

Sensitivity and Specificity of Lung Cancer Screening in Osaka, Japan

Tomotaka Sobue,^{1,8} Takaichiro Suzuki,² Minoru Matsuda,³ Takeshi Horai,³ Akiyoshi Kajita,⁴ Keiko Kuriyama,⁴ Masahiro Fukuoka,⁵ Yoko Kusunoki,⁵ Masanori Kikui,⁶ Shin-ei Ryu⁷ and Isaburo Fujimoto¹

¹Department of Field Research, ²Division of Cancer Epidemiology, Research Institute, ³Department of Respiratory Diseases, ⁴Department of Radiology, Center for Adult Diseases, Osaka, 1-3-3 Nakamichi, Higashinari-ku, Osaka 537, ⁵Second Department of Internal Medicine, ⁶Department of Pathology, Osaka Prefectural Habikino Hospital, 3-7-1 Habikino, Habikino-shi, Osaka 583 and ⁷Department of Internal Medicine, Shimonoseki Kousei Hospital, Shimonoseki-shi, Yamaguchi 750

Sensitivity and specificity were evaluated for lung cancer screening conducted at 8 municipalities in Osaka Prefecture during 1981-1985. As a screening policy, all attendants were examined by miniature chest X-ray, and the high-risk group, defined as those who smoked cigarettes or had bloody sputum, were also examined by 3-day pooled sputum cytology. A total of 33,599 screening tests for 19,028 people who were 40 years old or more at the time of screening were conducted, resulting in 33,490 miniature chest X-ray examinations for 18,992 people and 11,420 sputum cytologies for 7,070 people. As a result, 43 lung cancer cases were detected. All test-negatives were followed by means of record linkage with the files of the Osaka Cancer Registry up to the end of 1986. There were 24 cases who were diagnosed as having lung cancer without having given a positive screening result in 1981-1986. Assuming the preclinical detectable phase of lung cancer to be one year uniformly, the sensitivity and specificity for the lung cancer screening were estimated to be 71.6% and 95.3%, respectively. The feasibility of increasing the sensitivity is discussed.

Key words: Sensitivity — Specificity — Screening — Lung cancer

In 1987, the Japanese Ministry of Health and Welfare decided to include lung cancer screening as one of the cancer screening programs provided by the Health and Medical Services Law for the Aged. However, since it has been reported from the US¹⁻³⁾ and other countries^{4,5)} that lung cancer screening programs were not effective in terms of the reduction of mortality due to lung cancer, there have been arguments against the implementation of lung cancer screening as a public health measure. Unfortunately, these debates have not been based on results obtained in Japan, although a case-control study for evaluating the efficacy of lung cancer screening is now on-going in the Research Group supported by a Grant-in-Aid for Cancer Research (62-4) from the Japanese Ministry of Health and Welfare.

So far in Japan, no results have been published either on the validity of lung cancer screening or on the efficacy of screening in reducing lung cancer deaths. The validity of cancer screening should be measured in terms of sensitivity and specificity.⁶⁾ Moreover, since the developmental characteristics of cancer, such as the clinical stage at diagnosis, differ between asymptomatic patients detected by screening and symptomatic patients in hospitals, sensitivity measured for patients in hospitals tends to overestimate the actual value of sensitivity in screen settings.⁷⁾ Therefore, the sensitivity and specificity

of cancer screening should be measured by analyzing the actual data obtained from the screening. In order to accomplish this, however, it is necessary to follow up a large number of test-negatives and to detect interval cases, defined as those cases who were diagnosed as having lung cancer after they attended the screening at least once, and had not been found positive.

In Osaka, a population-based cancer registry has been in operation since 1962, with the cooperation of the Osaka Medical Association, the Osaka Prefectural Health Department, and the Center for Adult Diseases, Osaka. This enabled us to obtain information on lung cancer incidence among all test-negatives. The present study aimed to measure the sensitivity and specificity of lung cancer screening in Osaka by means of record linkage to the data in the Osaka Cancer Registry, and to explore ways to improve the validity of the tests.

MATERIALS AND METHODS

In 1981, a lung cancer screening trial was started in Osaka Prefecture in cooperation with local municipalities, Osaka Prefectural Habikino Hospital, and the Center for Adult Diseases, Osaka. At the beginning of the trial, 2 municipalities were involved. Since then, the study population has been gradually expanded, and 8 municipalities were involved in 1985. All residents over 40 years old were invited to the screening except those

⁸ To whom all correspondence should be addressed.

who had had the chance to be screened in occupational settings. As a screening policy, all attendants were examined by miniature chest X-ray (mostly 100 mm × 100 mm). The high risk group was defined as those who smoked more than 400 Cigarette Index (average number of cigarettes smoked per day × years smoked) or those who had had bloody sputum within the last 6 months, and this group was examined by 3-day pooled sputum cytology.⁸⁾ The miniature chest X-rays were independently examined by two experienced physicians, and previous X-rays of the same screenees were reviewed, when necessary. Those who were suspected as having lung cancer and/or active tuberculosis on the miniature chest X-ray or those who were classified as class C (squamous metaplasia with moderate atypia), D (squamous metaplasia with severe atypia), or E (cancer)⁹⁾ by sputum cytology, were asked to undergo further diagnostic evaluation.

The further diagnostic evaluation consisted of two steps. First, all those people asked to return were examined by full-size chest X-ray, and chest tomographs were taken, if necessary. Re-examinations of sputum cytology were conducted for those who had abnormal findings for sputum cytology during screening. After these examinations, those who were further suspected as having lung cancer were invited to a second-stage evaluation conducted in the 2 hospitals mentioned above, which consisted of diagnostic procedures such as transbronchial brushing cytology, and/or percutaneous needle aspiration cytology.

Analysis was focused on 33,599 screening tests conducted in 1981–85 for 19,028 screenees who were 40 years old or more at the time of screening. The screening test consisted of 33,490 miniature chest X-ray examinations for 18,992 people and 11,420 sputum cytologies for 7,070 people.

The files of the screenees were linked to the files of the Osaka Cancer Registry with the permission of the Osaka Medical Association (No. 880502), and all screenees were followed up through December 31, 1986. The indexes used to collate the two files were address, sex, and

date of birth. Confirmation of the matched cases was carried out by hand.¹⁰⁾

Information on lung cancer cases at diagnosis was collected by reviewing screening records for screen-detected cases. For interval cases, as much information as possible was collected, with the cooperation of doctors belonging to the Osaka Anti-Lung Cancer Association, by reviewing the files of the Osaka Cancer Registry or medical records in the hospitals where the diagnoses were made. Histologic type was classified according to the WHO lung carcinoma classification.¹¹⁾ Location of the tumor was classified as central type when the tumor was located at the 4th order of bronchus (subsegmental) or more proximal, and as peripheral type when the tumor was located at the 5th order of bronchus or more peripheral.

In order to calculate sensitivity and specificity, the gold standard, which refers to the presence or absence of lung cancer at the time of screening, was determined according to the results of the follow-up. For screen-detected cases, it was assumed that preclinical detectable lung cancer, in other words asymptomatic but detectable by screening, would have existed one year (practically 13 months) before the screening date which led to the diagnosis of lung cancer.¹²⁾ Also, for symptom-detected cases, it was assumed that preclinical detectable lung cancer would have existed one year before the date of diagnosis of lung cancer.

The confidence interval for sensitivity and specificity was calculated by the normal theory method for binomial distribution.¹³⁾

RESULTS

Table I shows the age distribution of screenees, allowing for duplicated counts if they were screened twice or more. The male-to-female ratio of screenees who were examined by miniature chest X-ray and sputum cytology was 1 to 1.8 and 2.7 to 1, respectively. Male screenees were slightly older than female screenees. Within males and females, there was no difference in the age distribu-

Table I. Sex and Age Distributions of Screenees

Age at screening	Male			Female		
	X-ray	Cytology	Either ^{a)}	X-ray	Cytology	Either
40–49	3,228	2,262	3,254	7,620	1,081	7,622
50–59	2,806	2,049	2,839	6,915	1,140	6,918
60–69	3,445	2,512	3,473	4,840	688	4,842
70–79	2,009	1,334	2,019	1,968	192	1,969
80–	350	143	351	309	19	312
Total	11,838	8,300	11,936	21,652	3,120	21,663

a) Number of persons who participated in either X-ray or cytology.

Table II. Distribution of Test Results at Screening

Result	X-ray		Cytology		Either ^{a)}	
	No.	(%)	No.	(%)	No.	(%)
Positive	1,349	(4.0)	332	(2.9)	1,615 ^{b)}	(4.8)
Negative	32,141	(96.0)	11,088	(97.1)	31,984	(95.2)
Total	33,490	(100.0)	11,420	(100.0)	33,599	(100.0)

a) Number of people who participated in either X-ray or cytology.

b) Number of people who were positive by either X-ray or cytology.

tion between those examined by chest X-ray and those by sputum cytology.

Table II shows the distribution of test results for miniature chest X-ray and sputum cytology at screening. The percent of test-positives at screening was 4.0% by chest X-ray, 2.9% by sputum cytology, and 4.8% by the two tests combined. A total of 1,615 people were determined as test-positives at screening, and were asked to undergo further evaluations. The attendance rate at these evaluations was 94.9%.

As a result, 43 lung cancer cases were detected through screening in 1981–85. Of these, 27 cases were detected at initial screening and 16 cases were detected at subsequent screening, where initial screening was defined as the screening that the patient attended for the first time, and subsequent screening was defined as the screening that the patient attended for the second time or subsequently, regardless of the number of years since initial screening. Of 43 screen-detected cases, 29 cases were detected by chest X-ray alone, 10 cases by sputum cytology alone, and 4 cases by both tests.

By linking the files of the screenees to the files of the Osaka Cancer Registry, 24 cases were identified as having been diagnosed without having given a positive screening result in 1981–86, and these are referred to as interval cases. Also, 3 out of 8 additional cases detected at screening in 1986, which was one year after the latest screening under this study, had a history of at least one previous screening.

Table III shows the characteristics of lung cancer cases at diagnosis for screen-detected cases in 1981–85, interval cases in 1981–86, and screen-detected cases in 1986. There was no statistically significant difference in sex distribution between screen-detected and interval cases by Fisher's exact test, although the numbers of subjects were rather limited. The age distribution proved to be slightly higher in interval cases than in screen-detected cases. The proportion of unknown clinical stage was significantly higher for interval cases. These unknown cases were mainly derived from the lung cancer cases who were registered in the Osaka Cancer Registry by

Table III. Distribution of Characteristics at Diagnosis for Lung Cancer Cases According to the Mode of Detection

	Screen-detected ^{a)} (N=43)		Interval (N=24)		Screen-detected ^{b)} (N=3)	
	No.	(%)	No.	(%)	No.	(%)
Sex						
Male	33	(76.7)	20	(83.3)	2	(66.7)
Female	10	(23.3)	4	(16.7)	1	(33.3)
Age						
40–49	0	(0.0)	1	(4.2)	0	(0.0)
50–59	12	(27.9)	5	(20.8)	1	(33.3)
60–69	19	(44.2)	7	(29.2)	1	(33.3)
70–79	11	(25.6)	7	(29.2)	0	(0.0)
80–	1	(2.3)	4	(16.7)	1	(33.3)
Clinical stage ^{c)}						
I	21	(48.8)	3	(12.5)	3	(100.0)
II	2	(4.7)	0	(0.0)		
III	10	(23.3)	3	(12.5)		
IV	10	(23.3)	8	(33.3)		
Unknown	0	(0.0)	10	(41.7)		
Histologic type						
Squamous	19	(44.2)	7	(29.2)	2	(66.7)
Adeno	16	(37.2)	4	(16.7)	1	(33.3)
Small	3	(7.0)	2	(8.3)		
Large	4	(9.3)	2	(8.3)		
Other	1	(2.3)	1	(4.2)		
Unknown	0	(0.0)	8	(33.3)		
Location						
Central ^{d)}	17	(39.5)	6	(25.0)	2	(66.7)
Peripheral ^{e)}	26	(60.5)	3	(12.5)	1	(33.3)
Unknown	0	(0.0)	15	(62.5)		

a) Detected by screening in 1981–1985.

b) Detected by screening in 1986 among the people who had participated in the screening in 1981–1985.

c) UICC 1978.

d) Located at 4th order of bronchus or more proximal.

e) Located at 5th order of bronchus or more distal.

death certificate only. However, considering the fact that all these unknown cases had died within 5 years from diagnosis, it is unlikely that many had been diagnosed in the early stage. Therefore, it can be said that more cases were diagnosed at an earlier stage among screen-detected cases than among interval cases. Regarding histologic type, the proportion of unknown cases was also high among interval cases. Of 8 cases with unknown histologic type, 7 cases died at hospitals with more than 200 beds. The remaining 1 case died at a hospital with 83 beds. If these unknown cases were ignored, there was a tendency for fewer adenocarcinoma (25.0%) cases and more small cell (12.5%) and large cell (12.5%) carcinoma cases to be observed among the interval cases, as compared to screen-detected cases. However, these comparisons of

Table IV. Sensitivity and Specificity for Chest X-Ray and Sputum Cytology at Screening

Test	Result	Cancer		Total	Sensitivity (%) [95%CI] ^{a)}	Specificity (%) [95%CI]
		+	-			
Chest X-ray	+	42	1,307	1,349	56.8	96.1
	-	32	32,109	32,141	[45.5-68.0]	[95.9-96.3]
	Total	74	33,416	33,490		
Cytology	+	15	317	332	30.6	97.2
	-	34	11,054	11,088	[17.7-43.5]	[96.9-97.5]
	Total	49	11,371	11,420		
Either ^{b)}	+	53	1,562	1,615	71.6	95.3
	-	21	31,963	31,984	[61.3-81.9]	[95.1-95.6]
	Total	74	33,525	33,599		

a) 95% confidence interval.

b) Either by X-ray or sputum cytology.

Table V. Sensitivity of Chest X-Ray and Sputum Cytology at Screening According to Histologic Type and Location of the Tumor

Test	Squamous				Adeno	
	Central (N=21)		Peripheral (N=7)		Peripheral (N=22)	
	Result	Se ^{a)}	Result	Se	Result	Se
	+/-	(%)	+/-	(%)	+/-	(%)
Chest X-ray	7/14	33.3	4/3	57.1	19/3	86.4
Cytology	14/6	70.0	1/5	16.7	0/9	0.0
Either ^{b)}	16/5	76.2	5/2	71.4	19/3	86.4

a) Sensitivity.

b) Either by X-ray or sputum cytology.

Table VI. Sensitivity of Chest X-Ray and Sputum Cytology at Screening According to Initial and Subsequent Screening

Test	Initial		Subsequent	
	Result	Se ^{a)}	Result	Se
	+/-	(%)	+/-	(%)
Chest X-ray	26/17	60.5	16/15	51.6
Cytology	10/21	32.3	5/13	27.8
Either ^{b)}	33/10	76.7	20/11	64.5

a) Sensitivity.

b) Either by X-ray or cytology.

previous screening within 1 year of the date of the screening which led to the diagnosis (13 for chest X-ray and 10 for sputum cytology). Out of 24 interval cases in 1981-86, 16 cases had undergone a previous screening within 1 year of the date of the diagnosis (16 for chest X-ray and 8 for sputum cytology). Out of 3 screen-detected cases in 1986, 2 cases had undergone a previous screening within 1 year of the date of the screening which lead to the diagnosis (2 for chest X-ray and 1 for sputum cytology). In total, 74 cases and 49 cases were assumed to have been examined by chest X-ray and sputum cytology, respectively, while having preclinical detectable lung cancer.

Table IV shows the test results and the presence of the disease, for chest X-ray, sputum cytology, and either test, in which test-positives were defined as those who were found positive by either chest X-ray or sputum cytology. As a result, sensitivity was estimated to be 56.8%, 30.6% and 71.6% for chest X-ray, sputum cytology, and either test, respectively; and the specificity was 96.1%, 97.2%, and 95.3%, respectively.

Furthermore, sensitivity was calculated according to the histologic type and the location of the tumor, focusing on the cases for which those characteristics were identified (Table V). There was no central-type adenocarcinoma. For central-type squamous cell carcinoma, sensitivity was higher for sputum cytology (70.0%) than for chest X-ray (33.3%). For peripheral-type squamous cell carcinoma, sensitivity was higher for chest X-ray (57.1%) than for sputum cytology (16.7%). For peripheral-type adenocarcinoma, sensitivity was higher for chest X-ray (86.4%) than for sputum cytology (0.0%). Regardless of the histologic type and the location of the tumor, the sensitivity for either test was over 70%. However, the values of these sensitivities were overes-

clinical stage and histologic type may be biased, if the distributions were different in unknown cases.

Of 43 screen-detected cases in 1981-85, all were examined by chest X-ray and 30 cases were examined by sputum cytology at the screening which led to the diagnosis of lung cancer. Out of these, 13 had undergone a

timated because unknown cases were ignored, especially for interval cases, for which all tests were categorized as false-negatives by definition.

Table VI shows sensitivity according to whether the screening was conducted initially or subsequently. The sensitivity of either test was higher for the initial screening than for subsequent screening, mainly because of chest X-rays.

DISCUSSION

This is the first study in Japan in which sensitivity and specificity of lung cancer screening were estimated from the data obtained from an actual screening program. It was estimated that sensitivity and specificity of annual lung cancer screening from miniature chest X-ray and sputum cytology were 71.6% and 95.3%, respectively. Sensitivity for chest X-ray alone and sputum cytology alone was estimated as 56.8% and 30.6%, respectively, which indicates the complementary qualities of the two tests.

Sensitivity varied mainly according to the location of the tumor. Chest X-rays showed relatively higher sensitivity for peripherally located tumors, while sputum cytology showed relatively higher sensitivity for centrally located tumors.

In this study, the sensitivity of the initial screening was higher than that of subsequent screenings, particularly for chest X-ray. This can be explained by the fact that cancer in the long, preclinical detectable phase will be more prevalent at the initial screening than at subsequent screenings, which will increase the proportion of screen-detected cases as compared to interval cases at the initial screening. On the other hand, however, since previous chest X-rays can be reviewed at subsequent screenings, this might cause the sensitivity value for subsequent screenings to become higher than that at the initial screening. This may be true only if special efforts are made to increase sensitivity, because reviewing previous

X-rays for required cases will usually increase specificity without any change in sensitivity. Also, some methodological problem exist in calculating sensitivity for subsequent screenings, as will be discussed later.

Since no previous data were available in Japan for comparing sensitivity and specificity of lung cancer screening, our data were compared to those reported from Johns Hopkins Lung Project (JHLP), which was one of the randomized controlled trials for evaluating lung cancer screening conducted in the US in the 1970s.³⁾ In the report from JHLP, however, the method used to calculate sensitivity and specificity was different from that in the present study. They assumed that all lung cancer diagnosed after entry would have existed at any screening before the diagnosis, with no regard for the interval between the time of diagnosis and previous screenings. Therefore, we recalculated sensitivity and specificity in this manner, focusing on all lung cancer cases diagnosed up to December 31, 1986 (Table VII). Also, there were some differences between the designs and methods of screening used by JHLP and those used by us. JHLP targeted only middle-aged male smokers, and offered 4-monthly sputum cytologies in addition to annual full-size chest X-rays and sputum cytologies. Despite these differences, almost the same sensitivity and specificity were observed in both studies. It should be noted however, that sensitivity in the JHLP study may be relatively underestimated, because the duration of follow-up after the beginning of the study was longer than in the present study.

In the report from JHLP, a comparison was made of sensitivity and specificity with the ACS-VA Study,¹⁴⁾ which was conducted in 1958–61 as a preliminary study for evaluating the efficacy of lung cancer screening. It was found that there was no substantial difference in sensitivity and specificity of chest X-rays and sputum cytology between the two studies, despite the fact that the JHLP study was conducted 20 years after the ACS-VA Study. This problem, the fact that the demerit

Table VII. Modified Sensitivity of Chest X-Ray and Sputum Cytology in Johns Hopkins Lung Project and in Osaka

	Chest X-ray		Cytology		Either ^{a)}	
	Result	Se(%) ^{b)}	Result	Se(%)	Result	Se(%)
	+/-	[95%CI] ^{c)}	+/-	[95%CI]	+/-	[95%CI]
JHLP	88/89	49.7 [42.4–57.1]	44/133	24.9 [18.5–31.2]	119/58	67.2 [60.3–74.1]
Osaka	33/37	47.1 [35.4–58.8]	13/34	29.2 [16.3–42.0]	43/27	61.4 [50.0–72.8]

a) Either by X-ray or cytology.

b) Sensitivity.

c) The 95% confidence interval.

of lung cancer screening is its low sensitivity, was also pointed out by Sasco.¹⁵⁾

In Japan, there have been many studies on sensitivity and specificity for stomach cancer screening. According to a study by Murakami *et al.*,¹⁶⁾ in which the same methods were used for calculating sensitivity and specificity, the sensitivity and specificity for stomach cancer screening by an image intensifier method, which is currently used in Japan, were estimated as 88.5% and 92.0%. This indicates that stomach cancer screening has a higher sensitivity but lower specificity when compared to present lung cancer screening. In the same study, it was reported that sensitivity and specificity for stomach cancer screening by a mirror camera method, used before the image intensifier method became widespread, were 89.6% and 85.8%. This indicates that stomach cancer screening initially had high sensitivity, allowing for low specificity in earlier periods, but its specificity has improved during its development.

Lung cancer screening has been developed primarily based on the screening system for tuberculosis. Specificity for lung cancer screening is rather high when compared to screenings for other cancers, partly because the specificity of screening for tuberculosis has been high. However, in the case of screening tests for serious diseases, such as cancer, emphasis should be placed on sensitivity rather than specificity. Clearly, much effort is needed to improve sensitivity in lung cancer screening, even at the expense of lowering specificity to some degree.

All 74 miniature chest X-rays which correspond to the tests used in calculating sensitivity and specificity were reviewed by two specialists in radiology who were not involved in the screening trial. In this review, 2 miniature chest X-rays of controls who were confirmed as not having lung cancer at least one year after the screening were also added per each case, matched by sex, age, and year of screening. Out of 32 false-negative chest X-rays, 5 cases were classified as test-positives in this review, which would increase the sensitivity of chest X-rays from 56.8% to 63.5%. This was counterbalanced by a decrease in specificity from 96.1% to 76.4%, which was defined as the proportion of test-positives in the controls. Also, 49 slides for sputum cytology used for calculating sensitivity and specificity were reviewed, but without controls. Out of 34 false-negative sputum cytologies, 6 cases were classified as test-positives in the review, which would increase sensitivity from 30.6% to 42.9%.

Therefore, if these figures are directly used, increasing the sensitivity from 71.6% to 81.1% can be achieved by combining the two tests and allowing for decreased specificity from 95.3% to 76.4%. Of course, since these results were obtained from the review in limited conditions, they cannot be directly applied to actual screen

settings. However, this implies the necessity not only for making further efforts to improve the sensitivity of the present screening system but also for developing new screening techniques which increase the sensitivity, while keeping reductions in specificity to a minimum.

This study contains some methodological points to be discussed. First, the duration of the preclinical detectable phase was assumed to be one year uniformly for all cases of lung cancer, although in actuality it varies case by case. This assumption theoretically causes two types of mistakes. One is that a cancer for which the preclinical detectable phase is shorter than one year might be defined as false-negative at screenings within one year prior to the diagnosis, but this should actually be true-negative in some cases. The other is that a cancer for which the preclinical detectable phase is longer than one year might be defined as true-negative at screenings two or more years prior to the diagnosis, but this should actually be false-negative in some cases.¹⁷⁾ In order to resolve this problem, a method was proposed by Day and Walter¹⁸⁾ in which the mean duration of the preclinical detectable phase and the sensitivity can be simultaneously estimated. However, this method requires sophisticated statistical skills as well as a huge amount of data, 100,000 screenees or more, in order to estimate sensitivity with a sufficiently narrow confidence interval. This is because a strong negative correlation exists between duration of preclinical phase and sensitivity. In the light of this, since screening is usually conducted annually, the method used in the present study is adequate to achieve its practical objective, because all interval cases diagnosed within one year after the screening are considered as missed cases regardless of the duration of the preclinical detectable phase. Moreover, since many studies have been done using this method, it is easy to compare results. Morrison proposed that sensitivity calculated by the method used in this study should be called "program sensitivity," which is different from real sensitivity, but which still has practical significance.¹⁹⁾

Second, the lead time for screen-detected cases was not considered in this study method. For screen-detected cases, the duration in which preclinical but detectable lung cancer exists before the time of detection is shorter than the duration of the preclinical detectable phase by a period of time equal to the lead time, while for symptom-detected cases, it is the duration of the preclinical detectable phase itself. Therefore, it is not appropriate to use the same duration for both screen-detected and symptom-detected cases in which preclinical but detectable lung cancer exists before the time of detection. On the other hand, the duration of the preclinical detectable phase may be longer for screen-detected cases than that for symptom-detected cases because of length-biased sampling. Further research is needed in this area.

Third, the sensitivity calculated by this study method is influenced by the attendance rates in the subsequent screenings, which are conducted annually after the initial screening. For example, if attendance rates for subsequent screenings are 100% and there are no false-negatives during the diagnostic evaluations, the sensitivity of subsequent screenings should theoretically be less than 50% in the long run because all screen-detected cases would be counted twice, as a false-negative and a true-positive, and all interval cases would be counted once as a false-negative. Accepting the assumptions mentioned above, there were 16 lung cancer cases detected at subsequent screenings among this study's subjects, and 5 had no screening history within 1 year before the time of detection. Therefore, if these 5 cases had had previous screening history within 1 year, the sensitivity of subsequent screening would have been 43.2% (16/37). In order to avoid this problem, it would be better to limit our consideration to initial screening for calculating the sensitivity, but this will cause other problems. First, since there has been a long history of screening for tuberculosis using chest X-ray examination in Japan, it is rather difficult to define when the screening started. Second, it will substantially reduce the sample size available for calculating sensitivity. Further discussion is needed on these points.

Fourth, higher sensitivity measured in this method may not directly mean greater efficacy to improve survival of lung cancer patients. Screen-detected cases who had screening history within 1 year before the time of detection ($n=13$) were counted twice as a false-negative and a true-positive, while screen-detected case who did not have screening history within 1 year before the time of detection ($n=30$) were counted once as a true-positive. Therefore, a higher proportion of the latter group will result in higher sensitivity. However, the 5-year survival rate was actually lower in the latter group (26.7%) than in the former group (38.5%), although the difference was not statistically significant by the log rank test. Both groups had higher 5-year survival than that of interval cases (8.3%). In order to avoid this type of controversy, it may be better to count only interval cases as false-negatives, and to count all screen-detected cases once as true-positives.

Fifth, the validity of the diagnosis of lung cancer for interval cases might not be the same as for screen-detected cases. Actually, there were 8 interval cases whose cause of death was described as primary lung cancer on death certificates, but whose histologic types could not be identified. Of these, 3 cases had undergone a previous screening within 1 year of the date of diagnosis (3 for chest X-ray and none for sputum cytology). Medical records in the hospitals where those diagnoses were made were reviewed for 2 interval cases. It was

found that there was no objective evidence of malignancy, such as cytologic examination, for 1 case. For the remaining case, it was determined from the autopsy that the tumor was not primary lung cancer, but rather mesothelioma. When these 3 interval cases were deleted, sensitivity for either chest X-ray and sputum cytology would be raised from 71.6% to 74.6%. In addition, there was one interval case whose histologic type was reported as adenocarcinoma of primary lung cancer, but a review of the medical records suggested that this was metastasis from hepatocellular carcinoma. Therefore, there is the possibility that the value of sensitivity was underestimated when all cases identified by death certificates only were included. However, there is also the possibility that some of the lung cancer cases would have been missed by the follow-up methods used in this study. For example, those who actually had primary lung cancer might have been misdiagnosed as having other disease. Therefore, we decided to include all interval cases without any histologic evidence of malignancy in this study.

A population-based cancer registry was used to find interval cases in this study. Out of 24 interval cases, only 2 were alive on July 31, 1989. This indicates that interval cases for lung cancer screening tend to develop rapidly and to have a poor prognosis. Therefore, even where a population-based cancer registry has not been operating, most of these interval cases in population-based screening programs can be detected by investigating death certificates of all inhabitants if the follow up period is sufficiently extended, such as for 3 to 4 years beyond the latest screening under study.

Sensitivity and specificity are two of the most important measures in the quality control of screening systems. Further effort is needed to measure sensitivity and specificity based on the data from actual screening in various places. By accumulating this information, adequate evaluation of the efficacy of lung cancer screening here in Japan should become possible.

ACKNOWLEDGMENTS

This study was partly supported by a Grant-in-Aid for Cancer Research (62-4) from the Japanese Ministry of Health and Welfare. We would like to thank the Osaka Medical Association, the Osaka Cancer Registry, and the doctors of the Osaka Anti-Lung Cancer Association for providing information concerning our study subjects. We also thank Mr. Y. Kuroda, Mr. M. Nishio, and Mr. Y. Nagai, who participated in data analysis as medical students in a laboratory course provided by the Department of Public Health, School of Medicine, Osaka University, and Ms. N. Nakayama, Ms. C. Inubushi, Ms. Y. Murai and Ms. M. Iwamoto for their technical assistance.

(Received May 27, 1991/Accepted July 19, 1991)

REFERENCES

- 1) Fontana, R. S., Sanderson, D. R., Woolner, L. B., Taylor, W. F., Miller, W. E. and Muhm, J. R. Lung cancer screening: The Mayo Program. *J. Occup. Med.*, **28**, 746-750 (1986).
- 2) Melamed, M. R., Flehinger, B. J., Zaman, M. B., Heelan, R. T., Perchick, W. A. and Martini, N. Screening for early lung cancer. Results of the Memorial Sloan-Kettering Study in New York. *Chest*, **86**, 44-53 (1984).
- 3) Tockman, M. S., Levin, M. L. and Frost, J. K. Screening and detection of lung cancer. In "Lung Cancer, Contemporary Issues in Clinical Oncology, Vol. 3," ed. J. Aisner, pp. 25-40 (1985). Churchill Livingstone, New York.
- 4) Kubik, A., Parkin, D. M., Khlát, M., Erban, J., Polak, J. and Adamec, M. Lack of benefit from semi-annual screening for cancer of the lung: follow-up report of a randomized controlled trial on a population of high-risk males in Czechoslovakia. *Int. J. Cancer*, **45**, 26-33 (1990).
- 5) Berndt, R., Nischan, P. and Ebeling, K. Screening for lung cancer in the middle-aged. *Int. J. Cancer*, **45**, 229-230 (1990).
- 6) Sackett, D. L., Haynes, R. B. and Tugwell, P. "Clinical Epidemiology. A Basic Science for Clinical Medicine," pp. 59-138 (1985). Little Brown, Boston.
- 7) Cole, P. and Morrison, A. S. Basic issues in population screening for cancer. *J. Natl. Cancer Inst.*, **64**, 1263-1272 (1980).
- 8) Saccomano, G., Saunders, R. P., Ellis, H., Archer, V. E., Wood, B. G. and Beckler, P. A. Concentration of carcinoma or atypical cells in sputum. *Acta Cytol.*, **7**, 305-310 (1963).
- 9) The Japan Lung Cancer Society. "General Rule for Clinical and Pathological Record of Lung Cancer," 3rd Ed., pp. 139-155 (1987). Kanehara Shuppan, Tokyo (in Japanese).
- 10) Oshima, A., Sakagami, F., Hanai, A. and Fujimoto, I. A method of record linkage. *Environ. Health Perspect.*, **32**, 221-230 (1979).
- 11) The World Health Organization. Histological typing of lung tumors, second edition. *Am. J. Clin. Pathol.*, **77**, 123-136 (1981).
- 12) Hisamichi, S., Sugawara, N., Fuchigami, A., Aikawa, K., Chuma, Y., Takeuchi, Y., Takahashi, K., Yoshikawa, K., Fujita, R., Iinuma, T., Yamada, T. and Ichikawa, H. False negative rate of stomach cancer in gastric mass survey. *Jpn. J. Cancer Clin.*, **24**, 189-194 (1978) (in Japanese).
- 13) Rosner, B. "Fundamentals of Biostatistics," 2nd Ed., pp. 165-169 (1986). Duxbury Press, Boston.
- 14) Lilienfeld, A., Archer, P. G., Burnet, C. H., Chamberlain, E. W., Chazin, B. J., Davies, D., Davis, R. L., Haber, P. A., Hodges, F. J., Koprowska, B., Kordan, B., Lane, J. T., Lawton, A. H., Lee, L., Jr., Maccallum, D. B., Macdonald, J. R., Milder, J. W., Naylor, B., Papanicolaou, G. N., Slutzker, B., Smith, R. T., Swepston, E. R. and Umiker, W. O. An evaluation of radiologic and cytologic screening for the early detection of lung cancer: a cooperative pilot study of the American Cancer Society and the Veterans Administration. *Cancer Res.*, **26**, 2083-2121 (1966).
- 15) Sasco, A. J. Screening for lung cancer. In "Lectures and Symposia of the 14th International Cancer Congress Vol. 6, Epidemiology, Prevention, Diagnosis," ed. K. Lapis and S. Eckhardt, pp. 177-185 (1987). Akademiai Kiado, Budapest.
- 16) Murakami, R., Tsukuma, H., Ubukata, T., Nakanishi, K., Fujimoto, I., Kawashima, T., Yamazaki, H. and Oshima, A. Estimation of validity of mass screening program for gastric cancer in Osaka, Japan. *Cancer*, **65**, 1255-1260 (1990).
- 17) Day, N. E. Estimating the sensitivity of a screening test. *J. Epidemiol. Community Health*, **39**, 364-366 (1985).
- 18) Day, N. E. and Walter, S. D. Simplified models of screening for chronic diseases: estimation procedures from mass screening programs. *Biometrics*, **40**, 1-14 (1984).
- 19) Morrison, A. S. "Screening in Chronic Disease," pp. 41-47 (1985). Oxford University Press, New York.