



Journal of Epidemiology

Contents lists available at ScienceDirect

Journal of Epidemiology

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

Original Article

Overview of BioBank Japan follow-up data in 32 diseases



Makoto Hirata ^a, Akiko Nagai ^b, Yoichiro Kamatani ^c, Toshiharu Ninomiya ^d, Akiko Tamakoshi ^e, Zentarō Yamagata ^f, Michiaki Kubo ^g, Kaori Muto ^b, Yutaka Kiyohara ^h, Taisei Mushiroda ⁱ, Yoshinori Murakami ^j, Koichiro Yuji ^k, Yoichi Furukawa ^l, Hitoshi Zembutsu ^{m, n}, Toshihiro Tanaka ^{o, p, q}, Yozo Ohnishi ^{o, r}, Yusuke Nakamura ^{m, s}, BioBank Japan Cooperative Hospital Group ^u, Koichi Matsuda ^{m, t, *}

^a Laboratory of Genome Technology, Institute of Medical Science, The University of Tokyo, Tokyo, Japan^b Department of Public Policy, Institute of Medical Science, The University of Tokyo, Tokyo, Japan^c Laboratory for Statistical Analysis, RIKEN Center for Integrative Medical Sciences, Yokohama, Japan^d Department of Epidemiology and Public Health, Graduate School of Medical Sciences, Kyushu University, Fukuoka, Japan^e Department of Public Health, Hokkaido University Graduate School of Medicine, Sapporo, Japan^f Department of Health Sciences, University of Yamanashi, Yamanashi, Japan^g RIKEN Center for Integrative Medical Sciences, Yokohama, Japan^h Hisayama Research Institute for Lifestyle Diseases, Fukuoka, Japanⁱ Laboratory for Pharmacogenomics, RIKEN Center for Integrative Medical Sciences, Yokohama, Japan^j Division of Molecular Pathology, Institute of Medical Science, The University of Tokyo, Tokyo, Japan^k Project Division of International Advanced Medical Research, Institute of Medical Science, The University of Tokyo, Tokyo, Japan^l Division of Clinical Genome Research, Institute of Medical Science, The University of Tokyo, Tokyo, Japan^m Laboratory of Molecular Medicine, Institute of Medical Science, The University of Tokyo, Tokyo, Japanⁿ Division of Genetics, National Cancer Center Research Institute, Tokyo, Japan^o SNP Research Center, RIKEN Yokohama Institute, Yokohama, Japan^p Department of Human Genetics and Disease Diversity, Graduate School of Medical and Dental Sciences, Tokyo Medical and Dental University, Tokyo, Japan^q Bioresearch Research Center, Tokyo Medical and Dental University, Tokyo, Japan^r Shinko Clinic, Medical Corporation Shinkokai, Tokyo, Japan^s Section of Hematology/Oncology, Department of Medicine, The University of Chicago, Chicago, USA^t Laboratory of Clinical Genome Sequencing, Graduate School of Frontier Sciences, The University of Tokyo, Tokyo, Japan

ARTICLE INFO

Article history:

Received 22 October 2016

Accepted 15 December 2016

Available online 10 February 2017

Keywords:

BioBank Japan project

Biobank

Common disease

Clinical information

Follow-up survey

ABSTRACT

Background: We established a patient-oriented biobank, BioBank Japan, with information on approximately 200,000 patients, suffering from any of 47 common diseases. This follow-up survey focused on 32 diseases, potentially associated with poor vital prognosis, and collected patient survival information, including cause of death. We performed a survival analysis for all subjects to get an overview of BioBank Japan follow-up data.

Methods: A total of 141,612 participants were included. The survival data were last updated in 2014. Kaplan–Meier survival analysis was performed after categorizing subjects according to sex, age group, and disease status. Relative survival rates were estimated using a survival-rate table of the Japanese general population.

Results: Of 141,612 subjects (56.48% male) with 1,087,434 person-years and a 97.0% follow-up rate, 35,482 patients died during follow-up. Mean age at enrollment was 64.24 years for male subjects and 63.98 years for female subjects. The 5-year and 10-year relative survival rates for all subjects were 0.944 and 0.911, respectively, with a median follow-up duration of 8.40 years. Patients with pancreatic cancer had the least favorable prognosis (10-year relative survival: 0.184) and patients with dyslipidemia had the most favorable prognosis (1.013). The most common cause of death was malignant neoplasms. A number of subjects died from diseases other than their registered disease(s).

* Corresponding author. Laboratory of Clinical Genome Sequencing, Graduate School of Frontier Sciences, The University of Tokyo, 4-6-1 Shirokanedai, Minato-ku, Tokyo 108-8639, Japan.

E-mail address: kmatsuda@k.u-tokyo.ac.jp (K. Matsuda).

Peer review under the responsibility of The Japan Epidemiological Association.

^u Hospital Group members are listed in Appendix A.

<http://dx.doi.org/10.1016/j.je.2016.12.006>

0917-5040/© 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/>).

Conclusions: This is the first report to perform follow-up survival analysis across various common diseases. Further studies should use detailed clinical and genomic information to identify predictors of mortality in patients with common diseases, contributing to the implementation of personalized medicine.

© 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

BioBank Japan (BBJ) was launched in 2003, establishing a large patient-oriented biobank in order to contribute to the implementation of personalized medicine. Approximately 200,000 patients diagnosed with any of 47 target common diseases were enrolled in the BBJ in the first 5-year period of the project.^{1–3} In this project, DNA, serum, and clinical information was collected from participants. This has been distributed to many researchers and companies, subject to an approved application, resulting in the publication of over 200 studies as of March 31 2016.⁴ Most studies have investigated genetic susceptibility to various diseases or genetic factors affecting drug efficacy and susceptibility to adverse drug reactions, providing important medical information for the implementation of personalized medicine.^{5–20} However, to date, no study has investigated the genes or serum markers associated with long term prognosis, despite the fact that survival analysis is one of the most important aspects for the implementation of personalized medicine. Of the 47 common diseases, we focused on 32 diseases, including malignant neoplasm and cardiovascular disease, which are potentially associated with a poor vital prognosis. In 2010, we commenced the follow-up survival survey involving BBJ participants with any of the 32 diseases and updated the follow-up data until 2014. In the present study, we identified subject characteristics for the follow-up survey using clinical information at enrollment, performed survival analysis, and investigated the distribution of causes of death in the deceased subjects.

Materials and methods

Study design

We enrolled patients with any one of the 47 target common diseases between fiscal year of 2003 and 2007. For the follow-up survival survey, we focused on 32 out of the 47 diseases, associated with participant vital status (Table 1). Data collection was started in 2010. Details of the protocols, participant recruitment, collection of clinical information, and follow-up survey, were described elsewhere.² Disease durations at the time of enrollment were calculated based on the date of onset or diagnosis of disease and the date of enrollment for 25 out of the 32 diseases.² The present protocols were reviewed and approved by the Ethics Committees of all participating institutions, including the Institute of Medical Science, the University of Tokyo, and the Center for Integrative Medical Sciences, RIKEN. Written informed consent was obtained from all participants. Baseline clinical dataset after data-cleansing, and follow-up dataset updated in 2014, were used for the present study.

Survival data analysis

We set the date of enrollment as a starting point for survival analysis. To examine whether the target 32 diseases affected vital prognosis, relative survival rates were estimated as the ratio of the observed cumulative survival rates of subjects to the expected survival rates of a group of people of the corresponding sex and age

Table 1

Baseline characteristics of subjects with 32 diseases for follow-up survey.

32 diseases	Number subjects	% of Male subjects	Mean age (y)	
			Male subjects	Female subjects
Whole cohort	141,612	56.48	64.24	63.98
Mn ^b Colorectal cancer	6180	62.88	66.87	66.20
Breast cancer	5827	0.72	63.43	57.52
Gastric cancer	5734	73.60	66.84	64.83
Prostate cancer	4669	100.00	72.45	N/A
Lung cancer	3557	64.75	67.54	65.88
Liver cancer	1757	75.98	67.46	69.63
Esophageal cancer	1193	86.67	65.64	65.42
Hematological cancer	1182	53.98	60.44	60.00
Uterine cervical cancer	1123	0.00	N/A	51.65
Uterine corpus cancer	961	0.00	N/A	58.91
Ovarian cancer	816	0.00	N/A	56.31
Pancreatic cancer	382	64.66	65.86	65.99
Gallbladder/ cholangiocarcinoma	366	62.30	67.42	68.57
Lv ^b Chronic hepatitis C	5180	53.90	63.31	64.40
Liver cirrhosis	2218	62.67	62.60	64.99
Chronic hepatitis B	1179	62.60	54.45	55.19
Cr ^b Cerebral infarction	13,955	63.55	68.46	70.90
Cerebral aneurysm	2426	35.57	60.36	62.55
Cv ^b Arrhythmia	14,184	65.01	66.82	68.78
Stable angina	13,117	71.13	67.65	70.57
Myocardial infarction	11,982	81.29	65.81	70.72
Heart failure	6888	62.66	65.78	70.94
Unstable angina	3944	74.62	66.62	71.08
Peripheral arterial diseases	2444	78.23	70.76	71.41
Rs ^b Bronchial asthma	7380	49.30	52.24	53.20
COPD ^a	2257	86.97	72.16	72.36
Interstitial lung disease/Pulmonary fibrosis	749	58.21	68.53	67.54
Mb ^b Dyslipidemia	37,478	51.71	62.00	65.93
Diabetes mellitus	34,523	64.03	63.13	65.39
Lc ^b Rheumatoid arthritis	3619	20.45	63.66	61.92
Osteoporosis	5028	7.52	71.37	73.08
Other Drug eruption	527	46.11	60.29	54.67

^a Chronic obstructive pulmonary disease.

^b Left boxes indicate disease categories for following analysis: Mn, malignant neoplasms; Lv, liver diseases; Cr, cerebrovascular diseases; Cv, cardiovascular diseases; Rs, respiratory diseases; Mb, metabolic diseases; Lc, locomotive diseases.

in the general population. To calculate expected survival rates, a survival-rate table of general Japanese population was obtained from the Cancer Registry and Statistics, Cancer Information Service, National Cancer Center, Japan.²¹ The survival-rate table 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.²² Fifteen subjects ≥ 100 years old were excluded due to the lack of data in the general population life table. Information on the cause of death, according to the ICD-10 code, was collected for deceased subjects by matching birth date, date of death, sex, and local government code, with the Vital Statistics by the Statistics and Information Department of the Ministry of Health, Labour and Welfare, Japan.²³ SAS 9.4 software were used for the data analysis. A p-value of <0.05 was considered statistically significant.

Results

Characteristics of the subjects in follow-up survey

Of 161,822 participants with any one of the 32 diseases in the BBJ cohort, 20,210 participants were excluded from this follow-up survey due to withdrawal of consent, withdrawal of hospitals, refusal of follow-up survey, and registration error of the target disease.² A total of 141,612 (87.51%) were included in the follow-up survey. Among the 32 diseases, dyslipidemia was the most common disease ($n = 37,478$), while gallbladder/cholangiocarcinoma were the least common ($n = 366$) (Table 1). Of all subjects, 56.48% were male and mean age at enrollment was 64.24 years for male subjects and 63.98 years for female subjects. Although the sex ratio varied among the diseases, there were no substantial differences of the mean age at enrollment between the sexes, according to disease. As all the participants were enrolled after onset or diagnosis of the target disease(s), we examined whether disease durations at enrollment would affect participants' survivals. More than half of the subjects with 8 of the 13 malignant neoplasms were enrolled within a year after diagnosis (eTable 1-1), while most of the subjects with other chronic diseases were enrolled more than 3 years (or even 5 years) after onset or diagnosis of the disease (eTables 1-2 and 3), after excluding the subjects with unknown disease durations.

Outlines of follow-up survival data

Analysis of follow-up data showed that median follow-up duration in all subjects was 8.40 years and follow-up rate was

97.0%. Of the 141,612 subjects with 1,087,434 person-years, 35,482 subjects died during follow-up. We could identify the causes of death in 31,054 (87.5%) cases. The follow-up rates across the diseases were comparable and even the lowest follow-up rate was 93.4% in drug eruption. In contrast, the median follow-up duration for each disease varied from 1.29 years for pancreatic cancer to 9.16 years for cerebral aneurysm, mostly dependent on the prognosis of each disease (Table 2). The 5-year and 10-year relative survival rates in all subjects were 0.944 and 0.911, respectively. The relative survival rates in each disease showed that pancreatic cancer exhibited the worst prognosis (0.291 over 5 years and 0.235 over 10 years), followed by other hepato-biliary-pancreatic diseases such as liver cancer, gallbladder/cholangiocarcinoma, and liver cirrhosis. Of note, the 10-year survival rates for dyslipidemia and prostate cancer were higher than those in the general population (10-year relative survival rate: 1.013 for both diseases) (Table 2). The observed survival analysis, stratified by disease durations at enrollment revealed that disease durations within one year were associated with low survival rates for most malignant neoplasms except liver, prostate and breast cancers. In contrast, longer disease durations were associated with lower survival rates for metabolic diseases: dyslipidemia and diabetes mellitus (eTable 1).

Survival analysis with sex and age

Observed cumulative survival rates of all subjects, analyzed by Kaplan–Meier methods, showed improved prognosis in female subjects, while relative survival rates were similar between both sexes (Fig. 1A).

Table 2
Follow-up characteristics of subjects with any one of 32 diseases.

32 diseases	Median follow-up duration (y)	Follow-up rate (%)	Observed survival rate (95% CI)		Relative survival rate (95% CI) ^a	
			5-year ^b	10-year ^b	5-year ^b	10-year ^b
Whole cohort	8.40	97.0	0.855 (0.853–0.857)	0.719 (0.716–0.721)	0.944 (0.942–0.946)	0.911 (0.907–0.914)
Colorectal cancer	7.95	97.7	0.759 (0.748–0.770)	0.625 (0.611–0.638)	0.849 (0.837–0.861)	0.816 (0.798–0.833)
Breast cancer	8.18	98.1	0.892 (0.883–0.899)	0.794 (0.782–0.806)	0.924 (0.915–0.931)	0.870 (0.857–0.883)
Gastric cancer	7.90	97.4	0.743 (0.732–0.755)	0.598 (0.584–0.612)	0.837 (0.824–0.850)	0.796 (0.778–0.815)
Prostate cancer	7.83	97.8	0.802 (0.790–0.813)	0.609 (0.592–0.625)	0.987 (0.972–1.000)	1.013 (0.985–1.040)
Lung cancer	6.52	97.5	0.591 (0.574–0.607)	0.446 (0.428–0.465)	0.703 (0.683–0.722)	0.685 (0.657–0.714)
Liver cancer	3.31	96.2	0.398 (0.375–0.422)	0.188 (0.167–0.211)	0.448 (0.422–0.475)	0.251 (0.223–0.281)
Esophageal cancer	5.49	97.5	0.539 (0.509–0.567)	0.380 (0.348–0.412)	0.599 (0.565–0.630)	0.496 (0.454–0.538)
Hematologic cancer	8.03	98.3	0.787 (0.762–0.809)	0.632 (0.599–0.664)	0.850 (0.823–0.874)	0.760 (0.720–0.798)
Uterine cervical cancer	8.97	96.3	0.890 (0.870–0.907)	0.839 (0.814–0.860)	0.912 (0.891–0.930)	0.889 (0.863–0.911)
Uterine corpus cancer	8.96	97.5	0.908 (0.887–0.924)	0.853 (0.827–0.876)	0.938 (0.917–0.955)	0.933 (0.904–0.958)
Ovarian cancer	8.52	97.9	0.771 (0.741–0.799)	0.685 (0.650–0.717)	0.792 (0.761–0.821)	0.733 (0.695–0.767)
Pancreatic cancer	1.29	95.8	0.262 (0.218–0.308)	0.184 (0.144–0.228)	0.291 (0.242–0.342)	0.235 (0.184–0.292)
Gallbladder/cholangiocarcinoma	2.36	97.5	0.406 (0.355–0.456)	0.292 (0.241–0.344)	0.455 (0.398–0.511)	0.384 (0.317–0.453)
Chronic hepatitis C	7.90	96.0	0.782 (0.770–0.793)	0.608 (0.592–0.623)	0.852 (0.839–0.864)	0.751 (0.731–0.769)
Liver cirrhosis	4.44	94.4	0.502 (0.481–0.523)	0.299 (0.278–0.321)	0.548 (0.525–0.571)	0.369 (0.343–0.396)
Chronic hepatitis B	8.22	95.3	0.857 (0.835–0.876)	0.770 (0.741–0.796)	0.899 (0.876–0.919)	0.871 (0.838–0.900)
Cerebral infarction	8.36	97.0	0.836 (0.829–0.842)	0.644 (0.635–0.653)	0.964 (0.956–0.971)	0.899 (0.887–0.912)
Cerebral aneurysm	9.16	98.2	0.931 (0.920–0.941)	0.834 (0.817–0.850)	0.988 (0.977–0.999)	0.965 (0.945–0.984)
Arrhythmia	8.24	97.4	0.847 (0.841–0.853)	0.667 (0.658–0.676)	0.968 (0.961–0.975)	0.915 (0.903–0.928)
Stable angina	8.54	97.5	0.864 (0.858–0.870)	0.703 (0.694–0.711)	0.988 (0.981–0.995)	0.967 (0.955–0.978)
Myocardial infarction	8.55	97.3	0.849 (0.842–0.855)	0.677 (0.667–0.686)	0.963 (0.955–0.970)	0.909 (0.896–0.921)
Heart failure	7.38	96.9	0.720 (0.709–0.731)	0.504 (0.489–0.518)	0.830 (0.817–0.842)	0.696 (0.675–0.715)
Unstable angina	8.86	97.2	0.862 (0.850–0.872)	0.687 (0.671–0.703)	0.978 (0.964–0.989)	0.929 (0.907–0.951)
Peripheral arterial diseases	6.94	96.8	0.700 (0.681–0.718)	0.443 (0.418–0.467)	0.832 (0.809–0.853)	0.668 (0.630–0.704)
Bronchial asthma	8.74	94.8	0.923 (0.917–0.929)	0.829 (0.819–0.839)	0.986 (0.980–0.993)	0.972 (0.960–0.984)
COPD ^c	6.89	96.1	0.664 (0.644–0.683)	0.403 (0.379–0.427)	0.812 (0.787–0.835)	0.655 (0.616–0.694)
Interstitial lung disease	5.81	98.0	0.558 (0.522–0.593)	0.338 (0.298–0.379)	0.636 (0.595–0.677)	0.462 (0.407–0.518)
Dyslipidemia	8.75	97.5	0.929 (0.927–0.932)	0.825 (0.820–0.829)	1.010 (1.008–1.013)	1.013 (1.007–1.018)
Diabetes mellitus	8.44	96.9	0.861 (0.857–0.864)	0.710 (0.704–0.715)	0.944 (0.940–0.947)	0.887 (0.880–0.893)
Rheumatoid arthritis	8.47	97.4	0.873 (0.861–0.883)	0.720 (0.702–0.736)	0.929 (0.916–0.939)	0.836 (0.816–0.855)
Osteoporosis	8.42	96.8	0.850 (0.839–0.859)	0.691 (0.676–0.706)	0.980 (0.967–0.990)	0.965 (0.944–0.986)
Drug eruption	8.51	93.4	0.858 (0.824–0.886)	0.743 (0.698–0.783)	0.921 (0.885–0.951)	0.875 (0.822–0.922)

^a Relative survival rates were estimated as the ratio of the observed cumulative survival rates of subjects to the expected survival rates, based on abridged life tables from Ministry of Health, Labour and Welfare, Japan. Subjects ≥ 100 years old were excluded because of lack of data in the reference life tables.

^b Each period describes duration after registration.

^c Chronic obstructive pulmonary disease.

Thirty-one of the 32 diseases were categorized into seven groups: malignant neoplasms, liver disease, cerebrovascular disease, cardiovascular disease, respiratory disease, metabolic disease, and locomotive disease (Table 1). Further survival analysis in each disease category also found a poorer prognosis among male subjects, across all disease categories (Fig. 2). Even in relative survival analysis, some disease categories, such as malignant neoplasms, liver disease, respiratory disease, and locomotive disease, exhibited a poorer prognosis among male subjects. Conversely, the relative survival rates in cardiovascular disease showed a poorer prognosis among female subjects. Furthermore, the relative survival rates showed a rapid decline during the first four years but a slower decline at a later period for malignant neoplasms and liver disease, while other diseases displayed a consistent decline throughout follow-up (Fig. 2). When the subjects were categorized according to

their age groups, cumulative survival rates showed a clear trend of poorer prognosis with advancing age (Fig. 1B).

Analysis of cause of death

The causes of death based on the ICD-10 code were also analyzed. The four most frequent underlying causes of death in Japan are malignant neoplasm (coded as C00–C97 by ICD-10), cardiovascular disease other than hypertension (I01–I02, I05–I09, I20–I25, and I30–I52), pneumonia (J12–J18), and cerebrovascular disease (I60–I69).²³ Therefore, the underlying causes of death in the present analysis followed the same categorization. The order of frequency among all subjects according to the four cause-of-death categories was comparable with that in Vital Statistics, Japan, 2014 (Fig. 3). Analysis of the cause of death by disease category revealed that the causes of death tended to be related to the registered disease; malignant neoplasm was the cause of death in 76.8% of subjects with malignant neoplasms (42.9% of all subjects), cardiovascular disease was the cause of death in 27.6% of subjects with cardiovascular disease (15.2% of all subjects), and cerebrovascular disease was the cause of death in 17.1% of subjects with cerebrovascular diseases (7.3% of all subjects) (Fig. 3).

Discussion

In the present follow-up survey, 141,612 BioBank Japan participants with any of the 32 diseases were included. Analysis of baseline clinical information of the subjects showed comparable characteristics (at least, according to sex and age) among all participants with the any one of the 32 diseases (Table 1),³ even though a number of participants were excluded for various reasons.² Of the 141,612 subjects with 1,087,434 person-years and 97.0% follow-up rate, 35,482 subjects died during the follow-up. The median follow-up duration was 8.40 years. Relative survival analysis revealed poor prognosis for some of the cancers such as pancreatic cancer, liver cancer, and gallbladder/cholangiocarcinoma (Table 2), which had already been reported as lethal malignancies.^{24–26} Furthermore, two other chronic diseases, liver cirrhosis and interstitial lung disease/lung fibrosis, were also associated with a poor prognosis, with a 10-year-relative survival rate of less than 50%. Analysis of the causes of death showed that the four most frequent underlying causes of death in all the subjects were malignant neoplasm, cardiovascular disease other than hypertension, pneumonia, and cerebrovascular disease.

To our knowledge, this is the first study to perform long-term follow-up survival surveys on various common diseases, including malignant neoplasms and other chronic diseases. Analysis of relative survival rates according to sex showed that male subjects had a poorer prognosis for some disease categories, such as malignant neoplasms, liver disease, respiratory disease, and locomotive disease, while female subjects had a poorer prognosis for cardiovascular disease. These results indicated that sex-linked factors may specifically affect survival in such diseases. Several earlier studies on malignant neoplasms described relative survival rates after stratifying for some clinical variables, such as sex, presenting status, histological grade, and disease stage. Further detailed survival analysis on malignant neoplasms among the BBJ subjects with malignant neoplasms was also performed.^{27–32} It is difficult to compare the results of relative survivals between the current and the previous studies, because of the different subject characteristics, such as ethnicity and life-style, and the differences in the health care system in other countries and the methods of enrollment used in Japan. The relative survival rates of malignant neoplasms in the current study appeared higher than those in cancer statistics in Japan, published by the Foundation for

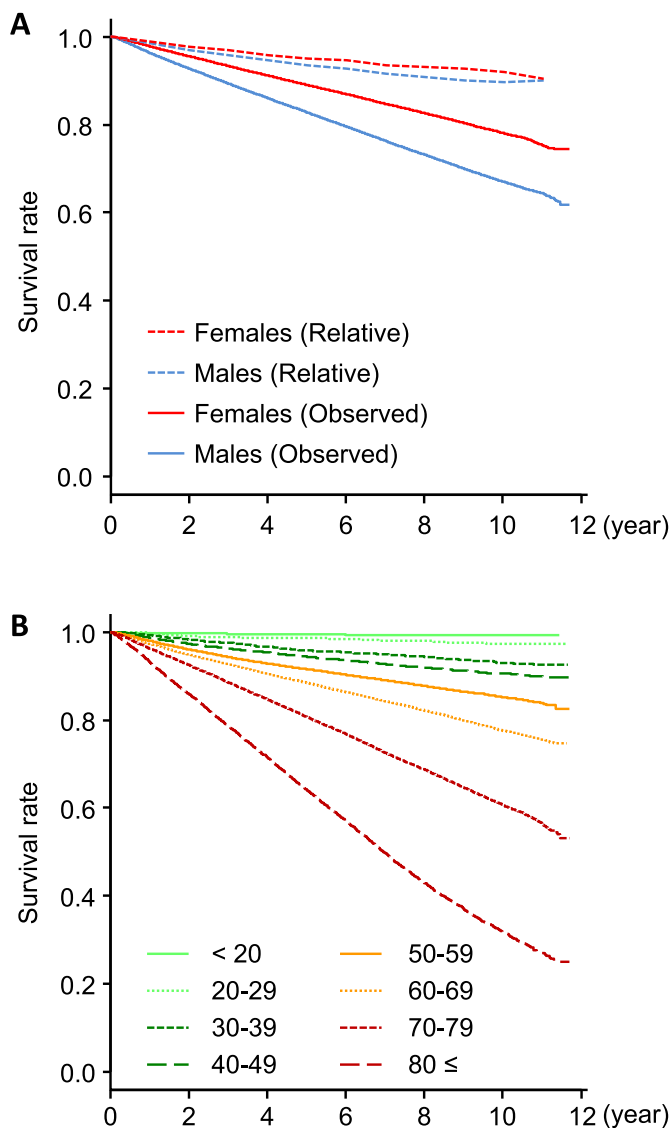


Fig. 1. Survival curves of all subjects based on sex and age group classification. (A) Observed cumulative survival curves (solid) were analyzed by Kaplan–Meier methods. Relative survival rates (dotted) were estimated as the ratio of the observed cumulative survival rates of subjects to the expected survival rates, based on a survival-rate table of general Japanese population from the Cancer Registry and Statistics, Cancer Information Service, National Cancer Center, Japan. (B) Observed cumulative survival rates of all subjects by age-groups were analyzed by Kaplan–Meier methods. Categorization of age groups was performed according to age at enrollment.

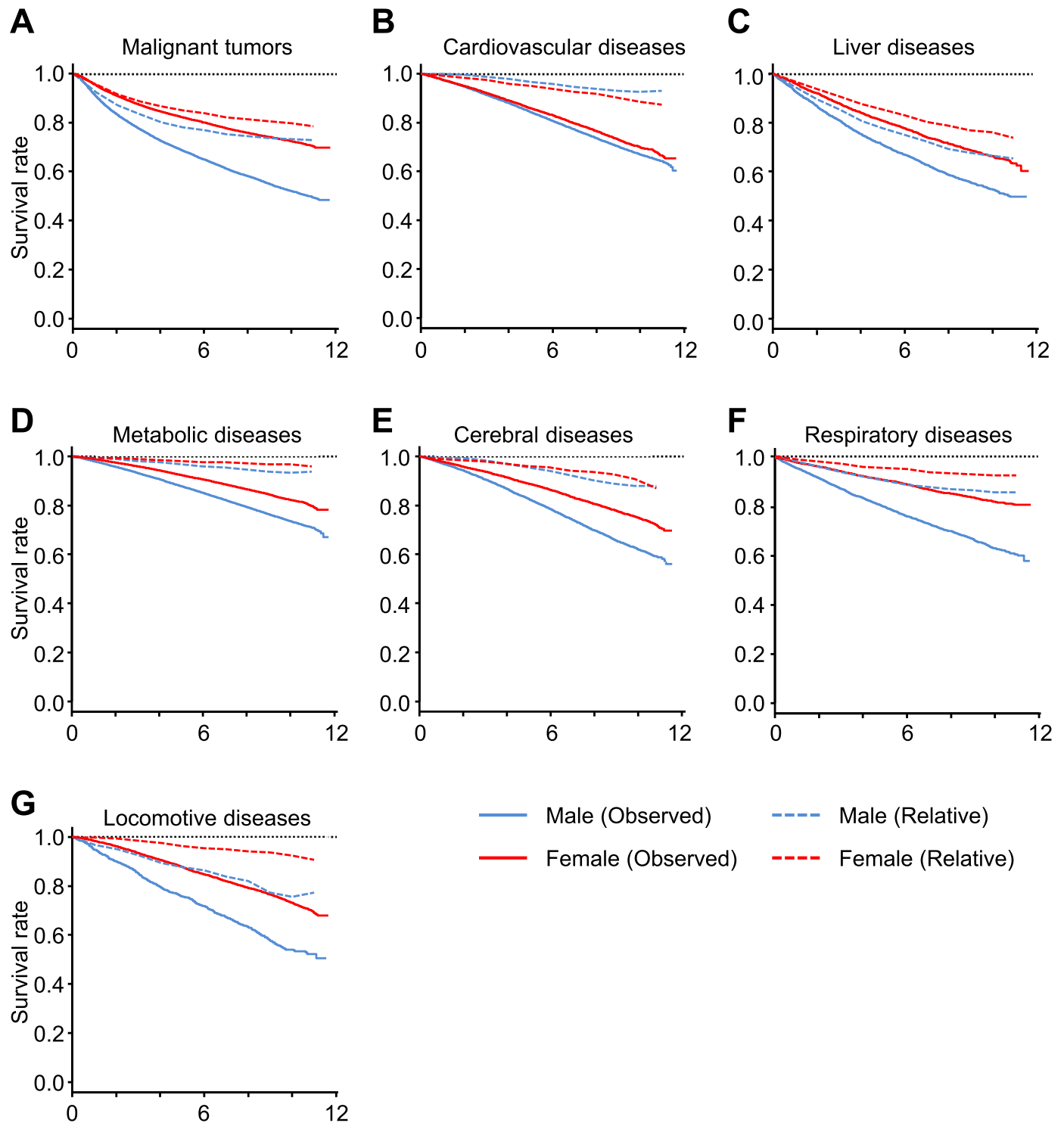


Fig. 2. Survival curves of patients by disease group. Observed cumulative survival curves (solid) and relative survival rates (dotted) of male and female BBJ patients for each disease group: Malignant neoplasms (A), Liver disease (B), Respiratory disease (C), Cardiovascular disease (D), Cerebrovascular diseases (E), Metabolic diseases (F), and Locomotive diseases (G). Observed cumulative survival curves and relative survival rates of subjects to the expected survival rates were analyzed by Kaplan–Meier methods. Relative survival rates were estimated as the ratio of the observed cumulative survival rates of subjects to the expected survival rates.

Promotion of Cancer Research.³³ This result is not surprising because a substantial proportion of cancer patients were registered more than one year after diagnosis, increasing the proportion of BBJ subjects with more favorable prognosis (eTable 1–1).

Relative survival analysis also showed significantly better outcome in subjects with dyslipidemia than in the general

population (Table 2), even though the population should be unbiased representative, based on the national public database Vital Statistics. As dyslipidemia was expected to be associated with poor clinical outcome, as it is a major risk factor for atherosclerotic disease,³⁴ further analysis is essential to identify the factors associated with a better prognosis in BBJ subjects with dyslipidemia.

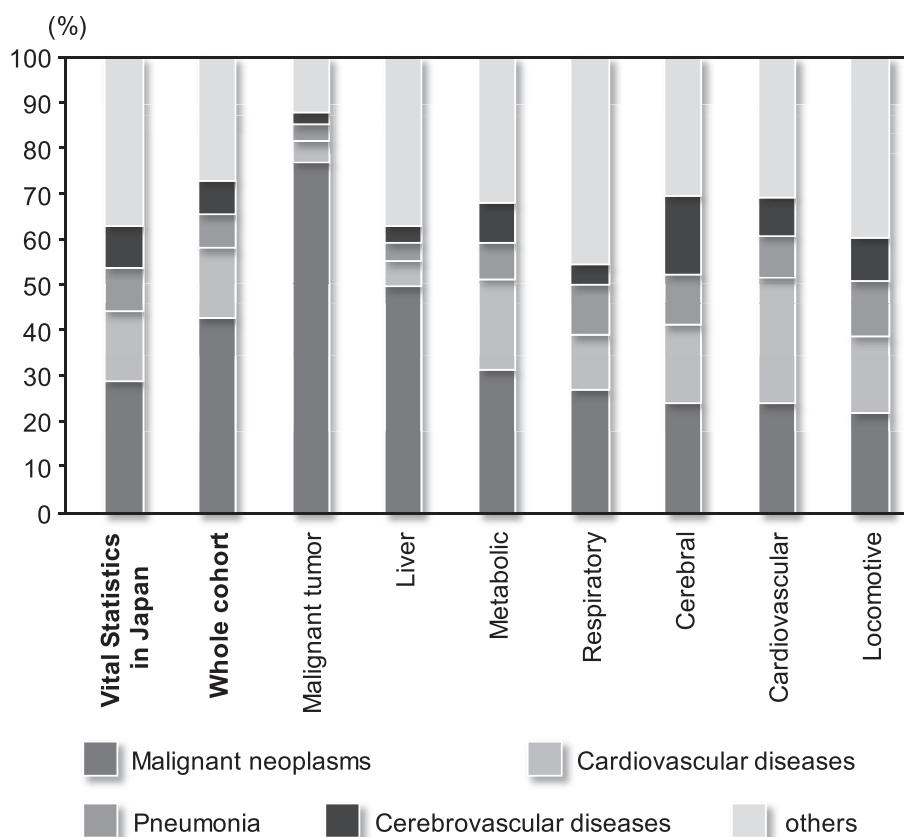


Fig. 3. Distribution of underlying cause of death in all subjects and disease categories. Each bar represents the distribution of underlying cause of death, categorized by ICD10 in all subjects or each disease category; malignant neoplasms were coded by C00–C97, cardiovascular diseases (other than hypertension) I01–I02, I05–I09, I20–I25, I27 and I30–I52, pneumonia J12–J18, cerebrovascular diseases I60–69. This categorization was based on Vital statistics (first from left), Japan in 2014, published by Ministry of Health Labour and Welfare, Japan.

The most frequent cause of death in the present study was malignant neoplasm, followed by cardiovascular disease (other than hypertension), pneumonia, and cerebrovascular disease, which were also the four most frequent causes of death in the Vital Statistics (Fig. 3).²³ However, a considerable number of subjects died from diseases other than the registered diseases. Indeed, as many participants were registered with multiple diseases, we need to consider the effect or interaction of comorbidity. Further investigation using detailed clinical information, including comorbidity, may explain how registered disease(s) are associated with a cause of death.

One of the limitations in the present study is the selection bias at the enrollment, as 12.5% of the BBJ participants with the 32 diseases were excluded from the present study as they did not agree to the follow-up survey or for other reasons,² and some of the subjects had been enrolled several years after diagnosis or onset of the disease(s). As it is generally challenging to set starting points for survival analysis follow-up in chronic diseases, due to their ambiguous onset, in the present study, we set the starting point for follow-up at enrollment. However, we still need to consider the possibility that such selection bias interferes with the natural history of results. A further limitation is incomplete information relating to the cause of death. We could not identify the cause of death in approximately 12.5% of the deceased subjects because Vital Statics data of deceased individuals in 2014 was not available for this survey.

In conclusion, a follow-up survey was performed for subjects with any one of the 32 target diseases in the BBJ Project. Characterization of the subjects enhanced the feasibility of the follow-up survey and survival analysis displayed different survival rates across the diseases. Further studies, using detailed clinical and survival data, as

well as genome-wide single-nucleotide-polymorphism data, already obtained from approximately 170,000 participants in BBJ,^{2,3} may prompt us to identify precise predictors for mortality and allow for the implementation of personalized medicine.

Conflicts of interest

None declared.

Acknowledgements

We express our gratitude to all the participants of 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 (from April 2015), and the Ministry of Education, Culture, Sports, Science, and Technology (from April 2003 to March 2015).

Appendix A

Author list for the BioBank Japan Cooperative Hospital Group

Members of medical institutions cooperating on the BioBank Japan Project who coauthored this paper include Masaki Shiono,

Kazuo Misumi, Reiji Kaieda, Hiromasa Harada (Tokushukai Hospitals); Shiro Minami, Atsushi Watanabe, Naoya Emoto (Nippon Medical School), Kazuhisa Takahashi, Satoru Takeda, Toshinari Funaki (Juntendo University), Satoshi Asai, Mitsuhiko Moriyama, Yasuo Takahashi (Nihon University), Tomoaki Fujioka, Wataru Obara (Iwate Medical University), Seiji Mori, Hideki Ito (Tokyo Metropolitan Institute of Gerontology), Satoshi Nagayama, Yoshio Miki (The Cancer Institute Hospital of JFCR), Akihito Masumoto, Akira Yamada (Aso Iizuka Hospital), Yasuko Nishizawa, Ken Kodama (Osaka Medical Center for Cancer and Cardiovascular Diseases), Hiromu Kutsumi, Yoshihisa Sugimoto (Shiga University of Medical Science), Yukihiko Koretsune, Hideo Kusuoka (National Hospital Organization, Osaka National Hospital), and Kozo Yoshimori (Fukujuji Hospital).

Appendix B. Supplementary data

Supplementary data related to this article can be found at <http://dx.doi.org/10.1016/j.je.2016.12.006>.

References

- Nakamura Y. The BioBank Japan project. *Clin Adv Hematol Oncol*. 2007;5:696–697.
- Nagai A, Hirata M, Kamatani Y, et al. Overview of the BioBank Japan project: study design and profile. *J Epidemiol*. 2017;27:S2–S8.
- 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.
- BioBank Japan. Publications from BioBank Japan. <https://biobankjp.org/work/public.html>; Updated 30.06.16. Accessed 25 July 2016.
- Kubo M. Recent advances in the genome-wide association study of common diseases. *Nihon Rinsho*. 2010;68(Suppl 8):129–133 (in Japanese).
- Arakawa S, Takahashi A, Ashikawa K, et al. Genome-wide association study identifies two susceptibility loci for exudative age-related macular degeneration in the Japanese population. *Nat Genet*. 2011;43:1001–1004.
- Hirota T, Takahashi A, Kubo M, et al. Genome-wide association study identifies three new susceptibility loci for adult asthma in the Japanese population. *Nat Genet*. 2011;43:893–896.
- Hirota T, Takahashi A, Kubo M, et al. Genome-wide association study identifies eight new susceptibility loci for atopic dermatitis in the Japanese population. *Nat Genet*. 2012;44:1222–1226.
- Kou I, Takahashi Y, Johnson TA, et al. Genetic variants in GPR126 are associated with adolescent idiopathic scoliosis. *Nat Genet*. 2013;45:676–679.
- Kumar V, Kato N, Urabe Y, et al. Genome-wide association study identifies a susceptibility locus for HCV-induced hepatocellular carcinoma. *Nat Genet*. 2011;43:455–458.
- Miki D, Ochi H, Hayes CN, et al. Variation in the DEPDC5 locus is associated with progression to hepatocellular carcinoma in chronic hepatitis C virus carriers. *Nat Genet*. 2011;43:797–800.
- Okada Y, Kubo M, Ohmiya H, et al. Common variants at CDKAL1 and KLF9 are associated with body mass index in East Asian populations. *Nat Genet*. 2012;44:302–306.
- Okada Y, Momozawa Y, Ashikawa K, et al. Construction of a population-specific HLA imputation reference panel and its application to Graves' disease risk in Japanese. *Nat Genet*. 2015;47:798–802.
- Okada Y, Sim X, Go MJ, et al. Meta-analysis identifies multiple loci associated with kidney function-related traits in East Asian populations. *Nat Genet*. 2012;44:904–909.
- Okada Y, Terao C, Ikari K, et al. Meta-analysis identifies nine new loci associated with rheumatoid arthritis in the Japanese population. *Nat Genet*. 2012;44:511–516.
- Okamoto K, Tokunaga K, Doi K, et al. Common variation in GPC5 is associated with acquired nephrotic syndrome. *Nat Genet*. 2011;43:459–463.
- Onouchi Y, Ozaki K, Burns JC, et al. A genome-wide association study identifies three new risk loci for Kawasaki disease. *Nat Genet*. 2012;44:517–521.
- Shiraishi K, Kunitoh H, Daigo Y, et al. A genome-wide association study identifies two new susceptibility loci for lung adenocarcinoma in the Japanese population. *Nat Genet*. 2012;44:900–903.
- Takahashi Y, Kou I, Takahashi A, et al. A genome-wide association study identifies common variants near LBX1 associated with adolescent idiopathic scoliosis. *Nat Genet*. 2011;43:1237–1240.
- Tanikawa C, Urabe Y, Matsuo K, et al. A genome-wide association study identifies two susceptibility loci for duodenal ulcer in the Japanese population. *Nat Genet*. 2012;44(430–434):S1–S2.
- Cancer Registry and Statistics, Cancer Information Service, National Cancer Center. Cohort Life Table. http://ganjoho.jp/reg_stat/statistics/qa_words/cohort01.html (in Japanese); Updated 11.03.16. Accessed 28 March 2016.
- Ministry of Health, Labour and Welfare, Japan. *Abridged Life Tables for Japan; 2014* (in Japanese) <http://www.mhlw.go.jp/toukei/saikin/hw/seimei/list54-57-02.html>. Accessed 28 March 2016.
- Ministry of Health, Labour and Welfare, Japan. Vital Statistics. http://www.mhlw.go.jp/toukei/saikin/hw/jinkou/kakutei14/dl/10_h6.pdf (in Japanese); Updated 01.06.15. Accessed 03 March 2016.
- Ansari D, Tingstedt B, Andersson B, et al. Pancreatic cancer: yesterday, today and tomorrow. *Future Oncol*. 2016;12:1929–1946.
- Blechacz B, Gores GJ. Cholangiocarcinoma: advances in pathogenesis, diagnosis, and treatment. *Hepatology*. 2008;48:308–321.
- Wallace MC, Preen D, Jeffrey GP, Adams LA. The evolving epidemiology of hepatocellular carcinoma: a global perspective. *Expert Rev Gastroenterol Hepatol*. 2015;9:765–779.
- Tamakoshi A, Nakamura K, Ukawa S, et al. Characteristics and prognosis of Japanese colorectal cancer patients: the BioBank Japan Project. *J Epidemiol*. 2017;27:S36–S42.
- Nakamura K, Ukawa S, Okada E, et al. Characteristics and prognosis of Japanese male and female lung cancer patients: the BioBank Japan Project. *J Epidemiol*. 2017;27:S49–S57.
- Nakamura K, Okada E, Ukawa S, et al. Characteristics and prognosis of Japanese female breast cancer patients: the BioBank Japan Project. *J Epidemiol*. 2017;27:S58–S64.
- Ukawa S, Nakamura K, Okada E, et al. Characteristics of the patients with liver cancer in the BioBank Japan Project. *J Epidemiol*. 2017;27:S43–S48.
- Ukawa S, Nakamura K, Okada E, et al. Clinical and histopathological characteristics of the patients with prostate cancer in the BioBank Japan Project. *J Epidemiol*. 2017;27:S65–S70.
- Okada E, Nakamura K, Ukawa S, et al. Demographic and lifestyle factors and survival among patients with esophageal and gastric cancer: the BioBank Japan Project. *J Epidemiol*. 2017;27:S29–S35.
- Foundation for Promotion of Cancer Research, Japan. *Cancer Statistics in Japan, 2015; 2005*. http://ganjoho.jp/en/professional/statistics/brochure/2015_en.html. Accessed 20 June 2016.
- Anderson TJ, Mancini GB, Genest Jr J, Gregoire J, Lonn EM, Hegele RA. The new dyslipidemia guidelines: what is the debate? *Can J Cardiol*. 2015;31:605–612.