

## Funding

Funding to pay the Open Access publication charges for this article was provided by 'The Third-Term Comprehensive 10-Year Strategy for Cancer Control', Ministry of Health, Labour and Welfare, Japan.

## Conflict of interest statement

None declared.

## References

- Sobue T, Ajiki W, Tsukuma H, Oshima A, Hanai A, Fujimoto I. Trend of lung cancer incidence rate by histological type: a population-based study in Osaka, Japan. *Jpn J Cancer Res* 1999;90:6-15.
- Yoshimi I, Ohshima A, Ajiki W, Tsukuma H, Sobue T. A comparison of trends in the incidence rate of lung cancer by histological type in the Osaka Cancer Registry, Japan and in the Surveillance, Epidemiology and End Results Program, USA. *Jpn J Clin Oncol* 2003;33:98-104.
- Thun MJ, Lally CA, Flannery JT, Calle EE, Flanders WD, Heath CW, Jr. Cigarette smoking and changes in the histopathology of lung cancer. *J Natl Cancer Inst* 1997;89:1580-6.
- Devesa SS, Bray F, Vizcaino AP, Parkin DM. International lung cancer trends by histological type: male:female differences diminishing and adenocarcinoma rising. *Int J Cancer* 2005;117:294-9.
- Janssen-Heijnen MLG, Coebergh JWW. The changing epidemiology of lung cancer in Europe. *Lung Cancer* 2003;41:245-58.
- Marugame T, Sobue T. Mortality trend of mouth and pharynx, esophagus, stomach, larynx and lung cancer in Japan by birth-cohort. *Jpn J Clin Oncol* 2004;34:432-8.
- Soda H, Oka M, Soda M, Nakatomi K, Kawabata S, Suenaga M, et al. Birth-cohort effects on incidence of lung cancers: a population-based study in Nagasaki, Japan. *Jpn J Cancer Res* 2000;91:960-5.
- Kaneko S, Ishikawa KB, Yoshimi I, Marugame T, Hamashima C, Kamo K, et al. Projection of lung cancer mortality in Japan. *Cancer Sci* 2003;94:919-23.
- Trends in lung cancer mortality in selected countries. IARC handbooks of cancer prevention, Tobacco control volume 11. International Agency for Research on Cancer, World Health Organization 2007,307-22.
- Tsukuma H, Ajiki W, Oshima A. Cancer incidence in Japan. *Gan To Kagaku Ryoho* 2004;31:840-6 (in Japanese).
- Marugame T, Kamo K, Sobue T, Akiba S, Mizuno S, Satoh H, et al. Trends in smoking by birth-cohorts born between 1900 and 1977 in Japan. *Preventive Med* 2006;42:120-7.
- Stellman SD, Muscat JE, Thompson S, Hoffmann D, Wynder EL. Risk of squamous cell carcinoma and adenocarcinoma of the lung in relation to lifetime filter cigarette smoking. *Cancer* 1997;80:382-8.
- Marugame T, Sobue T, Nakayama T, Suzuki T, Kuniyoshi H, Sunagawa K, et al. Filter cigarette smoking and lung cancer risk; a hospital-based case-control study in Japan. *Br J Cancer* 2004;90:646-51.
- Hanai A, Benn T, Fujimoto I, Muir CS. Comparison of lung cancer incidence rates by histological type in high and low incidence countries, with reference to the limited role of smoking. *Jpn J Cancer Res* 1988;79:445-52.
- Marugame T, Yoshimi I, Kamo K, Imamura Y, Kaneko S, Mizuno S, et al. Trends in lung cancer mortality among young adults in Japan. *Jpn J Clin Oncol* 2005;35:177-80.
- Curado MP, Edwards B, Shin HR, Storm H, Ferlay J, Heanue M, et al. Cancer Incidence in Five Continents. Vol. IX. Lyon: IARC 2007, IARC Scientific Publications No. 160. Available at [http://www-dep.iarc.fr/CIS-IX/PDF/INDICES/I\\_07.pdf](http://www-dep.iarc.fr/CIS-IX/PDF/INDICES/I_07.pdf)

# Sensitivity and specificity of lung cancer screening using chest low-dose computed tomography

Y Toyoda<sup>1,2</sup>, T Nakayama<sup>\*,1</sup>, Y Kusunoki<sup>1,3</sup>, H Iso<sup>2</sup> and T Suzuki<sup>1</sup>

<sup>1</sup>Department of Cancer Control and Statistics, Osaka Medical Center for Cancer and Cardiovascular Diseases, Osaka, Japan; <sup>2</sup>Public Health, Department of Social and Environmental Medicine, Graduate School of Medicine, Osaka University, Osaka, Japan; <sup>3</sup>Osaka Prefectural Medical Center for Respiratory and Allergic Diseases, Osaka, Japan

Lung cancer screening programmes using chest X-ray and sputum cytology are routinely performed in Japan; however, the efficacy is insufficient. Screening using low-dose computed tomography (CT) is a more effective approach and has the potential to detect the disease more accurately. A total of 7183 low-dose CT screening tests for 4689 participants and 36 085 chest X-ray screening tests for 13 381 participants were conducted between August 1998 and May 2002. Sensitivity and specificity of lung cancer screening were calculated by both the detection method and the incidence method by linkage of the screening database and the Cancer Registry database. The preclinical detectable phase was assumed to be 1 year. Sensitivity and specificity by the detection method were 88.9 and 92.6% for low-dose CT and 78.3 and 97.0% for chest X-ray, respectively. Sensitivity of low-dose CT by the incidence method was 79.5%, whereas that of chest X-ray was 86.5%. Lung cancer screening using low-dose CT resulted in higher sensitivity and lower specificity than traditional screening according to the detection method. However, sensitivity by the incidence method was not as high as this. These findings demonstrate the potential for overdiagnosis in CT screening-detected cases.

*British Journal of Cancer* (2008) **98**, 1602–1607. doi:10.1038/sj.bjc.6604351 www.bjcancer.com

Published online 6 May 2008

© 2008 Cancer Research UK

**Keywords:** sensitivity; specificity; lung cancer; screening

Lung cancer is the leading cause of cancer death in Japan, with 45 927 men and 17 307 women dying from lung cancer in 2006. Since 1987, lung cancer screening programme using chest X-ray and sputum cytology for all residents aged 40 years of age and older regardless of smoking status has been conducted by the Ministry of Health and Welfare. Unfortunately, the efficacy of lung cancer screening using chest X-ray and sputum cytology is insufficient (Fontana *et al*, 1986; Marcus *et al*, 2000a, 2006b; Sagawa *et al*, 2003a). Therefore, a more effective approach is required to decrease lung cancer deaths.

Annual lung cancer screening using low-dose computed tomography (CT) has been performed as an opportunistic screening method since the early 1990s in Japan. Several study groups introduced low-dose CT for population-based screening in clinical trials. These previous studies reported a high detection rate, an ability to detect small tumours and a high survival rate in detected cases (Henschke *et al*, 2001, 2006; Sone *et al*, 2001; Nawa *et al*, 2002; Sobue *et al*, 2002a; Swensen *et al*, 2002; Diederich *et al*, 2004; Jett, 2005; Libby *et al*, 2006). Some studies referred to interval cancer cases of lung cancer screening using low-dose CT, and one study referred to the sensitivity of screening (Sone *et al*, 2001; Diederich *et al*, 2004). However, screening databases are yet to be linked to a cancer registry, which is essential for accurate evaluation of screening, including the confirmation of all interval cancer cases. To date, no study has been conducted on sensitivity

and specificity of annual lung cancer screening using low-dose CT and cancer registry data. Therefore, the present study was conducted to evaluate sensitivity and specificity of annual lung cancer screening using low-dose CT and data from screening and local cancer registry databases.

## MATERIALS AND METHODS

### Study setting

Since 1998, annual population-based lung cancer screening using low-dose CT has been conducted at five municipalities in Osaka prefecture: A (city), B (city), C (town), D (town) and E (town). All residents aged 40 years of age and older were recruited by mail using a letter from the public health division of each municipality regardless of smoking status. Subjects recruited to the lung cancer screening programme underwent either miniature chest X-ray or low-dose chest CT.

As a principle, heavy smokers were recommended to undergo low-dose CT screening. In addition, the persons who want to undergo low-dose CT screening also underwent low-dose CT screening. Others underwent chest X-ray screening.

A high-risk group for lung cancer, smokers with over a 20 pack index or who had haemosputum, was examined by 3-day pooled sputum cytology.

Low-dose CT or chest X-ray images were reviewed and classified by two trained physicians to determine the need for further clinical examination. Sputum cytology was also performed by a certified cytopathologist to determine the need for further clinical examination.

\*Correspondence: Dr T Nakayama;

E-mail: nakayama-to@mc.pref.osaka.jp

Received 24 September 2007; revised 18 February 2008; accepted 26 March 2008; published online 6 May 2008



Those diagnosed with the need for further clinical examination were regarded as screen-positive. These individuals were asked to undergo further diagnostic evaluation at Osaka Medical Center for Cancer and Cardiovascular Disease. All individuals with positive chest X-ray screening were asked to undergo chest CT as a further examination.

### Data collection

All subjects were individuals who had undergone either low-dose CT or chest X-ray screening tests between August 1998 and May 2002. The following participants were excluded from the analyses: (1) participants who had a past history of lung cancer, (2) participants who were suspected of having lung cancer by a previous screening or other medical examination and had received medical treatment and (3) participants who were suspected of having lung cancer at the previous screening or by other medical examination, but had refused further examinations.

Participants were divided into two groups: (a) low-dose CT group and (b) chest X-ray group. The low-dose CT group consisted of persons who had undergone low-dose CT at least one time during the study period, whereas the chest X-ray group consisted of persons who had undergone only chest X-ray. The low-dose CT group included those who had undergone both CT screening and chest X-ray screening within the study period. For these cases, screenings using chest X-ray were ignored to evaluate low-dose CT screening.

All data were entered into the screening database that was linked to the Osaka Cancer Registry (OCR) database with data reflecting incidence cases through December 2003. The indices used to collate the two databases were name, sex, address and date of birth. Information about lung cancer cases was extracted from hospital medical records or the OCR file.

We assumed that the preclinical detectable phase was 1 year for interval cancer cases. For death certificate-only cases, the date of 3 months before death was regarded as the date of diagnosis. Using these parameters, all lung cancer cases diagnosed within 1 year after a negative screen were regarded as interval lung cancers. Screen-detected cases were considered as true-positive cases regardless of the time between the date of screening and the date of diagnosis.

### Statistical analyses

The sensitivity of screening was calculated by both the detection method and the incidence method. Although the detection method is simple and widely used, sensitivity estimated by detection method is affected by length and overdiagnosis biases (Day, 1985). The incidence method is not affected by length or overdiagnosis bias and is often used for breast cancer screening or colorectal cancer screening (Fletcher *et al*, 1993; Zappa *et al*, 2001).

### Detection method

Sensitivity and specificity were calculated by the detection method using the following formulae.

	True disease state	
	+	-
Screening test	+ a	b
	- c	d
	Sensitivity = $a/(a + c)$	
	Specificity = $d/(b + d)$	

Sensitivity and specificity calculated by the detection method were stratified by smoking status, histological type and screening rank. The screening rank was classified as the initial and repeated

screenings, regardless of the number of years since the initial screening.

### Incidence method

In addition, we calculated sensitivity by the incidence method using the following approximate formula (Day, 1985; Zappa *et al*, 2001):

$$\text{Sensitivity} = 1 - [I(t)/I]$$

Where  $I(t)$  = the observed number of interval cancer cases during time  $t$  and  $I$  = the expected number of cases in the absence of screening.

We calculated the number of expected lung cancer cases in the absence of screening based on the following data. Age-specific lung cancer incidence rates provided from the OCR in 2001 were 16.3, 61.6, 180.9, 477.3 and 770.2 (per 100 000 person-years) for men, and 6.3, 25.9, 53.4, 116.7 and 241.3 for women, for age groups 40–49, 50–59, 60–69, 70–79 and  $\geq 80$ , respectively. Lung cancer incidence rates in the OCR were weighted by smoking status. According to the previous large-scale cohort study in Japan, the lung cancer incidence rates among ex-smokers and current smokers were assumed to be 2.2 times and 4.5 times of that of nonsmokers, respectively, among men, and 3.7 times and 4.2 times, respectively, among women (Sobue *et al*, 2002b). According to an official report from Osaka prefecture in 2003, the proportions of current smokers, ex-smokers and nonsmokers were 40, 30 and 30% among men, and 11, 7 and 82% among women, respectively (Department of public health, Osaka Prefecture, 2006).

We assumed that smoking status proportions were the same across all age groups, so the expected incidence rate according to sex and smoking status was modified using the following formulae:

$$\begin{aligned} \text{Expected incidence rate among male nonsmokers} \\ = \text{Incidence rate in OCR}/(4.5 \times 0.40 + 2.2 \times 0.30 + 1 \times 0.30) \end{aligned}$$

$$\begin{aligned} \text{Expected incidence rate among female nonsmokers} \\ = \text{Incidence rate in OCR}/(4.2 \times 0.11 + 3.7 \times 0.07 + 1 \times 0.82) \end{aligned}$$

The expected incidence rates for ex-smokers and current smokers were assumed to be 2.2 times and 4.5 times of that of nonsmokers, respectively, among men, and 3.7 times and 4.2 times, respectively, among women.

The differences in sensitivity and specificity among the stratified variables were tested by  $\chi^2$  test. All statistical analyses were performed using SAS software, version 8.01 (SAS Institute Inc, Cary, NC, USA).

### Ethical approval

The protocol for the present study was approved by the Ethics Committee of Osaka Medical Center for Cancer and Cardiovascular Disease, Osaka, Japan. Informed consent for participation in the clinical trial, including CT screening, was obtained from all individuals.

### RESULTS

From August 1998 to May 2002, a total of 7190 low-dose CT screening tests and a total of 36 085 chest X-ray screening tests were performed. Seven screening participants were excluded from analysis because they did not meet the eligibility criteria. Participants were ineligible for the following reasons: two participants were under follow-up care, one was suspected of having lung cancer but refused further examination and four had a history of lung cancer. A total of 7183 low-dose CT screening tests for 4689 participants (2765 men and 1924 women) and 36 085 chest



X-ray screening tests for 13 381 participants (4180 men and 9201 women) enrolled in the study.

Table 1 shows the number of screening tests by sex, age group, smoking status and rank of screening tests. Most of the

**Table 1** Number of screening tests performed by age group, smoking status and rank; (a) low-dose CT group and (b) chest X-ray group: Osaka, 1998–2002

	Male	Female	Total
<b>(a) Low-dose CT group</b>			
Age (years)			
40–49	700	490	1190
50–59	1147	1132	2279
60–69	1885	886	2771
70–79	690	194	884
80–	43	16	59
Smoking status			
Nonsmoker	362	2048	2410
Ex-smoker	1012	113	1125
Current smoker	3091	557	3648
Rank			
Initial	2765	1924	4689
Repeated	1700	794	2494
Total	4465	2718	7183
<b>(b) Chest X-ray group</b>			
Age (years)			
40–49	1258	4862	6120
50–59	1679	8632	10311
60–69	4163	7910	12073
70–79	2695	3670	6365
80–	573	643	1216
Smoking status			
Nonsmoker	2807	23790	26597
Ex-smoker	4328	740	5068
Current smoker	3233	1187	4420
Rank			
Initial	4180	9201	13381
Repeated	6188	16516	22704
Total	10368	25717	36085

CT = computed tomography.

participants who underwent low-dose CT screening were male current smokers or ex-smokers. Sputum cytology was additionally performed for 3539 screening tests for the low-dose CT group and 5417 screening tests for the chest X-ray group.

Forty cases in the low-dose CT group and 29 cases in the chest X-ray group were detected by the screening. Five interval cases in the low-dose CT group and eight interval cases in the chest X-ray group were confirmed by linkage to OCR (Table 2). All of the interval cancer cases for both the low-dose CT group and the chest X-ray group were smokers. As for the low-dose CT group, all of them were nonadenocarcinoma. Two cases, one in the low-dose CT group and one in the chest X-ray group, were detected by sputum cytology on negative radiological screen.

Table 3 shows sensitivity and specificity by the detection method according to histological type, smoking status and rank of screening. As a result, sensitivity and specificity (95% confidence interval) of screening were 88.9% (79.7–98.1%) and 92.6% (92.0–93.2%) for the low-dose CT group, and 78.3% (65.1–91.6%) and 97.0% (96.9–97.2%) for the chest X-ray group, respectively. Specificity of chest X-ray screening was significantly higher than that of low-dose CT screening ( $P < 0.001$ ). The difference in sensitivity by the detection method was not significant.

As for histological type, sensitivity for adenocarcinoma was significantly higher than that for nonadenocarcinoma (low-dose CT: 100 vs 61.5%;  $P < 0.001$ , and chest X-ray: 95.8 and 50.0%;  $P < 0.001$ ); however, the histological type of three interval cases in the chest X-ray group was unknown. As for screening rank, specificity for the repeated screenings was significantly higher than that for the initial screenings (low-dose CT: 95.7 vs 91.0%;  $P < 0.001$ , and chest X-ray: 97.7 vs 95.9%;  $P < 0.001$ ). As for sex, specificity for men was significantly lower than that for women (low-dose CT: 92.1 vs 93.5%;  $P < 0.05$ , and chest X-ray: 95.7 vs 97.6%;  $P < 0.001$ ). Sensitivity of chest X-ray screening for women was significantly higher than that for men (100 vs 68.2%;  $P < 0.05$ ). As for smoking status, sensitivity of both low-dose CT and chest X-ray for nonsmokers was 100%.

Table 4 shows sensitivity estimated by the incidence method. Until the end of December 2003, a total of 14 434 person-years (total for men: 9173 person-years; total for women: 5512 person-years) for the low-dose CT group and a total of 59 725 person-years (total for men: 17 962 person-years; total for women: 41 763 person-years) for the chest X-ray group had been followed up for. The mean follow-up terms were 3.1 person-years and 4.5 person-years, respectively. The number of expected lung cancer cases was calculated to be 24.4 persons for the low-dose CT group and 59.3 persons for the chest X-ray group. As a result, sensitivity

**Table 2** Interval cancer cases of screening; (a) low-dose CT group and (b) chest X-ray group

	Sex	Age (years)	Pack index	Smoking status	Histological type	Location	Rank	Clinical stage
<b>(a) Low-dose CT group</b>								
1	F	71	48	Current	Squamous	Unknown	Initial	III
2	M	60	43	Current	Large cell	Peripheral	Initial	III
3	M	72	48	Current	Small cell	Unknown	Repeated	IV
4	M	72	45	Current	Squamous	Unknown	Repeated	I
5 <sup>a</sup>	F	59	29	Ex	Squamous	Central	Initial	I
<b>(b) Chest X-ray group</b>								
1	M	68	48	Current	Squamous	Unknown	Repeated	III
2	M	83	61	Current	Small cell	Unknown	Repeated	Unknown
3	M	72	21	Ex	Adeno	Unknown	Repeated	I
4	M	69	25	Current	Undifferentiated	Unknown	Initial	III
5	M	60	50	Current	Unknown	Unknown	Initial	Unknown
6 <sup>a</sup>	M	63	68	Ex	Squamous	Central	Repeated	I
7	M	59	80	Current	Unknown	Unknown	Repeated	Unknown
8	M	85	15	Ex	Unknown	Unknown	Repeated	Unknown

CT = computed tomography. <sup>a</sup>Detected by sputum cytology.

**Table 3** Sensitivity and specificity by the detection method according to histological type, smoking status and rank of screening: (a) low-dose CT group and (b) chest X-ray group

	No. of screenings	Screen-detected cases	Interval cases	Sensitivity (%) (95% CI)	Specificity (%) (95% CI)
<b>(a) Low-dose CT group</b>					
Sex					
Men	4465	29	3	90.6 (80.5–100)	92.1 (91.3–92.9)
Women	2718	11	2	84.6 (65.0–100)	93.5 (92.6–94.4)
Smoking status					
Nonsmoker	2410	13	0	100	93.5 (92.5–94.4)
Ex-smoker	1125	6	1	85.7 (59.8–100)	91.5 (89.9–93.1)
Current smoker	3648	21	4	84.0 (69.6–98.4)	92.4 (91.6–93.3)
Histological type					
Adenocarcinoma	—	32	0	100	—
Nonadenocarcinoma	—	8	5	61.5 (35.1–88.0)	—
Rank					
Initial	4688	32	3	91.4 (82.2–100)	91.0 (90.2–91.8)
Repeated	2494	8	2	80.0 (55.2–100)	95.7 (94.9–96.5)
Total	7183	40	5	88.9 (79.7–98.1)	92.6 (92.0–93.2)
<b>(b) Chest X-ray group</b>					
Sex					
Men	10368	15	8	65.2 (45.8–84.7)	95.7 (95.3–96.1)
Women	25717	14	0	100	97.6 (97.3–97.8)
Smoking status					
Nonsmoker	26597	13	0	100	97.4 (97.3–97.7)
Ex-smoker	5068	4	3	57.1 (20.5–93.8)	95.9 (95.3–96.4)
Current smoker	4420	12	5	70.6 (48.9–92.2)	95.7 (95.1–96.3)
Histological type					
Adenocarcinoma	—	23	1	95.8 (87.8–100)	—
Nonadenocarcinoma	—	6	4	50.0 (21.7–78.3)	—
Unknown	—	0	3	—	—
Rank					
Initial	13381	13	2	86.7 (69.5–100)	95.9 (95.5–96.2)
Repeated	22704	16	6	76.2 (58.0–94.4)	97.7 (97.5–97.9)
Total	36085	29	8	78.3 (65.1–91.6)	97.0 (96.9–97.2)

CI = confidence interval; CT = computed tomography.

(95% confidence interval) estimated by the incidence method was 79.5% (63.5–95.5%) and 86.5% (77.8–95.2%), respectively. The difference in sensitivity by the incidence method was not statistically significant.

## Discussion

The present study is the first report on sensitivity and specificity of annual lung cancer screening using low-dose CT and data from a local Cancer Registry. Sensitivity and specificity of low-dose CT screening according to the detection method were 88.9 and 92.6%. The sensitivity estimated by the incidence method resulted in a value of 79.5%. On the other hand, sensitivity and specificity of chest X-ray in the same time frame by the detection method were 78.3 and 97.0%, respectively. Furthermore, sensitivity of chest X-ray screening by the incidence method was 86.5%.

In previous studies conducted in the 1980s, sensitivity and specificity of annual lung cancer screening using chest X-ray and sputum cytology were also evaluated by the detection method. In those studies, sensitivity and specificity for usual screening were 63.6–88.0% and 94.7–99.6%, respectively (Sobue *et al*, 1991; Soda *et al*, 1993; Sagawa *et al*, 1994b; Tsukada *et al*, 2002). The use of

low-dose CT screening resulted in a higher sensitivity and lower specificity than usual screening. The reported high sensitivity in participants undergoing low-dose CT screening is the result of improvement in the detection of small tumours. The lower specificity value indicates the difficulty of diagnosing nodules detected by screening.

Several points must be considered when the present study results are compared with previous results. Since 1980s, lung cancer incidence by histological type has undergone a change over time. With a large decline in the smoking rate among men, the proportion of squamous cell carcinoma or small cell carcinoma has decreased, whereas the proportion of adenocarcinoma has increased (Yoshimi *et al*, 2003). The current environment may be more advantageous for lung cancer screening because adenocarcinoma occurring in the peripheral lung has a longer doubling time than squamous cell carcinoma (Arai *et al*, 1994). In addition, as most low-dose CT screening-detected lung cancer lesions are too small to detect by chest X-ray and have a longer preclinical phase, simple comparison of low-dose CT screening with chest X-ray screening is difficult.

We used the detection method and stratified analyses by screening rank and histological type. As for screening rank, specificity of both low-dose CT and chest X-ray for the repeated



**Table 4** Sensitivity of screening by the incidence method; (a) low-dose CT group and (b) chest X-ray group

	Person-years	Expected incidence	Screen-detected cases	Interval cases	Sensitivity (%) (95% CI)
<i>(a) Low-dose CT group</i>					
<i>Sex</i>					
Men	9173	21.8	29	3	86.2 (71.8–100)
Women	5512	2.5	11	2	20.0 (0–69.6)
<i>Smoking status</i>					
Nonsmokers	4878	1.7	13	0	100
Ex-smokers	2388	4.2	6	1	76.2 (35.5–100)
Current smokers	7419	18.6	21	4	78.5 (59.8–97.2)
Total	14 685	24.4	40	5	79.5 (63.5–95.5)
<i>(b) Chest X-ray group</i>					
<i>Sex</i>					
Men	17 962	42.1	15	8	81.0 (69.1–92.8)
Women	41 763	17.2	14	0	100
<i>Smoking status</i>					
Nonsmokers	42 976	17.4	13	0	100
Ex-smokers	8452	19.2	4	3	84.3 (68.1–100)
Current smokers	8297	22.8	12	5	78.1 (61.1–95.1)
Total	59 725	59.3	29	8	86.5 (77.8–95.2)

CI = confidence interval; CT = computed tomography.

screenings was significantly higher than that of the initial screenings. The high specificity associated with repeated screenings is due to the fact that the review of previous images facilitates ruling out benign nodules. Sensitivity of low-dose CT and chest X-ray for the repeated screenings was lower than that of the initial screenings; however, the difference was not statistically significant. Sensitivity for the initial screenings was affected by length bias and overestimation because lung cancers with long preclinical detectable phases were more prevalent. Regarding histological type, adenocarcinoma sensitivity estimated by the detection method was significantly higher than that for nonadenocarcinoma for both low-dose CT and chest X-ray. In the previous study, sensitivity of chest X-ray was 86.4% for adenocarcinoma and 44.2% for nonadenocarcinoma (Sobue *et al*, 1991c). Both low-dose CT screening and chest X-ray screening have a high sensitivity for the detection of adenocarcinoma. In contrast, sensitivity estimated by the detection method for nonadenocarcinoma remained low. As for smoking status, both low-dose CT and chest X-ray had superior performance for nonsmokers.

Although the detection method is simple and widely used, it is affected by overdiagnosis or length bias because cancers with long preclinical detectable phases are included in the denominator. In the 1980s, lung cancer was considered to be an aggressive and rapid-growing cancer; however, it has been reported that low-dose CT screening-detected lung cancer has a long doubling time and good prognosis (Sone *et al*, 2001; Nawa *et al*, 2002; Sobue *et al*, 2002a; Swensen *et al*, 2002; Henschke *et al*, 2006; Libby *et al*, 2006). The incidence method, which is not affected by overdiagnosis bias and length bias, is preferred for the correct evaluation of low-dose CT screening. Screening for breast cancers or colorectal cancers, with long doubling times, has been evaluated using the incidence method whereas lung cancer screening has been evaluated using the detection method only (Fletcher *et al*, 1993; Zappa *et al*, 2001).

In this study, we calculated expected lung cancer incidence to be 24.4 persons for the low-dose CT group and 59.3 persons for the chest X-ray group according to age-specific lung cancer incidence rate in the OCR, smoking status in Osaka prefecture and the relative risk of lung cancer incidence associated with smoking according to a large-scale cohort study in Japan. Unexpectedly, the sensitivity of low-dose CT screening estimated by the incidence

method (79.5%) was lower than that of chest X-ray screening (86.5%); however, the difference was not statistically significant. There are several possible explanations for this contradiction. First, the mean follow-up term of the low-dose CT group (3.1 person-years) was shorter than that of the chest X-ray group (4.5 person-years). Furthermore, the mean pack index of current smokers among the low-dose CT group (42 for men and 23 for women) was somewhat higher than that of the chest X-ray group (38 for men and 16 for women). Therefore, expected lung cancer incidence for the low-dose CT group might be underestimated. Second, four screen-detected cases among the chest X-ray group were checked with lesions other than cancer. These lung cancer cases were incidentally detected by the subsequent chest CT as a further examination on positive tests; all of them were adenocarcinoma. When these cases were regarded as interval cases, sensitivity (95% confidence interval) of chest X-ray screening by the incidence method resulted in 79.7% (69.5–90.0%). Considering these points, sensitivity of low-dose CT screening according to the incidence method with 3–5 person-years of follow-up period would be almost equal to that of chest X-ray screening. These findings suggest that the efficacy of low-dose CT screening might be limited to rapid-growing lung cancer with a short preclinical detectable phase ( $\leq 1$  year). Since low-dose CT screening-detected lung cancer is slow growing, further research with a longer follow-up period is required.

A total of 40 lung cancer cases were detected by low-dose CT screening, suggesting the possibility of overdiagnosis by low-dose CT screening. In particular, low-dose CT screening detected 13 lung cancer cases in nonsmokers whereas expected incidence in nonsmokers was only 1.7 persons. All of these cases were peripheral adenocarcinoma. In contrast, expected lung cancer incidence for the chest X-ray group was higher than the number of screen-detected cases. This fact might suggest that there is little possibility of overdiagnosis by chest X-ray screening.

Of the five interval cancer cases in the low-dose CT group, four cases were squamous cell carcinoma or small cell carcinoma, which are strongly associated with smoking (Sobue *et al*, 1991d; Shimizu *et al*, 1994; Stellman *et al*, 2001). Three cases had remarkable emphysema. These interval cancer cases associated with smoking indicate the limitation of low-dose CT screening for



nonadenocarcinoma among smokers. In other words, the high sensitivity of low-dose CT screening identified using the detection method is due to the detection of adenocarcinoma with a long preclinical detectable phase.

This study has some limitations. First, many nodules were detected by low-dose CT screening, but subsequent pathological examinations were not performed. In this study, small pure ground-glass opacity nodules (<10 mm) were carefully observed, and no invasive treatment was performed. In these cases, lung cancer was highly suspected, but a lung cancer diagnosis was not made and the cases were not registered in the OCR. Given the presence of such cases, the sensitivity according to the detection method might be underestimated. Second, to compare usual screening with low-dose CT screening, the preclinical detectable phase was assumed to be 1 year. We need to assess a longer preclinical detectable phase, because most of the low-dose CT screening-detected lung cancer cases were slow growing. Third, the sample size was relatively small for proper evaluation, particularly for stratified analyses.

In summary, the present findings suggest that lung cancer screening using low-dose CT has a higher sensitivity and a lower specificity than usual lung cancer screening by chest X-ray, when using the detection method analysis. However, sensitivity estimated by the incidence method was not as high as that estimated by detection method. As all interval cancer cases were associated with smoking, low-dose CT screening showed limited efficacy for nonadenocarcinoma in smokers. Furthermore, these findings demonstrate the potential for overdiagnosis in low-dose CT screening-detected cases.

## ACKNOWLEDGEMENTS

This study was supported by a grant-in-aid from the Ministry of Health, Labor and Welfare of Japan for a Third term Comprehensive Control Research for Cancer (Grant no. 19141901). We would like to thank the staff of the Division of Epidemiology and the Osaka Cancer Registry, Department of Cancer Control and Statistics, Osaka Medical Center for Cancer and Cardiovascular Disease.

## REFERENCES

- Arai T, Kuroishi T, Saito Y, Kurita Y, Naruke T, Kaneko M, The Japanese lung cancer research group (1994) Tumor doubling time and prognosis in lung cancer: evaluation from chest films and clinical follow-up study. *Jpn J Clin Oncol* 24: 199–204
- Day NE (1985) Estimating the sensitivity of a screening test. *J Epidemiol Community Health* 39: 364–366
- Department of public health, Osaka Prefecture (2006) *Smoking, Kenko Osaka 21 Report*, pp 26–64. Osaka Prefecture: Osaka
- Diederich S, Thomas M, Semik M, Lenzen H, Roos N, Weber A, Heindel W, Wormanns D (2004) Screening for early lung cancer with low-dose spiral computed tomography: results of annual follow-up examinations in asymptomatic smokers. *Eur Radiol* 14: 691–702
- Fletcher SW, Black W, Harris R, Rimer BK, Shapiro S (1993) Report of the international workshop on screening for breast cancer. *J Natl Cancer Inst* 85: 1644–1656
- Fontana RS, Sanderson DR, Woolner LB, Taylor WF, Miller WE, Muhm JR (1986) Lung cancer screening: the Mayo Program. *J Occup Med* 28: 746–750
- Henschke CI, Naidich DP, Yankelevitz DF, McGinness G, McCauley DI, Smith JP, Libby D, Pasmantier M, Koizumi J, Flieder D, Altorki N, Miettinen OS (2001) Early lung cancer action project: initial findings on repeat screenings. *Cancer* 92: 153–159
- Henschke CI, Yankelevitz DF, Libby DM, Pasmantier MW, Smith JP, Miettinen OS (2006) Survival of patients with Stage I lung cancer detected on CT screening. *New Eng J Med* 355: 1763–1772
- Jett JR (2005) Limitation of screening for lung cancer with low-dose spiral computed tomography. *Clin Cancer Res* 11: 4988–4992
- Libby DM, Wu N, Lee IJ, Farooqi A, Smith JP, Pasmantier MW, MacCauley D, Yankelevitz DF, Henschke CI (2006) CT screening for lung cancer: the value of short-term CT follow up. *Chest* 129: 1039–1042
- Marcus PM, Bergstralh EJ, Fagerstrom RM, Williams DE, Fontana R, Taylor WF, Prorok PC (2000a) Lung cancer mortality in Mayo Lung Project: impact of extended follow-up. *J Natl Cancer Inst* 92: 1308–1316
- Marcus PM, Bergstralh EJ, Zweig MH, Harris A, Offord KP, Fontana RS (2006b) Extended lung cancer incidence follow-up in the Mayo Lung Projects and overdiagnosis. *J Nat Cancer Inst* 98: 748–756
- Nawa T, Nakagawa T, Kusano S, Kawasaki Y, Sugawara Y, Nakata H (2002) Lung cancer screening using low-dose spiral CT: results of baseline and 1-year follow up studies. *Chest* 122: 15–20
- Sagawa M, Nakayama T, Tsukada H, Nshii K, Baba T, Kurita Y, Saito Y, Kaneko M, Sakuma T, Suzuki T, Fujimura S (2003a) The efficacy of lung cancer screening conducted in 1990s: four case-control studies in Japan. *Lung Cancer* 41: 29–36
- Sagawa M, Saito Y, Takahashi S, Endo C, Uda K, Kamma K, Sato M, Ohkuda K, Sato H, Fujimura S (1994b) Sensitivity and specificity of lung cancer screening with sputum cytology and chest X-ray film in high risk group. *Hagan* 34: 1–5 (in Japanese)
- Shimizu H, Nagata C, Tsuchiya E, Nakagawa K, Weng SY (1994) Risk of lung cancer among cigarette smokers in relation to tumor location. *Jpn J Cancer Res* 85: 1196–1199
- Sobue T, Moriyama N, Kaneko M, Kusumoto M, Kobayashi T, Tsuchiya R, Kakimura R, Ohmatsu H, Nagai K, Nishiyama H, Matsui E, Eguchi K (2002a) Screening for lung cancer with low-dose helical computed tomography: anti-lung cancer project. *J Clin Oncol* 20: 911–920
- Sobue T, Suzuki T, Fujimoto I, Matsuda M, Doi O, Mori T, Furuse K, Fukuoka M, Yasumitsu T, Kuwahara O, Ichitani M, Taki T, Kuwabara M, Nakahara K, Endo S, Sawamura K, Kurata M, Hattori S (1991d) Lung cancer risk among ex-smokers. *Jpn J Cancer Res* 82: 273–279
- Sobue T, Suzuki T, Matsuda M, Horai T, Kajita A, Kuriyama K, Fukuoka M, Kusunoki Y, Kikui M, Ryu S, Fujimoto I (1991c) Sensitivity and specificity of lung cancer screening in Osaka, Japan. *Jpn J Cancer Res* 82: 1069–1076
- Sobue T, Yamamoto S, Hara M, Sasaduki S, Sasaki S, Tsugane S (2002b) Cigarette smoking and subsequent risk of lung cancer by histologic type in middle-aged Japanese men and women: the JPHC study. *Int J Cancer* 99: 245–251
- Soda H, Tomita H, Kohno S, Oka M (1993) Limitation of annual screening chest radiography for the diagnosis of lung cancer. *Cancer* 72: 2341–2346
- Sone S, Li F, Yang ZG, Honda T, Maruyama Y, Takashima S, Hasegawa M, Kawakami S, Kubo K, Haniuda M, Yamada T (2001) Results of three-year mass screening programme for lung cancer using mobile low-dose spiral computed tomography scanner. *Br J Cancer* 84: 25–32
- Stellman SD, Takezaki T, Wang L, Chen Y, Citron ML, Djordjevic MV, Harlap S, Muscat JE, Neugut AI, Wynder EL, Ogawa H, Tajima K, Aoki K (2001) Smoking and lung cancer risk in American and Japanese men: An international case-control study. *Cancer Epidemiol Biomarkers Prev* 10: 1193–1199
- Swensen SJ, Jett JR, Sloan JA, Midthun DE, Hartman TE, Slykes AM, Aughenbaugh GL, Zink FE, Hillman SL, Noetzel GR, Marks RS, Clayton AC, Pairolero PC (2002) Screening for lung cancer with low-dose spiral computed tomography. *Am J Respir Crit Care Med* 165: 508–513
- Tsukada H, Yokoyama A, Kurita Y, Misawa H (2002) Evaluation of population based lung cancer screening in Niigata and analysis of interval cases based on comparison lung cancer registry with screening records. *Nihon Kokyuki Gakkai Zasshi* 38: 501–508 (in Japanese)
- Yoshimi I, Ohshima A, Ajiki W, Tsukuma H, Sobue T (2003) A comparison of trends in the incidence rate of lung cancer by histological type in the Osaka Cancer Registry, Japan and in the Surveillance, Epidemiology and End Results Program, USA. *Jpn J Clin Oncol* 33(2): 98–104
- Zappa M, Castiglione G, Paci E, Grazzini G, Rubeca T, Turco P, Crocetti E, Ciatto S (2001) Measuring interval cancers in population-based screening using different assays of fecal occult blood testing: the district of Florence experience. *Int J Cancer* 92: 151–154

特集

低線量 CT 検診の現状と展望

CT 肺がん検診の有効性の検討

中山 富雄

呼吸と循環

第56巻 第5号 別刷

2008年5月15日 発行

医学書院



# 特集

## 低線量 CT 検診の現状と展望

### CT 肺がん検診の有効性の検討\*

中山 富雄<sup>1</sup>

低線量 CT を用いた肺がん検診は、わが国で 1993 年より「東京から肺がんをなくす会」において世界で初めて行われた<sup>1)</sup>。従来の胸部 X 線検査では発見することができなかつた微小腺がんを高い頻度で発見できることが報告され、またそのほとんどが切除可能であり、予後も極めて良好であることが明らかとされた<sup>2-4)</sup>。この報告から、低線量 CT 肺がん検診は、肺がん対策の切り札として注目され、国内外で数多くの研究が行われ、また人間ドックを中心に国内で広く普及してきている。ここでは、疫学的な立場から低線量 CT 肺がん検診の有効性に関する評価について概説する。

#### がん検診の評価

がん検診手法の評価を行う場合、最も簡単な方法は発見率を比較するものである。従来の方法と比べて、新しい方法の発見率が著明に高い場合は、新しい方法が“有益”であると結論している報告が多い。発見率に加えて発見がんの病期割合が I 期にシフトしていたり、切除率が高かったり、生存率が高いという成績を知ると新しい方法が“有益”であると“確信”してしまいがちである。しかし、これらの情報は症例研究(case series)の範疇に入るものであり、がん検診の評価方法としては最も信頼性の低いものとして位置づ

表 1 がん検診の評価手法

研究方法	評価指標
実験的手法(前向き) ランダム化比較試験	
観察的手法(前向き) コホート研究	死亡率減少
(後向き) 症例対照研究	
症例研究	感度・特異度
	5年生存率
	病期早期化
	腫瘍径短縮

けられている(表 1)。

それはこのような症例研究においてはバイアスが混入しやすいためである。がん検診に関する症例研究で混入しやすいバイアスとしては、length bias, overdiagnosis bias などがある。CT 肺がん検診の場合、発見される小型肺がんの多くが極めて遅い進行速度を示すことが明らかになっており、腫瘍の進行速度に関係するこの 2 つのバイアスが問題となる。length bias は図 1 のように、定期的な検診を行った場合、進行速度の遅いがんのみが発見され、早いがんは検診で発見されにくいという偏りである。一方、overdiagnosis bias はそれを極端にしたもので、進行速度の極端に長いもののなかには、放置したとしてもその腫瘍によって命を損なうことのないものが存在し、それが検診発見がんのなかに混入するという偏りであ

\* An Evaluation of the Efficacy of CT Screening for Lung Cancer

<sup>1</sup> 大阪府立成人病センター調査部疫学課(〒537-8511 大阪市東成区中道1-3-3) Tomio Nakayama: Division of Epidemiology, Department of Cancer Control and Statistics, Osaka Medical Center for Cancer and Cardiovascular Diseases

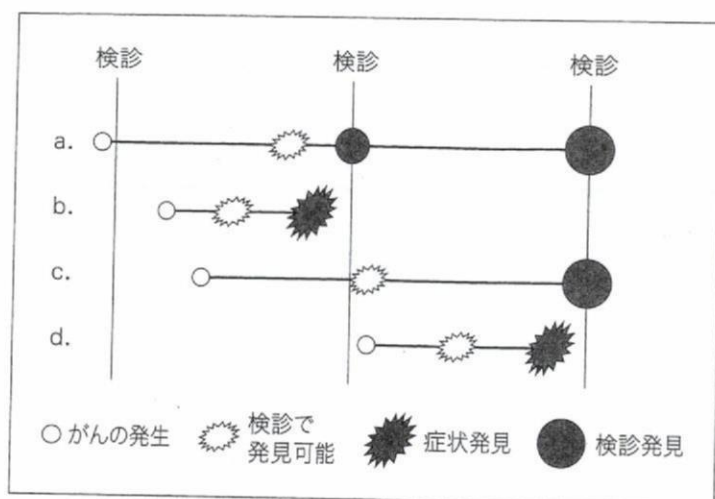


図1 length bias

定期的な検診を行った場合、進行速度の遅いもの(a, c)は、何回かの検診の受診機会のどこかで発見される可能性があるが、進行速度の速いもの(b, d)は、検診の受診機会が来る前に症状が出て検診外で発見されてしまう。このことにより検診発見例だけに注目すると進行速度の遅いもの(=予後の良いもの)に偏ってしまい、進行速度の速いもの(=予後の悪いもの)は見逃されてしまう。

る。これらの2つのバイアスは、実際にCT肺がん検診で見つかった肺野末梢病変をfollow-upする際に、しばしば経験される。Kodamaらは、2cm以下のすりガラス状陰影を呈する病変の約半数が2年以上の経過で増大せず、そのなかには切除後肺がんと確定した例も含まれると報告している<sup>5)</sup>。

このように、overdiagnosisと考えられる病変が検診発見例のなかにならかなり高い割合で混入していることが推定される場合は、発見率やI期率、生存率など発見例に限定した症例研究のみでは、検診の評価を過大に評価することにつながる。検診の場合は、検診と検診の間に発症する検診間発見がん(interval case)が少なからず存在する。したがって、発見例のみならず検診間発見がんも含めた評価(受診者集団全体の死亡率)が必須となる。

#### 科学的根拠を踏まえた肺がん検診の評価に関するガイドライン

現在わが国で、がん検診の有効性評価を行っているのは、厚生労働省がん研究助成金「がん検診

の適切な方法とその評価法の確立に関する研究」班(主任研究者：平成15～18年度 祖父江友孝、平成19年度～現在 濱島ちさと)である。この研究班ではUS Preventive Service Task Forceなどの諸外国で培われた有効性評価に関する手法を元に、既存の文献をEBMの手法を用いて評価し、各種がん検診に関するガイドラインを作成している<sup>6)</sup>。2006年に公開された「有効性評価に基づく肺がん検診ガイドライン」<sup>7)</sup>においては、検査法として、「非高危険群に対する胸部X線検査および高危険群に対する胸部X線検査と喀痰細胞診併用法」、および「低線量CT」の2つの手法が検討された。方法の詳細およびガイドラインについては、「科学的評価を踏まえたがん検診推進のページ」(<http://canscreen.ncc.go.jp>)から入手が可能である。

この2つの検診手法について、Medlineなどから1985年以降2005年までの1,038文献が抽出され、最終的に72文献が採用された。検診手法別に証拠のまとめが作成され、それを元に証拠のレベルと推奨がそれぞれ作成された。

「非高危険群に対する胸部X線検査および高危険群に対する胸部X線検査と喀痰細胞診併用法」に関しては、1970年代に行われた欧米でのランダム化比較試験の成績よりも、1990年代に行われた国内の症例対照研究の成績を重視し、証拠のレベルとしては2+ (死亡率減少効果の有無を示す、中等度の質の症例対照研究・コホート研究が行われている)と判定された。

一方「低線量CT」に関しては、死亡率減少効果を評価した直接的証拠はほとんどなく、Mayo Clinicで行われた4年間のCT検診受診者の肺がん死亡率を、1970年代に行われたMayo Lung Projectの肺がん死亡率と比較し、差を認めなかったというSwensenらの成績が唯一であった<sup>8)</sup>。間接的証拠としても追跡法を用いたCT検診の感度・特異度に関する報告はなく、stage shiftを示した論文もなく、証拠としてはあまりにも不十分な状況であった。このため証拠のレベルとしては、2- (死亡率減少効果に関する、質の低い症例対照研究・コホート研究が行われている)と判定された。



表2 各検診手法の証拠と推奨のまとめ

検査方法	証拠	推奨	表現
非高危険群に対する胸部X線検査, および高危険群に対する胸部X線検査と喀痰細胞診併用法	2+	B	死亡率減少効果を示す相応な証拠があるので, 対策型検診および任意型検診として, 非高危険群に対する胸部X線検査, および高危険群に対する胸部X線検査と喀痰細胞診併用法による肺がん検診を実施することを勧める。ただし, 死亡率減少効果を認めるのは, 二重読影, 比較読影などを含む標準的な方法を行った場合に限定される。標準的な方法が行われていない場合には, 死亡率減少効果の根拠はあるとはいえず, 肺がん検診としては勧められない。また, 事前に不利益に関する十分な説明が必要である。
低線量CT	2-	I	死亡率減少効果の有無を判断する証拠が不十分であるため, 対策型検診として実施することは勧められない。任意型検診として実施する場合には, 効果が不明であることと不利益について適切に説明する必要がある。なお, 臨床現場での撮影条件を用いた非低線量CTは, 被曝の面から健常者への検診として用いるべきではない。

表3 評価指標に肺がん死亡率を含めた研究

	JLSS	NLST	NELSON study
国名	日本	米国	オランダ-ベルギー
開始年	2001年 (1997~2002年まで登録)	2002年(2004年に登録終了)	2003年
対象者	40歳以上男女 CT検診受診者46,700人 通常検診受診者84,000人	55~74歳 重喫煙者 53,500人	50~75歳 男性喫煙者 28,000人
肺がん死亡率減少の想定		20% reduction	25% reduction (検出力80%)
方法	コホート研究 CT検診群 少なくとも一度受診 通常検診群 CT検診の受診歴なし	ランダム化比較試験 研究群 CT検診1,2,3年目 対照群 単純X線検診	ランダム化比較試験 研究群 CT検診1,2,4年目 対照群 無検診
追跡年数	平均7年	4.5年	10年
最終解析予定	2008年	2009年	記載なし

不利益として, 各検査法の偽陰性率・偽陽性率・放射線被曝・精密検査の偶発症などが比較されている。特に放射線被曝に関しては, 胸部単純X線による被曝は極めて小さいものの, 低線量CTの被曝は経年検診を長期間続けた場合は無視できないものであることが示されているが, 対象年齢と受診間隔の設定, 利益と不利益のバランスの評価が必要であるという見解が示されている。なお, 臨床条件での非低線量CTによるスクリーニング検査は健常者に対して行うべきではないと明確に示されている。

このガイドラインの結論である推奨のレベルは, 「非高危険群に対する胸部X線検査および高危険群に対する胸部X線検査と喀痰細胞診併用

法」はB(死亡率減少効果を示す相応な証拠があるので, 実施することを勧める), 「低線量CT」がI(死亡率減少効果の有無を判断する証拠が不十分であるため, 対策型検診として実施することは勧められない。任意型検診として実施する場合には, 効果が不明であることと不利益について適切に説明する必要がある)と判定されている(表2)。

#### ガイドラインに含まれなかった研究

前章で述べたガイドラインは, 2005年末までご報告されたものを評価したものである。以後2007年末までに死亡率に関して言及した論文は認められないものの, 進行中の研究がある。国内

では厚生労働省第3次対がん総合戦略研究事業に基づき筆者が主任研究者をしている Japan Lung Screening Study が進行中である。本研究は、CT 肺がん検診を一度でも受診したことのあるもの約4万人と単純X線肺がん検診を受診した8万人を登録し追跡したコホート研究である。本研究は登録後約6年弱の追跡結果を現在解析している最中であり、現時点で結果の公開には至らないが、少なくとも一度のみのCT肺がん検診の受診では大幅な死亡率減少効果は期待できないようである。

また国外においては、2つのランダム化比較試験が進行中である(表3)。National Lung Screening Trial (NLST) は、米国で進行中の53,500人を対象とした研究であり、研究群は年1回計3回の低線量CT検診を、対照群は単純X線検診をそれぞれ提供されるデザインである。ただし、この研究は追跡期間がわずか5年であり、国内で経験されるような高分化な小型腺がんの発見および救命を評価することはできないと危惧されている。一方、オランダ・ベルギーで行われている NELSON trial は28,000人規模の研究であり、対照群には検診を一切提供しないデザインである。こちらは追跡期間を10年と予定しており、NLST に比べると成果は期待できるデザインではある。

#### まとめ

低線量CT検診は、従来の単純X線検査では発見不可能であった肺野の小型腺がんを高頻度に発見しうること世界的に注目されたものの、その代名詞であった充実性部分を有さないすりガラス状陰影を呈する肺がんを切除する意義自体に疑問が投げかけられている。検診として実施する場

合には受診者集団の死亡率減少効果が証明される必要があるが、未だその結果は公表されていない。したがって、現時点では公的資金が用いられる地域住民を対象とした対策型検診での利用は正当化されない。人間ドックにおける任意型検診においては、効果が確認されていないことや過剰診断などの不利益が存在することを、受診者に説明し同意が得られたうえで実施されることが必要であろう。

#### 文献

- 1) Sobue T, Moriyama N, Kaneko M, et al: Screening for lung cancer with low-dose helical computed tomography: anti-lung cancer association project. *J Clin Oncol* 20: 911-920, 2002
- 2) Sone S, Li F, Yang G, et al: Results of three-year mass screening programme for lung cancer using mobile low-dose spiral computed tomography scanner. *Br J Cancer* 84: 25-32, 2001
- 3) Henschke CI, Naidich DP, Yankelevitz DF, et al: Early lung cancer action project: initial findings on repeat screenings. *Cancer* 92: 153-159, 2001
- 4) Nawa T, Nakagawa T, Kusano S, et al: Lung cancer screening using low-dose spiral CT: results of baseline and 1-year follow-up studies. *Chest* 122: 15-20, 2002
- 5) Kodama K, Higashiyama Y, Yokouchi H, et al: Natural history of pure ground-glass opacity after long-term follow-up of more than 2 years. *Ann Thorac Surg* 73: 386-393, 2002
- 6) 平成16年度厚生労働省がん研究助成金「がん検診の適切な方法とその評価法の確立に関する研究」班: 有効性評価に基づくがん検診ガイドライン作成手順書, 2004
- 7) 平成18年度厚生労働省がん研究助成金「がん検診の適切な方法とその評価法の確立に関する研究」班: 有効性評価に基づく肺がん検診ガイドライン, 2006
- 8) Swensen SJ, Jett JR, Hartman TE, et al: CT screening for lung cancer: five-year prospective experience. *Radiology* 235: 259-265, 2005



## 肺がん CT 検診の検証

中山 富雄\*  
なかやま とみお

- がん検診の有効性評価としては、発見率や生存率はバイアスが混入し過大評価となるため受診者集団全体の死亡率減少効果を示すことが必須である。
- 低線量 CT については、感度・特異度はもちろん死亡率減少効果を示した論文もまだ報告されていない。
- 現在死亡率減少効果を評価する研究として、国内ではコホート研究である Japan Lung Cancer Screening Study, 国外ではランダム化比較試験である National Lung Screening Trial と NELSON trial が進行中である。

**Key Words** 肺がん検診, CT 検診, 有効性評価, ランダム化比較試験, コホート研究

### はじめに

がん対策基本法制定時の附帯決議の第十七項において「最新の知見に基づき有効性が高いと認められるがん検診を地域における検診の項目に位置づけること」という記載がある。さて、「有効性が高いと認められるがん検診」とはいかなるものであろうか？ 肺がんではここ 10 年間低線量 CT を用いた肺がん検診（以下、CT 検診）の有効性が“期待されてきた”が<sup>1-3)</sup>、その評価の現状について報告する。

### □ がん検診の有効性評価

がんの診断に用いられる新しい診断機器が開発された場合に、従来の手法では見つけることができなかった小型の病巣の検出力（発見率・I 期

率・切除率など）をもって、「有用である」という評価を示す報告が多く見られる。確かに、小型の病巣の検出力があがることを根拠に、新しい検査法を臨床に導入することは正しい。しかし診療に導入する場合と、がん検診に導入する場合とでは、評価する指標や理念は大きく異なる。表 1 に診療と検診の相違点を示した。注意しなければならないことは、がん検診が健常者を対象としている点にある。健常者の有病率は、症状をもって病院に来院する受診者集団の有病率の数百分の 1 にすぎない。病院受診者集団では、有病率が高いために個人に換算した場合の検査法のメリットがデメリットを大幅に上回ることが多いため、小型の病巣の検出力をもって「有効」と結論してもそれほど問題にはならない。一方、検診の

表 1 診療と検診の相違点

	診 療	検 診
受診者	有症状者もしくは異常所見を指摘されている。有病率が高い。	無症状である。有病率は低い。
緊急性	症状があることから、病気の進行度が進んでいることが多く、緊急性が高い。	病気の進行度は低く、緊急性が低い。
診断の際の優先課題	緊急性が高いことから、正しく診断することが最優先	健康な人に病気であるという誤った判定をつける割合を減免すること。
検査法	侵襲性のある検査や、高額の検査もある程度までは許される。	侵襲性のない、安価な検査

注) 検診での最重要課題は、健常者に害を与えない（与える確率を最小化する）ということである。要精検率が数 .0% というような検査は、がん検診ではあり得ない。

\*大阪府立成人病センター 調査部疫学課



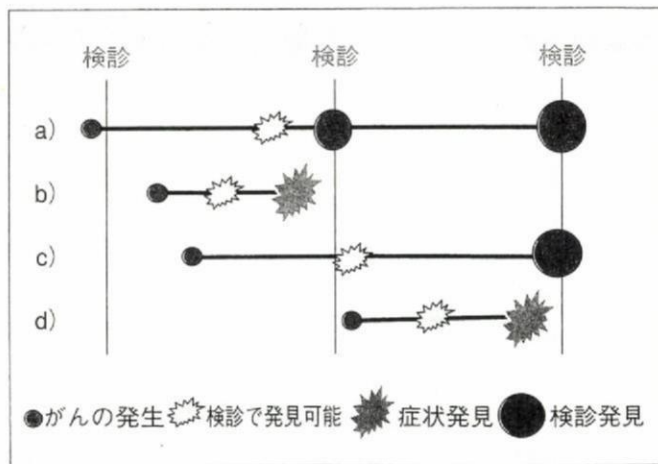


図1 Length bias

定期的な検診を行った場合、進行速度の遅いもの(a, c)は、何回かの受診機会のどこかで発見される可能性があるが、進行速度の速いもの(b, d)は、検診の受診機会が来る前に、症状が出て検診外で発見されてしまう。このことにより検診発見例だけに着目すると進行速度の遅いもの(=予後のよいもの)に偏ってしまい、進行速度の速いもの(=予後の悪いもの)は見逃されてしまう。

場合は、受診者個人に換算するとメリットは小さく、デメリットと近接してくるため、その是非については、より厳密な評価が要求される。

検診の場合、発見率や生存率などの発見がに限った指標は、バイアス(偏り)が混入しやすく、検診の効果を過大評価することにつながるといわれている。図1に示すように、定期的な検診により、進行速度の緩やかながんが発見されやすく、進行速度の速いがんは発見されにくいという傾向がある(レンジス・バイアス)。CT検診の場合は、このバイアスの影響を強く受けている。一方、進行速度の速いがんは、検診と検診との間に症状を伴って発見されることが多い。検診の評価とすれば、“発見できたがん”のみならず、“発見できなかった”がんの割合およびその予後を併せて評価することが不可欠である。このため受診者集団全体の当該がん死亡率こそが、がん検診の有効性を示すもっとも重要な指標と国際的に位置づけられている。

がん検診の死亡率減少効果を評価する研究デザインとしては、ランダム化比較試験、コホート研究、症例対照研究などがあり、もっとも偏りが混入しにくいといわれているのは、ランダム化比較

試験である。治療の評価と異なり、がん検診の評価の場合、がん死亡自体が稀なため、サンプルサイズは1万人規模、追跡期間も10年程度となり、国際的に見ても容易には実施し得ない状況にある。

## □ 肺がん CT 検診の現状

CT検診に関しては、2008年3月の時点で感度・特異度を報告した論文はない。あくまで単純X線で見えないものをCTで発見できたという意味でのものしかなく、がん登録との照合などの標準的な方法を用いた報告はない。また受診者集団の死亡率を報告した論文も、現状ではSwensenらの報告のみに留まる<sup>4)</sup>。この報告はMayo Clinicの4年間のCT検診受診者の肺がん死亡率を、70年代に行われたMayo Lung Projectの肺がん死亡率と比較し、差を認めなかったというものであり、30年前のhistorical controlとの比較という点と、追跡期間が登録後4年と短い点で十分な評価とはいえない。Henschkeらが設立した、I-ELCAPという国際的CT検診の共同研究がもっとも有名ではあるが<sup>5)</sup>、生存率の報告しかなされていらない。この研究はCT検診受診群のみで、対照群が設けられておらず、検診の評価研究としては、はなはだ不十分といわざるをえない。

CT検診受診者の死亡率を評価指標とする進行中の研究がいくつかある。国内では厚生労働省第3次対がん総合戦略研究事業に基づき筆者が主任研究者をしているJapan Lung Screening Studyが進行中である。本研究は、CT肺がん検診を一度でも受診したことのあるもの約4万人と単純X線による肺がん検診を受診しCT検診を受診していない8万人を登録し追跡したコホート研究である。本研究は登録後約6年弱の追跡結果を現在解析している最中であり、現時点で結果の公開には至らないが、少なくとも一度のみのCT肺がん検診の受診では大幅な死亡率減少効果は期待できないようである。

また国外においては、二つのランダム化比較試験が進行中である(表2)。National Lung Screening Trial (NLST)は、米国で進行中の53500人を対象とした研究であり、研究群は年1回計3回の



表2 国内外における低線量肺がん CT 検診の有効性評価を目的とした現在進行中の研究

	Japan Lung Screening Study (JLSS)	National Lung Screening Trial (NLST)	NELSON study
国名	日本	米国	オランダ・ベルギー
開始年	2001 (1997~2002年まで登録)	2002 (2004年に登録終了)	2003
対象者	40歳以上男女 CT検診受診者46700人 通常検診受診者84000人	55~74歳 重喫煙者 53500人	50~75歳 男性喫煙者 28000人
肺がん死亡率減少の想定		20% reduction	25% reduction (検出力80%)
方法	コホート研究 CT検診群 少なくとも一度受診 通常検診群 CT検診の受診歴なし	ランダム化比較試験 研究群 CT検診1, 2, 3年目 対照群 単純X線検診	ランダム化比較試験 研究群 CT検診1, 2, 4年目 対照群 無検診
追跡年数	平均7年	4.5年	10年
最終解析予定	2008年	2009年	記載なし

低線量 CT 検診を、対照群は単純 X 線検診をそれぞれ提供されるデザインである。ただしこの研究は追跡期間がわずか5年であり、国内で経験されるような高分化な小型腺がんの発見および救命を評価することはできないと危惧されている。一方、オランダ・ベルギーで行われている NELSON trial は 28000 人規模の研究であり、対照群には検診を一切提供しないデザインである。こちらは追跡期間を 10 年と予定しており、NLST に比べると成果は期待できるデザインではある。

### まとめ

低線量 CT 検診は、従来の単純 X 線検査では発見不可能であった肺野の小型腺がんを高頻度に発見し得ることで世界的に注目されたものの、その代名詞であった充実性部分を有さないすりガラス状陰影を呈する肺がんを切除する意義自体に疑問が投げかけられている<sup>6)</sup>。検診として実施する場合には受診者集団の死亡率減少効果が証明される必要があるが、いまだその結果は公表されていない。したがって現時点では公的資金が用いられる地域住民を対象とした対策型検診での利用は正当化されない。人間ドックにおける任意型検診においては、効果が確認されていないことや過剰

診断などの不利益が存在することを、受診者に説明し同意が得られたうえで、実施されることが必要であろう<sup>7)</sup>。

### 文献

- 1) Sobue T, Moriyama N, Kaneko M, et al : Screening for lung cancer with low-dose helical computed tomography : anti-lung cancer association project. J Clin Oncol 20 : 911-920, 2002
- 2) Sone S, F Li, G Yang, et al : Results of three-year mass screening programme for lung cancer using mobile low-dose spiral computed tomography scanner. British J Cancer 84 : 25-32, 2001
- 3) Nawa T, Nakagawa T, Kusano S, et al : Lung cancer screening using low-dose spiral CT : results of baseline and 1-year follow-up studies. Chest 122 : 15-20, 2002
- 4) Swensen SJ, Jett JR, Hartman TE, et al : CT screening for lung cancer : five-year prospective experience. Radiology 235 : 259-265, 2005
- 5) Henschke CI, Naidich DP, Yankelevitz DF, et al : Early lung cancer action project : initial findings on repeat screenings. Cancer 92 : 153-159, 2001
- 6) Kodama K, Higashiyama Y, Yokouchi H, et al : Natural history of pure ground-glass opacity after long-term follow-up of more than 2 years. Ann Thorac Surg 73 : 386-393, 2002
- 7) 平成 18 年度厚生労働省がん研究助成金「がん検診の適切な方法とその評価法の確立に関する研究」班：有効性評価に基づく肺がん検診ガイドライン。2006

日本臨牀 66 卷 増刊号 6 (2008 年 8 月 28 日発行) 別刷

# 肺 癌

—基礎・臨床研究のアップデート—

臨床研究 II. 検 診

胸部単純 X 線と喀痰細胞診を用いた肺癌検診の評価

中山富雄



## 胸部単純X線と喀痰細胞診を用いた肺癌検診の評価

Evaluation of the effectiveness lung cancer screening used by plain chest X-ray and sputum cytology

中山富雄

**Key words** : 肺癌検診, ランダム化比較試験, 症例対照研究, 過剰診断

## はじめに

肺癌の早期発見法としては, 肺野末梢に発生する肺癌を標的とした画像診断と, 中枢気管支に発生する肺癌(主に扁平上皮癌)を標的とした喀痰細胞診が, 広く用いられてきた. 近年画像診断法として低線量CTを肺癌検診に導入することが期待されているが, 従来胸部単純X線も, いまだ広く用いられている. この胸部単純X線と喀痰細胞診を用いた肺癌検診の有効性については, 古くから世界的に議論が行われてきたが, いまだ解決をみていない.

本稿では, まずがん検診に関する評価方法を概説し, 肺癌検診の評価に関する国内外の研究と, 現状での評価について解説する.

## 1. がん検診の評価指標

がん検診の評価指標としては, 発見率や発見がんの切除率・I期率・腫瘍径の小型化などがよく用いられている. しかし, これらの発見がんに限った指標は, バイアス(偏り)が混入しやすく, 検診の効果を過大評価することにつながる. 図1に示すように, 定期的な検診により, 進行速度の緩やかながんが発見されやすく, 進行速度の速いがんは発見されにくいという傾向がある(レンジス・バイアス). したがって進行

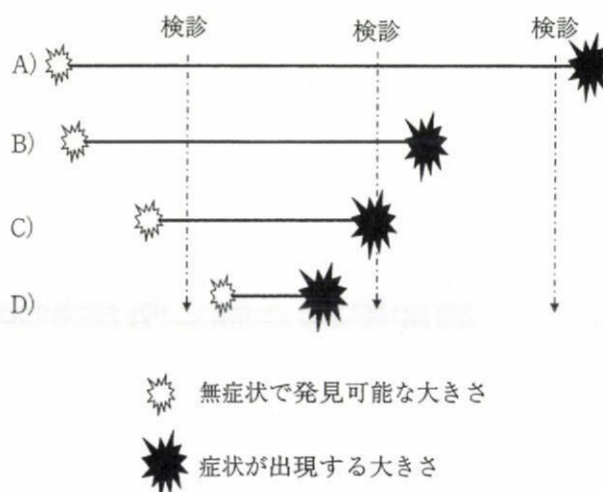


図1 レンダス・バイアス

定期的検診が行われた場合, A)では3回の検診の機会いずれでも早期で発見可能, B)では2回の検診の機会でも早期発見可能, C)では早期発見可能なのは1回のみ, D)では定期検診による発見自体困難で中間期がんとして診断される. このように進行速度の遅いがん(=予後の良いがん)が検診では発見されやすい.

速度の速いがんは, 検診と検診との間に症状を伴って検診とは別に医療機関を受診し発見される(中間期がん)ことが多い. 中間期がんの把握は容易ではないが, これらの割合およびその予後を発見がんの指標と併せて評価しないと, がん検診の正しい評価にはならない. したがって,

表1 70年代に行われた胸部X線と喀痰細胞診を用いた肺癌検診の有効性評価のためのランダム化比較試験

	文献	開始年	検診方法		参加者数	RR (95%CI)
			検診群	対照群		
Johns-Hopkins Lung Project	1	1973	胸部X線年1回+ 喀痰4カ月ごと	胸部X線 年1回	5,250: 5,171	0.91 (0.72-1.16)
Memorial Sloan-Kettering Lung Project	2	1,974	胸部X線年1回+ 喀痰4カ月ごと	胸部X線 年1回	4,968: 5,072	0.92 (0.67-1.26)
Mayo Lung Project	4	1971	X線と喀痰 4カ月ごと	X線と喀痰 年1回を推奨	4,618: 4,593	1.06 (0.82-1.36)
Czechoslovakian Lung Project	5	1976	X線と喀痰 6カ月ごと	無検診	3,171: 3,174	1.36 (0.94-1.97)

発見がんという限られた集団ではなく、検診受診者全体の予後の指標であるがん死亡率が、がん検診の有効性を示す最も重要な指標となる。

がん検診の死亡率減少効果を評価する研究デザインとしては、ランダム化比較試験、コホート研究、症例対照研究などがあり、最も偏りが混入しにくいといわれているのは、ランダム化比較試験である。近年がん患者に対する抗がん剤の臨床試験などでは、ランダム化比較試験が広く普及したものの、健常者に対するがん検診の評価の場合は、がん死亡というイベント自体がまれなため、サンプルサイズは1万人規模、追跡期間も10年程度となるため、容易には実施しえない。肺癌については1970年代に米国を中心に4つのランダム化比較試験が行われ、90年代以降我が国で6つの症例対照研究が行われた。

## 2. 70年代に行われた肺癌検診を評価するランダム化比較試験

1970年代には、胸部単純X線撮影は、肺癌の早期発見に欠かせないものとして、世界的に認識されていた。一方、喀痰細胞診を用いることで、単純X線で指摘不可能な肺門部早期扁平上皮癌を発見可能であること、その診断に当時我が国で開発された気管支ファイバースコープが必要であることが、世界的トピックスであった。このため、主に喀痰細胞診の評価を行うことを

目的として、4つのランダム化比較試験が米国を中心に行われた(表1)。

Johns-Hopkins および Sloan-Kettering で行われた研究は、いずれも胸部X線写真に喀痰細胞診を上乗せする効果を評価した研究である<sup>1,2)</sup>。この2つの研究は、いずれも喀痰細胞診の上乗せ効果を証明することはできなかったが、最近の佐藤らの報告によれば、日本の検診の成績と比べて、精度が不良であったことが報告されている<sup>3)</sup>。

Mayoの研究とチェコの研究は、胸部単純X線撮影と喀痰細胞診の併用法を評価した研究である<sup>4,5)</sup>。チェコの研究は、腺癌(大細胞癌を含めて)の割合が全体の18%しかなく、また術後30日以内の死亡率が11%と際立って高いという当時の東ヨーロッパの貧困な医療状況を示す結果であり、現在の日本の医療と比較することは無理といわざるを得ない。したがって、肺癌検診の評価として参考にすべきは、規模が最も大きいMayo Clinicの研究が中心となる。

Mayo Lung Projectは、そのユニークなデザインと、得られた結果の矛盾点について、いまだに論争の絶えない研究である。図2に示すように、この研究は45歳以上の重喫煙男性10,933人をリクルートし、まず胸部単純X線と喀痰細胞診でスクリーニングし、発見された肺癌例(prevalence case)および全身状態不良例1,722人を研究から除外した。これは初回検診の受診者の中に



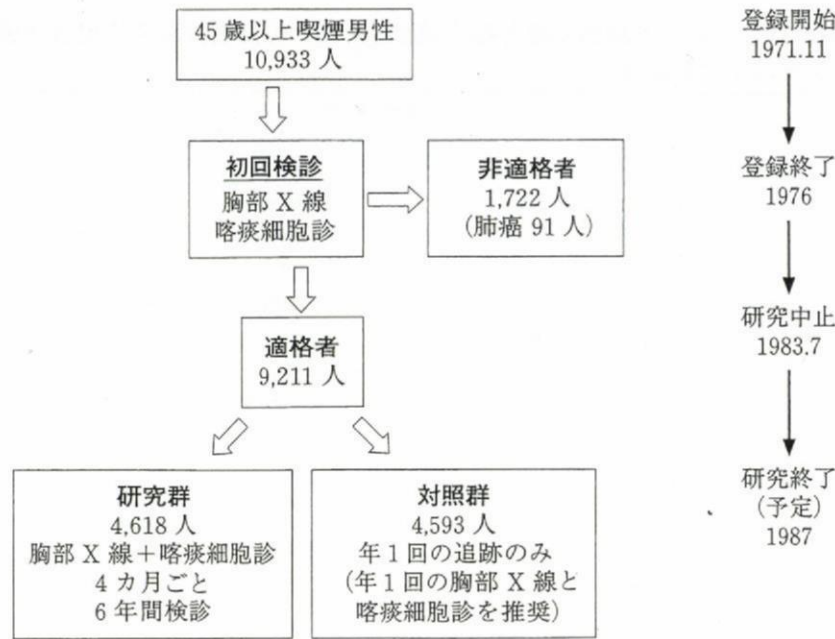


図2 Mayo Lung Projectの構造

研究群には4カ月ごと6年間検診が提供され、対照群には年1回の検診受診が推奨されるだけにとどまった。しかし研究群のコンプライアンスは75%と低く、対照群のコンタミネーションは73%(最後の2年間のうち1回以上検査を受診した割合)と高かった。

は、進行がんの占める割合が高いことから、これを除外して、その後の検診の効果を評価することに専念したためである。適格者9,211人を2群に割り付け、研究群は4カ月に1度の検診が提供されたが、対照群には、年に1度検診を受けるようにアドバイスがなされた。これは当時のMayo Clinicの外来に、「喫煙者は年に1回検診を受診するように！」というポスターが掲示されていたように、検診を年1回受診することの妥当性には問題ないと考えられていた事情を反映している。1971年から開始された研究は、研究群と対照群の間で肺癌死亡率に差がないという83年の中間報告をきっかけに、研究は中止に追い込まれた。

当時は疑いもされていなかった肺癌検診の効果を否定する結果は、大きな衝撃を与えたが、更にランダム化比較試験でありながら両群中の肺癌罹患数に明らかな差があったことが、その後の議論の焦点となった(表2)。すなわち、対照群で把握された肺癌(当然すべて検診外発見)は160例であったが、研究群で把握された肺癌(検診発見に検診外発見を加えたもの)は206例

表2 Mayo Lung Projectの肺癌罹患の状況(文献<sup>6)</sup>より引用)

	研究群	対照群
検診発見	90	
胸部X線	66	
喀痰細胞診	18	
双方発見	6	
検診外発見	116	160
胸部X線*	43	48
症状	73	112
計	206	160

\*検診以外で撮影された胸部X線を示す。研究群においても年3回の研究の枠内の検査以外に、更に胸部X線を受診しているものが少なかったことを示している。

であり、46例も研究群の罹患数が多かったのである。Eddyは、この差を過剰診断(本来放置しても死に至らないがんを検診で発見していた)と解釈し、本研究の結果の解釈の中心となっている<sup>6)</sup>。一方Straussらは、追跡が不十分であった、あるいは割り付けが不適切であったと

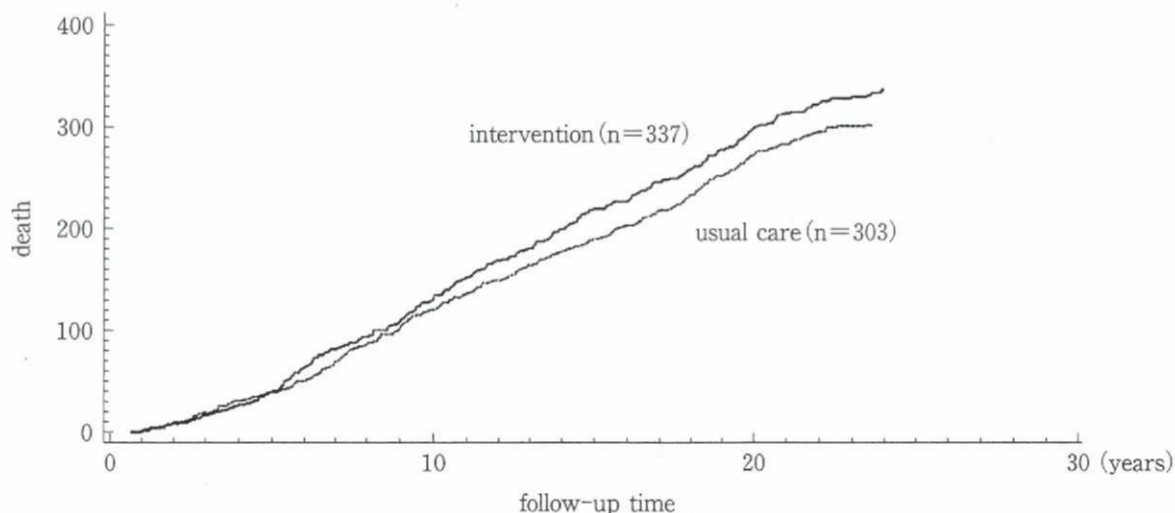


図3 Mayo Lung Projectの長期追跡研究における累積肺癌死亡数の推移  
intervention: 研究群, usual care: 対照群.  
(文献<sup>8)</sup>より引用)

表3 Mayo Lung Projectの長期追跡研究での組織型別死亡数・生存率(文献<sup>8)</sup>より引用)

	扁平上皮癌		腺癌		大細胞癌		小細胞癌	
	研究群	対照群	研究群	対照群	研究群	対照群	研究群	対照群
発見肺癌数	68	51	59	38	29	24	48	45
肺癌死亡	35	33	39	25	19	20	40	40
全死因死亡	61	49	55	37	37	26	47	44
5年生存率								
(肺癌死)	55	26	30	33	35	9	6	6
(全死亡)	40	14	21	21	29	6	4	2

いう解釈を提案している<sup>7)</sup>.

最近 Marcus らは、研究対象者の20年にわたる長期追跡結果を明らかにしたが、研究開始後20年を経過しても肺癌死亡率に差はなかったと報告しており、追跡不十分説や割り付け不均等説を否定し、過剰診断のみで説明可能としている<sup>8)</sup>。しかし Marcus らの論文においても、2つの問題がある。まずこの論文で示されている肺癌累積死亡率の図(図3)をみると、統計学的有意差はないものの、研究群の方の肺癌死亡率が高いことが示されている。また組織型別死亡率・生存率の表(表3)によると、腺癌では研究群も対照群も5年生存率に差はなく、扁平上皮癌で研究群の方が有意に対照群よりも高い成績が示されている。Marcus らの説のとおり、46例の差が過剰診断によるとすれば、その大半は

進行速度の遅い腺癌によるものと考えられ、腺癌の生存率が研究群で上昇しているはずであるが、そうではなかった。もし扁平上皮癌で過剰診断が起こるとすれば、喀痰細胞診発見による肺門部扁平上皮癌のはずであるが、Mayoの扁平上皮癌の7割が胸部X線発見であったと報告されている<sup>4)</sup>。

このように、いまだに疑問が残った結果であり、ランダム化比較試験とはいえ、偏りを制御できなかった可能性が残る。

### 3. 国内で行われた肺癌検診を評価する症例対照研究

国内でも1980年代に肺癌検診に関するランダム化比較試験を行うべきであるという議論が行われたようであるが、当時は結核予防法によ