	27.5	158	30.3
Unknown	217		83		18		65	

Abbreviations: CEA, carcinoembryonic antigen; COPD, chronic obstructive pulmonary disease; IP/PF, interstitial pneumonia/pulmonary fibrosis.

except for adenocarcinoma in females. In males, small cell carcinoma was detected at more advanced stages compared with the other types, while in females, squamous cell carcinoma and small cell carcinoma was detected at more advanced stages compared with the remaining two types. In both sexes, adenocarcinoma and large cell carcinoma were distributed at younger age levels, as compared to squamous cell carcinoma or small cell carcinoma. Furthermore, participants with small cell carcinoma had the highest frequency of serum CEA ≥5 ng/mL among all histological types.

Prognosis of participants registered within 90 days year after diagnosis

Among the 1120 participants involved in the survival survey, the mean (standard deviation) follow-up period was 5.2 (3.4) years in a total of 5811 person-years of follow-up. During the follow-up

period, 572 deaths were identified in the total period and 471 were identified during the 5 years of follow-up. Consequently, the 5-year cumulative survival rate was 57.1% (95% CI, 54.2–60.0%). The 5-year relative survival rate was 62.8% (95% CI, 59.5–66.0%). In males, the 5-year cumulative survival rate was 49.2% (95% CI, 45.5–52.9%), while the 5-year relative survival rate was 55.4% (95% CI, 51.3–59.6%). In females, the 5-year cumulative survival rate was 71.9% (95% CI, 67.4–76.4%), while the 5-year relative survival rate was 75.6% (95% CI, 70.9–80.4%).

Table 5 shows which clinical characteristics will affect all-cause mortality in male and female-combined participants. The risk of all-cause mortality increased with increasing age at study. In addition, low body mass index, and ever smoker, more advanced stage, squamous cell or small cell carcinoma, and elevated serum CEA level were crudely associated with an increased risk of all-cause mortality after adjustment for age at study entry and entry year (Table 5).

Table 2

Characteristics of the 1283 overall female lung cancer patients in the BioBank Japan Project. Data are also presented specifically for the 415 female lung cancer patients registered within 90 days after their diagnosis and those grouped according to age at study entry.

	Overall participants		Those registered within 90 days after diagnosis					
	Total (n = 1283)		Total (n = 415)		Age at study entry			
	n	%	n	%	20–59 yr (n = 124)		≥60 yr (n = 291)	
				n	%	n	%	
Body mass index								
<18.5 kg/m ²	158	12.7	52	12.8	16	12.9	36	12.8
18.5–24.9 kg/m ²	900	72.2	287	70.7	94	75.8	193	68.4
≥25.0 kg/m ²	189	15.2	67	16.5	14	11.3	53	18.8
Unknown	36		9		0		9	
Smoking habit								
Never	942	74.1	281	68.2	74	59.7	207	71.9
Ever (current/former)	330	25.9	131	31.8	50	40.3	81	28.1
Unknown	11		3		0		3	
Alcohol drinking habit								
Never	887	70.0	281	68.5	62	50.0	219	76.6
Ever (current/former)	381	30.0	129	31.5	62	50.0	67	23.4
Unknown	15		5		0		5	
History of COPD								
Absence	1274	99.3	413	99.5	124	100.0	289	99.3
Presence	9	0.7	2	0.5	0	0.0	2	0.7
History of IP/PF								
Absence	1270	99.0	411	99.0	123	99.2	288	99.0
Presence	13	1.0	4	1.0	1	0.8	3	1.0
Family history of lung cancer								
Absence	1093	85.2	352	84.8	93	75.0	259	89.0
Presence	190	14.8	63	15.2	31	25.0	32	11.0
Stage								
Occult	5	0.5	1	0.3	0	0.0	1	0.5
0	1	0.1	0	0.0	0	0.0	0	0.0
IA	459	48.5	139	43.2	41	41.0	98	44.1
IB	178	18.8	62	19.3	19	19.0	43	19.4
IIA	24	2.5	9	2.8	4	4.0	5	2.3
IB	52	5.5	23	7.1	7	7.0	16	7.2
IIIA	80	8.5	32	9.9	12	12.0	20	9.0
IIIB	67	7.1	23	7.1	9	9.0	14	6.3
IV	80	8.5	33	10.2	8	8.0	25	11.3
Unknown	337		93		24		69	
Histological type								
Adenocarcinoma	1006	85.9	322	82.4	98	82.4	224	82.4
Squamous cell carcinoma	80	6.8	32	8.2	9	7.6	23	8.5
Large cell carcinoma	11	0.9	6	1.5	2	1.7	4	1.5
Small cell carcinoma	34	2.9	16	4.1	4	3.4	12	4.4
Others	40	3.4	15	3.8	6	5.0	9	3.3
Unknown	112		24		5		19	
Serum CEA level								
<5 ng/mL	954	79.2	286	73.3	91	79.1	195	70.9
≥5 ng/mL	251	20.8	104	26.7	24	20.9	80	29.1
Unknown	78		25		9		16	

Abbreviations: CEA, carcinoembryonic antigen; COPD, chronic obstructive pulmonary disease; IP/PF, interstitial pneumonia/pulmonary fibrosis.

When we focused on the participants with adenocarcinoma, similar relationships between clinical characteristics and all-cause mortality were observed (Table 5).

Discussion

In this report, we described the lifestyle and clinical characteristics of Japanese male and female lung cancer patients involved in the BBJ Project. As compared to the results of lung cancer patients in the Patient Survey in Japan, 2005 (Fig. 1),¹⁸ our registry contained a greater proportion of younger age distribution in both sexes. In both sexes, stage and serum CEA level were worse in those registered within 90 days compared with the overall participants. As anticipated from the evidence that smoking is a strong risk factor for lung cancer,⁵ our study population had a much higher frequency of ever smoker than a general Japanese population for both sexes.¹⁹

A report of the Japanese Lung Cancer (JLC) Registry showed that the frequency of each stage was 0.3% for stage 0, 29.3% for stage IA,

15.5% for stage IB, 1.4% for stage IIA, 6.3% for stage IIB, 11.7% for stage IIIA, 14.5% for stage IIIB, and 20.9% for stage IV in newly diagnosed lung cancer patients registered in 2002 throughout Japan.¹⁰ Another Japanese registry, which was comprised of patients at two major hospitals in Tokushima prefecture between 1999 and 2009, reported that the frequency was 22% for stage IA, 10% for stage IB, 1% for stage IIA, 5% for stage IIB, 11% for stage IIIA, 17% for stage IIIB, and 31% for stage IV.¹¹ Our participants, even those registered within 90 days after diagnosis, had less advanced stages of lung cancer, compared with these two registries. One possible explanation may be the difference in the definition of cancer stages. Stage used in this report was determined using either pathologic stage or clinical stage, whereas two registries used clinical stages. Another explanation for this difference may be the differences in participating hospitals among the registries.

The JLC Registry reported that the proportion of histological types of lung cancer was 57.5% for adenocarcinoma, 26.1% for squamous cell carcinoma, 0.8% for adeno-squamous cell

Table 3

Characteristics of the male lung cancer patients registered in the BioBank Japan Project within 90 days after their diagnosis, grouped according to histological type.

	Histological type							
	Adenocarcinoma (n = 388)		Squamous cell carcinoma (n = 185)		Large cell carcinoma (n = 31)		Small cell carcinoma (n = 53)	
	n	%	n	%	n	%	n	%
Age at study entry								
20–29 yr	0	0.0	0	0.0	0	0.0	0	0.0
30–39 yr	5	1.3	2	1.1	0	0.0	0	0.0
40–49 yr	20	5.2	5	2.7	2	6.5	1	1.9
50–59 yr	74	19.1	31	16.8	10	32.3	9	17.0
60–69 yr	157	40.5	62	33.5	9	29.0	18	34.0
70–79 yr	114	29.4	75	40.5	10	32.3	23	43.4
80–89 yr	18	4.6	10	5.4	0	0.0	2	3.8
≥90 yr	0	0.0	0	0.0	0	0.0	0	0.0
Body mass index								
<18.5 kg/m ²	31	8.1	19	10.4	3	9.7	6	11.8
18.5–24.9 kg/m ²	285	74.6	130	71.0	25	80.6	35	68.6
≥25.0 kg/m ²	66	17.3	34	18.6	3	9.7	10	19.6
Unknown	6		2		0		2	
Smoking habit								
Never	57	14.7	12	6.5	0	0.0	1	1.9
Ever (current/former)	331	85.3	173	93.5	31	100.0	51	98.1
Unknown	0		0		0		1	
Alcohol drinking habit								
Never	88	22.7	55	29.9	12	38.7	15	28.8
Ever (current/former)	299	77.3	129	70.1	19	61.3	37	71.2
Unknown	1		1		0		1	
History of COPD								
Absence	380	97.9	172	93.0	28	90.3	49	92.5
Presence	8	2.1	13	7.0	3	9.7	4	7.5
History of IP/PF								
Absence	379	97.7	181	97.8	30	96.8	50	94.3
Presence	9	2.3	4	2.2	1	3.2	3	5.7
Family history of lung cancer								
Absence	332	85.6	158	85.4	25	80.6	49	92.5
Presence	56	14.4	27	14.6	6	19.4	4	7.5
Stage								
Occult	0	0.0	1	0.7	0	0.0	0	0.0
0	1	0.3	0	0.0	0	0.0	0	0.0
IA	132	42.4	29	21.2	5	21.7	5	16.1
IB	62	19.9	40	29.2	9	39.1	3	9.7
IIA	12	3.9	2	1.5	0	0	2	6.5
IIB	13	4.2	21	15.3	3	13.0	0	0.0
IIIA	26	8.4	23	16.8	0	0	6	19.4
IIIB	30	9.6	14	10.2	3	13.0	7	22.6
IV	35	11.3	7	5.1	3	13.0	8	25.8
Unknown	77		48		8		22	
Serum CEA level								
<5 ng/mL	252	68.9	125	76.2	20	74.1	27	61.4
≥5 ng/mL	114	31.1	39	23.8	7	25.9	17	38.6
Unknown	22		21		4		9	

Abbreviations: CEA, carcinoembryonic antigen; COPD, chronic obstructive pulmonary disease; IP/PF, interstitial pneumonia/pulmonary fibrosis.

carcinoma, 3.0% for large cell carcinoma, 9.3% for small cell carcinoma, and 3.3% for others.¹⁰ The Tokushima Registry showed that the proportion was 51% for adenocarcinoma, 25% for squamous cell carcinoma, 3% for large cell carcinoma, 13% for small cell carcinoma, and 8% for others.¹¹ Our registry showed a higher proportion of adenocarcinoma and a lower proportion of squamous cell carcinoma and small cell carcinoma compared with these two registries, when males and females were combined. Histological type is known to influence the choice of therapeutic strategy.^{10,11} Therefore, the frequency of histological types may be different among the participating hospitals, which may have resulted in the differences in the proportion of histological types among the registries.

These two relevant registries permitted calculation of the prevalence of each disease stage in patients grouped according to histological type, and reported that small cell carcinoma was more

frequently diagnosed at an advanced stage compared with the other histological types of lung cancer.^{10,11} The JLC Registry reported that stage IV accounted for 40.3% of small cell carcinoma cases, while stage IV accounted for 18.6% of non-small cell carcinoma cases.¹⁰ The Tokushima Registry showed that stage IV accounted for 49.8% of small cell carcinoma cases, while stage IV accounted for 29.6% of non-small cell carcinoma cases.¹¹ Although our registry generally showed a similar pattern, stage IV accounted for a smaller percentage of small cell carcinoma cases in our registry compared with these two registries possibly due to the issue mentioned earlier.

The Japanese Association of Clinical Cancer Centers reported a 5-year relative survival rate of 43.8% in 10512 lung cancer patients who were registered between 2004 and 2007 (stage I, 39.3%; stage II, 7.2%; stage III, 24.9%; stage IV, 27.1%; and others or unknown, 1.5%).²⁰ Our lung cancer participants had a better 5-year relative

Table 4

Characteristics of the female lung cancer patients registered in the BioBank Japan Project within 90 days after their diagnosis, grouped according to histological type.

	Histological type							
	Adenocarcinoma (n = 322)		Squamous cell carcinoma (n = 32)		Large cell carcinoma (n = 6)		Small cell carcinoma (n = 16)	
	n	%	n	%	n	%	n	%
Age at study entry								
20–29 yr	0	0.0	0	0.0	0	0.0	0	0.0
30–39 yr	3	0.9	1	3.1	0	0.0	0	0.0
40–49 yr	12	3.7	0	0.0	1	16.7	1	6.3
50–59 yr	83	25.8	8	25.0	1	16.7	3	18.8
60–69 yr	119	37.0	11	34.4	3	50.0	3	18.8
70–79 yr	95	29.5	9	28.1	1	16.7	9	56.3
80–89 yr	10	3.1	3	9.4	0	0.0	0	0.0
≥90 yr	0	0.0	0	0.0	0	0.0	0	0.0
Body mass index								
<18.5 kg/m ²	37	11.7	4	12.5	2	33.3	3	20.0
18.5–24.9 kg/m ²	227	71.8	19	59.4	4	66.7	11	73.3
≥25.0 kg/m ²	52	16.5	9	28.1	0	0.0	1	6.7
Unknown	6		0		0		1	
Smoking habit								
Never	240	75.0	6	18.8	1	16.7	7	46.7
Ever (current/former)	80	25.0	26	81.3	5	83.3	8	53.3
Unknown	2		0		0		1	
Alcohol drinking habit								
Never	227	71.4	18	56.3	2	33.3	8	53.3
Ever (current/former)	91	28.6	14	43.8	4	66.7	7	46.7
Unknown	4		0		0		1	
History of COPD								
Absence	321	99.7	32	100.0	6	100.0	16	100.0
Presence	1	0.3	0	0.0	0	0.0	0	0.0
History of IP/PF								
Absence	319	99.1	32	100.0	6	100.0	16	100.0
Presence	3	0.9	0	0.0	0	0.0	0	0.0
Family history of lung cancer								
Absence	273	84.8	28	87.5	3	50.0	14	87.5
Presence	49	15.2	4	12.5	3	50.0	2	12.5
Stage								
Occult	1	0.4	0	0.0	0	0.0	0	0.0
0	0	0.0	0	0.0	0	0.0	0	0.0
IA	134	48.7	1	5.9	0	0.0	2	16.7
IB	52	18.9	3	17.6	3	50.0	2	16.7
IIA	8	2.9	0	0.0	1	16.7	0	0.0
IIB	13	4.7	4	23.5	2	33.3	2	16.7
IIIA	27	9.8	2	11.8	0	0.0	1	8.3
IIIB	15	5.5	4	23.5	0	0.0	3	25.0
IV	25	9.1	3	17.6	0	0.0	2	16.7
Unknown	47		15		0		4	
Serum CEA level								
<5 ng/mL	239	77.6	19	67.9	4	66.7	7	43.8
≥5 ng/mL	69	22.4	9	32.1	2	33.3	9	56.3
Unknown	14		4		0		0	

Abbreviations: CEA, carcinoembryonic antigen; COPD, chronic obstructive pulmonary disease; IP/PF, interstitial pneumonia/pulmonary fibrosis.

survival rate than the lung cancer patients involved in the Japanese Association of Clinical Cancer Centers, because our population had a higher frequency of stage I patients.

The JLC Registry demonstrated that more advanced stage was associated with poorer prognosis in lung cancer patients after 5 years of follow-up.¹⁰ Furthermore, small cell carcinoma had the worst prognosis, followed by large cell carcinoma, squamous cell carcinoma/adeno-squamous cell carcinoma, and adenocarcinoma.¹⁰ Our results were consistent with the results of this registry. Although our participants had more favorable clinical characteristics than the participants of this registry, our survival analysis also identified similar predictive factors for all-cause mortality in lung cancer patients. In addition to cancer stage and histological type, we observed that ever smoker, low body mass index, and elevated serum CEA level were also crudely associated with poorer prognosis after adjustment for age and study entry year. Factors associated with poorer prognosis identified among the total cases of lung cancer were similar to those of adenocarcinoma.

The strength of the present study is the enrollment of male and female lung cancer patients from many participating hospitals nationwide. In addition, we provided sex-specific information on the characteristics of Japanese lung cancer patients. On the other hand, the present study has several limitations. First, the survival analysis was conducted among cases registered within 90 days after their diagnosis, not definite incident cases. There might have been a bias resulting from selecting survival patients in our prognostic assessment. Therefore, the true prognosis of lung cancer may be worse than we observed. Second, we identified factors predicting all-cause mortality in lung cancer patients after adjustment only for age and entry year. These predictive factors may be inter-related, while other factors such as performance status and treatment may confound the associations we observed. Therefore, caution should be taken in interpreting our results. Finally, there were missing data on some variables in the BBJ Project database. The main reason for missing data was that the data were collected mostly from medical records in which limited clinical information

Table 5
Age-adjusted hazard ratios for all-cause mortality in male and female-combined lung cancer patients registered in the BioBank Japan Project within 90 days after their diagnosis, grouped according to patient characteristics, after a mean follow-up of 5.2 years. Data are also presented specifically for male and female-combined lung cancer patients with adenocarcinoma.

	Those registered within 90 days after diagnosis				Those with adenocarcinoma			
	Participants	Cases	PYFU	Age-adjusted HR (95% CI)	Participants	Cases	PYFU	Age-adjusted HR (95% CI)
Age at study entry								
Additional 1 yr increase				1.03 (1.02–1.04)				1.03 (1.02–1.05)
Body mass index								
<18.5 kg/m ²	128	75	592	1.09 (0.83–1.43)	66	34	365	1.21 (0.79–1.83)
18.5–24.9 kg/m ²	791	404	4132	1.00 (Reference)	484	206	2912	1.00 (Reference)
≥25.0 kg/m ²	185	80	1039	0.78 (0.60–1.02)	111	40	680	0.84 (0.57–1.25)
Additional 1 kg/m ² increase				0.97 (0.94–0.99)				0.95 (0.91–0.99)
Smoking habit								
Never	336	129	2056	1.00 (Reference)	277	94	1804	1.00 (Reference)
Ever (current/former)	781	442	3748	1.42 (1.07–1.88)	392	193	2183	1.72 (1.17–2.54)
Alcohol drinking habit								
Never	460	227	2401	1.00 (Reference)	300	124	1772	1.00 (Reference)
Ever (current/former)	654	342	3392	0.86 (0.70–1.07)	367	162	2204	0.85 (0.61–1.19)
Stage								
I	468	122	3365	1.00 (Reference)	355	74	2655	1.00 (Reference)
II	86	44	509	2.57 (1.63–4.06)	44	25	259	3.24 (1.64–6.40)
III	172	125	643	5.63 (4.05–7.82)	97	65	418	6.59 (4.05–10.73)
IV	87	80	173	14.83 (10.15–21.66)	57	54	122	21.20 (12.37–36.33)
Histological type								
Adenocarcinoma	671	288	3994	1.00 (Reference)				
Squamous cell carcinoma	208	123	950	1.27 (1.00–1.62)				
Large cell carcinoma	35	17	190	0.96 (0.50–1.84)				
Small cell carcinoma	66	50	201	2.03 (1.44–2.85)				
Serum CEA level								
<5 ng/mL	739	283	4535	1.00 (Reference)	474	151	3179	1.00 (Reference)
≥5 ng/mL	294	215	1095	2.55 (2.07–3.13)	173	118	758	2.75 (2.05–3.69)

Abbreviations: CEA, carcinoembryonic antigen; CI, confidence interval; HR, hazard ratio; PYFU, person-years of follow-up. The hazard ratios were calculated using a Cox proportional hazards regression model stratified by sex and institution and adjusted for age at study entry and entry year.

was available. Another possible reason for missing data on cancer stage was that the data were collected through medical records regarding each component of the TNM classification (included in the Japanese Classification of Lung Cancer, the 6th edition (2003)), but not stage. Furthermore, some cancer patients admitted participating hospitals for only therapeutic procedures or follow-up after being diagnosed at other hospitals. It is unclear whether these reasons for missing data had a crucial effect on our results.

In conclusion, the findings of lung cancer patient characteristics detailed in this report are to some extent in accordance with other nationwide or local registries on lung cancer, with the requirement of caution that our patients had less advanced stages of lung cancer, compared with other registries. In the survival analysis, we found that some characteristics at study entry would be crudely associated with all-cause mortality in Japanese lung cancer patients. In spite of the caution mentioned earlier, the BBJ Project can be expected to provide reliable and valuable evidence on lung cancer.

Conflicts of interest

All authors declare no conflicts of interest.

Acknowledgements

We express our gratitude to all the participants in the BioBank Japan Project. We thank all the medical coordinators of the cooperating hospitals for collecting samples and clinical information, as well as Yasushi Yamashita and staff members of the BioBank Japan Project for administrative support. We also thank Dr. Kumao Toyoshima for his overall supervision of the BioBank Japan project. This study was supported by funding from the Tailor-Made Medical Treatment with the BBJ Project from Japan Agency for Medical Research and development, AMED (since April 2015), and the

Ministry of Education, Culture, Sports, Science, and Technology (from April 2003 to March 2015).

Appendix. Author list for BioBank Japan Cooperative Hospital Group

Members of medical institutions cooperating on the BioBank Japan Project who coauthored this paper include Hiromasa Harada, Makoto Hibino, Atsushi Okuyama and Nobuyasu Kano (Tokushukai Hospitals); Shiro Minami, Akihiko Genma and Jitsuo Usuda (Nippon Medical School); Kenji Suzuki, Mitsuaki Sekiya and Satoru Takeda (Juntendo University); Satoshi Asai, Mitsuhiro Moriyama and Yasuo Takahashi (Nihon University); Tomoaki Fujioka and Wataru Obara (Iwate Medical University); Seiji Mori and Hideki Ito (Tokyo Metropolitan Institute of Gerontology); Satoshi Nagayama and Yoshio Miki (The Cancer Institute Hospital of JFCR); Akihiko Masumoto and Akira Yamada (Aso Iizuka Hospital); Yasuko Nishizawa and Ken Kodama (Osaka Medical Center for Cancer and Cardiovascular Diseases); Noriaki Tezuka and Yasutaka Nakano (Shiga University of Medical Science); Yukihiro Koretsune and Mitsumasa Ogawara (National Hospital Organization, Osaka National Hospital); and Kazunari Yamana (Fukujuji Hospital).

References

- Global Burden of Disease Cancer Collaboration. The global burden of cancer 2013. *JAMA Oncol.* 2015;1:505–527.
- GBD 2013 Mortality and Causes of Death Collaborators. Global, regional, and national age-sex specific all-cause and cause-specific mortality for 240 causes of death, 1990–2013: a systematic analysis for the Global Burden of Disease Study 2013. *Lancet.* 2015;385:117–171.
- Katanoda K, Hori M, Matsuda T, et al. An updated report on the trends in cancer incidence and mortality in Japan, 1958–2013. *Jpn J Clin Oncol.* 2015;45:390–401.

4. Hori M, Matsuda T, Shibata A, Katanoda K, Sobue T, Nishimoto H, Japan Cancer Surveillance Research Group. Cancer incidence and incidence rates in Japan in 2009: a study of 32 population-based cancer registries for the Monitoring of Cancer Incidence in Japan (MCIJ) project. *Jpn J Clin Oncol*. 2015;45:884–891.
5. Sobue T, Yamamoto S, Hara M, Sasazuki S, Sasaki S, Tsugane S, JPHC Study Group. Japanese Public Health Center. Cigarette smoking and subsequent risk of lung cancer by histologic type in middle-aged Japanese men and women: the JPHC study. *Int J Cancer*. 2002;99:245–251.
6. Chan M. *WHO Report on the Global Tobacco Epidemic 2008: Fresh and Alive*. Geneva: World Health Organization; 2008.
7. Ministry of Health, Labour, and Welfare, Japan. National Vital Statistics in 2014. <http://www.mhlw.go.jp/toukei/saikin/hw/jinkou/kakutei14/index.html> Accessed 25 July 2016 [in Japanese].
8. Sekine I, Nishiwaki Y, Yokose T, Nagai K, Suzuki K, Kodama T. Young lung cancer patients in Japan: different characteristics between the sexes. *Ann Thorac Surg*. 1999;67:1451–1455.
9. Nakaya N, Goto K, Saito-Nakaya K, et al. The lung cancer database project at the National Cancer Center, Japan: study design, corresponding rate and profiles of cohort. *Jpn J Clin Oncol*. 2006;36:280–284.
10. Sawabata N, Asamura H, Goya T, et al. Japanese Joint Committee for Lung Cancer Registry. Japanese Lung Cancer Registry Study: first prospective enrollment of a large number of surgical and nonsurgical cases in 2002. *J Thorac Oncol*. 2010;5:1369–1375.
11. Kanematsu T, Hanibuchi M, Tomimoto H, et al. Epidemiological and clinical features of lung cancer patients from 1999 to 2009 in Tokushima Prefecture of Japan. *J Med Invest*. 2010;57:326–333.
12. Nakamura Y. The BioBank Japan project. *Clin Adv Hematol Oncol*. 2007;5: 696–697.
13. Nagai A, Hirata M, Kamatani Y, et al. Overview of the BioBank Japan Project: study design and profile. *J Epidemiol*. 2017;27:S2–S8.
14. Hirata M, Kamatani Y, Nagai A, et al. Cross-sectional analysis of BioBank Japan clinical data: a large cohort of 200,000 patients with 47 common diseases. *J Epidemiol*. 2017;27:S9–S21.
15. Hirata M, Nagai A, Kamatani Y, et al. Overview of BioBank Japan follow-up data in 32 diseases. *J Epidemiol*. 2017;27:S22–S28.
16. Cancer Registry and Statistics, Cancer Information Service, National Cancer Center, Japan. Cohort Life Table. http://ganjoho.jp/reg_stat/statistics/qa_words/cohort01.html Accessed 25 July 2016 [in Japanese].
17. Ministry of Health, Labour and Welfare, Japan. Abridged Life Tables for Japan. <http://www.mhlw.go.jp/toukei/saikin/hw/seimei/list54-57-02.html> Accessed 25 July 2016 [in Japanese].
18. Ministry of Health, Labour and Welfare, Japan. Patient Survey in 2005. <http://www.e-stat.go.jp/SG1/estat/List.do?lid=000001047095> Accessed 25 July 2016 [in Japanese].
19. Ministry of Health, Labour and Welfare, Japan. National Health and Nutrition Survey in 2005. <http://www.mhlw.go.jp/bunya/kenkou/eiyou07/01.html> Accessed 25 July 2016 [in Japanese].
20. Japanese Association of Clinical Cancer Centers. Five-year relative survival rate in all cases in 2004–2007. <http://www.gunma-cc.jp/sarukihan/seizonritu/seizonritu2007.html#10> Accessed 25 July 2016 [in Japanese].