KJIM

Risk factors for primary lung cancer among never-smoking women in South Korea: a retrospective nationwide population-based cohort study

Yoon Ho Ko[1,2,*], Seung Joon Kim[3,*], Wan-Seop Kim[4], Chan Kwon Park[5], Cheol-Kyu Park[6], Yang-Gun Suh[7], Jung Seop Eom[8], Sukki Cho[9], Jae Young Hur[4], Sung Ho Hwang[10], and Jun-Pyo Myong[11]

[1]Division of Oncology, Department of Internal Medicine and [2]Cancer Research Institute, College of Medicine, The Catholic University of Korea, Seoul; [3]Division of Pulmonology, Department of Internal Medicine, Seoul St. Mary's Hospital, College of Medicine, The Catholic University of Korea, Seoul; [4]Department of Pathology, Konkuk University Medical Center, Seoul; [5]Division of Pulmonology, Allergy and Critical Care Medicine, Department of Internal Medicine, Yeouido St. Mary's Hospital, College of Medicine, The Catholic University of Korea, Seoul; [6]Department of Internal Medicine, Chonnam National University Hwasun Hospital, Hwasun; [7]Department of Radiation Oncology, National Cancer Center, Goyang; [8]Department of Internal Medicine, Pusan National University Hospital, Busan; [9]Department of Thoracic and Cardiovascular Surgery, Seoul National University Bundang Hospital, Seongnam; [10]Department of Radiology, Korea University Anam Hospital, Seoul; [11]Department of Occupational and Environmental Medicine, Seoul St. Mary's Hospital, of Medicine, The Catholic University of Korea, Seoul, Korea

Background/Aims: We performed a large-scale, retrospective, nationwide, cohort study to investigate the risk factors for lung cancer among never-smoking Korean females.

Methods: The study data were collected from a general health examination and questionnaire survey of eligible populations conducted between January 1, 2003 and December 31, 2004; the data were acquired from the tailored big data distribution service of the National Health Insurance Service. After a 1-year clearance period, 5,860,922 of 6,318,878 never-smoking female participants with no previous history of lung cancer were investigated. After a median follow-up of 11.4 years, 43,473 (0.74%) participants were defined as "newly diagnosed lung cancer".

Results: After adjusting for all variables at baseline, the variables older age, lower body mass index (BMI), less exercise, frequent alcohol drinking, meat-based diet, rural residence, and previous history of cancer were associated with a higher incidence of lung cancer. Low BMI (< 18.5 kg/m²: hazard ratio [HR], 1.33; 95% confidence interval [CI], 1.27 to 1.40) was a significant independent risk factor; as BMI decreased, HR increased. Negative associations between BMI and lung-cancer development were also observed after controlling for age (*p* for trend < 0.001). Drinking alcohol one to two times a week (HR, 1.25; 95% CI, 1.21 to 1.28) and eating a meat-based diet (HR, 1.08; 95% CI, 1.01 to 1.15) were associated with lung-cancer incidence.

Conclusions: Modifiable baseline characteristics, such as BMI, exercise, alcohol consumption, and diet, are risk factors for lung-cancer development among never-smoking females. Thus, lifestyle modifications may help prevent lung cancer.

Keywords: Lung neoplasm; Never-smoking women; Risk factor; Cohort studies

Correspondence to Jun-Pyo Myong, M.D.
Department of Occupational and Environmental Medicine, Seoul St. Mary's Hospital, College of Medicine, The Catholic University of Korea, 222 Banpo-daero, Seocho-gu, Seoul 06591, Korea
Tel: +82-2-2258-6267, Fax: +82-2-2258-6691, E-mail: medical001@catholic.ac.kr
https://orcid.org/0000-0001-8674-1034

*These authors contributed equally to this work.

Ko YH, et al. Lung cancer and never-smoking women

KJIM

INTRODUCTION

Lung cancer is one of the most common cancers worldwide, in terms of both incidence and cancer-related mortality. Although the age-standardized incidence rate in many developed countries has decreased in males, it increased in females between 1990 and 2016 [1]. In Korea, the population-standardized rate of death due to lung cancer was 34.1 per 100,000 persons (crude: n = 17,399) in 2015 [2]. Approximately 30% of all lung-cancer cases occurred in women [2], and 79.7% of females with lung cancer did not have any history of smoking [3,4]. The high incidence of lung cancer in never-smoking women in South Korea requires epidemiological investigation to identify risk factors for the development of lung cancer in never-smoking females.

The smoking rate decreased from 66.3% in 1998 to 42.1% in 2013 in men but did not change significantly in women, with the rate remaining at 5% to 8% [5]. However, the causes of lung cancer in never-smoking females are not well understood. Possible risk factors have been suggested, including occupational exposure, secondhand tobacco smoke, radon, indoor or outdoor air pollution, history of lung disease, family history of lung cancer, use of menopausal hormonal therapy, human papillomavirus infection, and dietary factors, such as a high intake of red or processed meat and a low intake of fruits and vegetables [6]. However, a large-scale study of the potential risk factors for lung-cancer development in never-smoking Korean women has not been conducted. Thus, this nationwide cohort study investigated the risk factors for lung cancer among never-smoking Korean females.

METHODS

Study design

This was a retrospective cohort study to evaluate potential risk factors for lung-cancer development among never-smoking Korean females. It was based on data derived from a general health examination (GHE) and questionnaire survey of an eligible population conducted on the GHE participation date. National Health Insurance (NHI) provides a database of never-smoking Korean females for the period January 1, 2002 through December 31, 2016. The clearance period of lung cancer was set to 1 year before and after participation in GHE (overall, a 2-year clearance period). The results of medical claims of participants during the period January 1, 2016, to December 31, 2016, were discarded. The follow-up duration was calculated from the date of GHE participation to the date of lung-cancer diagnosis (for the lung-cancer group), date of study completion (December 31, 2015), or date of death. The detailed scheme of the study design is shown in Fig. 1. This study was approved by the Institutional Review Board of the Catholic Medical Center (No. KC17ZESE0457), and was performed in accordance with the guidelines of human research. The requirement for written informed consent was waived by the Institutional Review Board (Catholic Medical Center) because of the retrospective nature of the study.
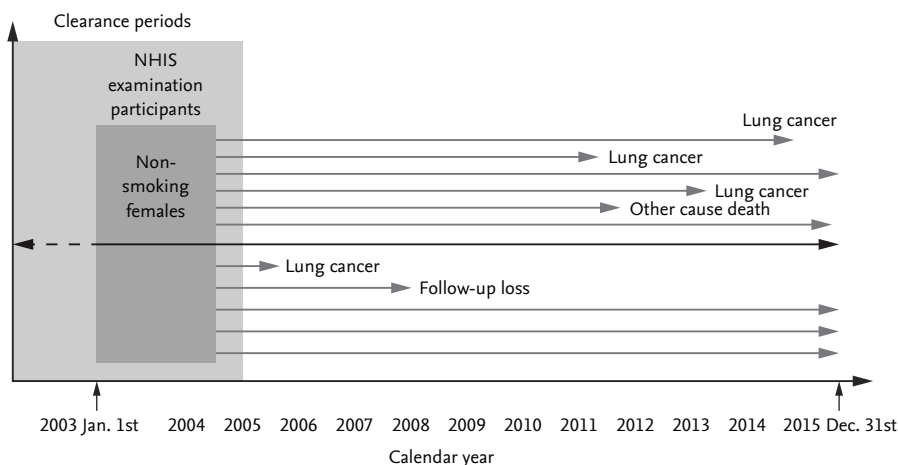
Figure 1. Scheme of the present study. NHIS, National Health Insurance Service.

Study population

The data were acquired from the tailored Big Data distribution service of the National Health Insurance Service (NHIS), which targets all employees or members of the community in the Republic of Korea. Data from individuals who participated in a GHE between January 1, 2003 and December 31, 2004 were collected, and subjects classified as current smokers during the same period were removed from the analysis. In total, 6,318,878 female participants with no previous history of lung cancer were identified from the NHIS data. Of these, 312,046 were excluded because they were not deemed to be non-smokers, as validated using a previously rated questionnaire in NHIS data for ex-smokers or current smokers. Of the remaining 6,006,832 participants, 132,164 with missing data for independent variables were excluded. In addition, those diagnosed with lung cancer in clearance periods (prevalent cases) were excluded (n = 13,747). Finally, 5,874,668 eligible female participants were selected for the final analysis. A flow chart depicting the selection process is shown in Fig. 2.
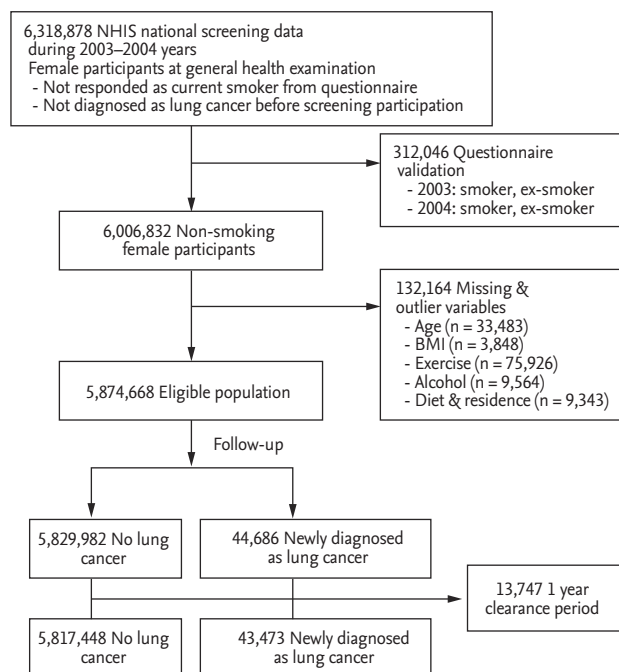
Definition of the study outcome

"Newly diagnosed lung cancer" is defined as C34 in the International Classification of Diseases 10th Revision (ICD-10), and the disease codes were derived from the Disease Code dataset (T40) of the NHIS from at least one or more in-patient or out-patient clinics (n = 100,242) (C34 criteria 1 in Fig. 2) during follow-up. To provide more reliability for the diagnosis, "newly diagnosed lung cancer" was established if C34 was recorded at least twice in the in-patient and/or out-patient databases (n = 44,686) (C34 criteria 2 in Fig. 2) during the follow-up period. The number of patients diagnosed with lung cancer during the clearance period was 13,747. Finally, 5,817,448 participants were defined as "no incidental lung cancer" and 43,473 as "newly diagnosed lung cancer." The date of death was extracted from the qualification NHIS dataset.

Baseline characteristics

Patient age was classified as < 40, 40 to 49, 50 to 59, 60 to 69, or ≥ 70 years and body mass index (BMI) as < 18.5, 18.5 to 22.9, 23 to 24.9, 25 to 29.9, and ≥ 30 kg/m². Health behaviors were obtained from the NHIS questionnaire dataset. Exercise and alcohol consumption were classified as binary variables and assessed by the following questions: exercise, "How many times do you exercise for 30 minutes or more a week?" (< 3 to 4 times vs. ≥ 3 to 4 times); alcohol consumption: "How many times do you drink alcohol a week?" (< 1 to 2 times vs. ≥ 1 to 2 times). Dietary habits were assessed by the question "What is your dietary style?" (vegetables vs. vegetables and meat vs. meat). Residence was categorized into rural and urban areas according to the registry database of the NHIS. Previous history of cancer was extracted from previous disease codes (T40) recorded in the NHIS registry (ICD-10 codes C54 and C55 for uterine cancer, C56 and C57 for ovarian and adnexa cancer, C53 for cervical cancer, C50 for breast cancer, and other C_{XX} for other cancers).

Figure 2. Study design and subject characteristics. NHIS, National Health Insurance Service; BMI, body mass index.

Statistical analysis

Baseline characteristics are presented as means (± standard deviation [SD]) and medians (ranges) for continuous variables and frequencies (%) for categorical variables. Student's *t* test was used to compare age and BMI, and analysis of variance was used to compare the follow-up periods. Chi-square tests were performed

Ko YH, et al. Lung cancer and never-smoking women

KJIM

for comparisons of categorical variables. Risk factors for newly diagnosed lung cancer were assessed using a Cox proportional hazards model. Cox regression analysis was performed after adjustment for all independent variables (age, BMI, alcohol consumption, diet habits, exercise, residence, previous cancer history). Age was stratified into five groups, and additional multiple Cox regression analyses of BMI and alcohol consumption according to age were performed. A $p < 0.05$ was considered statistically significant, and hazard ratios (HRs) are shown with their respective 95% confidence intervals (CIs). The statistical package SAS for Windows version 9.4 (SAS Inc., Cary, NC, USA) was used to perform all statistical analyses.

RESULTS

General characteristics of the study population
A review of the medical claims database revealed that 43,473 (0.74%) participants among the 5,860,922 never-smoking Korean females were newly diagnosed with lung cancer during the study follow-up period. The median follow-up period was 11.4 years (range, 1 to 12.8) for the entire study population, 6.8 years (1.0 to 12.6) for the lung-cancer group, and 11.5 years (1.0 to 13.0) for the no-lung-cancer group. Overall, 337,180 were censored from never-smoking Korean females due to death unrelated to lung cancer. Regarding the study population demographics, the mean age was 46.8 years and the mean BMI 23.4 kg/m². The elderly group exhibited a higher incidence of lung cancer than that of the younger group ($p < 0.001$). Subjects who exercised less than three to four or more times per week, consumed alcohol one to two or more times per week, or consumed a diet low in fruits and vegetables were more likely to develop lung cancer ($p < 0.001$). Living in a rural area and a history of other cancer were also associated with lung cancer (Table 1).

Factors associated with lung-cancer development
Table 2 shows the results of Cox proportional hazards analyses of the risk factors for newly diagnosed lung cancer. Age was found to be the most important risk factor for developing lung cancer. In addition, less exercise, more frequent alcohol drinking, living in a rural area, and a previous history of cancer were identified as risk factors in univariate and multivariate Cox regression analyses after adjusting for all independent variables.

After adjusting for all variables at baseline, lower BMI (< 18.5 kg/m²; HR, 1.33; 95% CI, 1.27 to 1.40) was a significant independent risk factor for lung cancer. The HR increased as the BMI decreased: the HRs were 1.33 (95% CI, 1.27 to 1.40) in the BMI < 18.5 kg/m² group, 0.88 (95% CI, 0.86 to 0.90) in the BMI range 23.0 to 24.9 kg/m² group, 0.80 (95% CI, 0.78 to 0.82) in the BMI range 25 to 29.9 kg/m² group, and 0.75 (95% CI, 0.70 to 0.79) in the BMI ≥ 30 kg/m² group. Similar negative associations between BMI and lung-cancer development were also observed after controlling for the effect of age (Fig. 3); statistical significance was strongest in subjects over 40 years of age (p for trend < 0.001).

As shown in Fig. 4, drinking alcohol one to two times per week was a risk factor for an increased risk of lung cancer among those aged 50 or more after age stratification. The HRs for lung cancer among subjects aged 50 to 59, 60 to 69, and ≥ 70 years were 1.13 (95% CI, 1.07 to 1.19), 1.33 (95% CI, 1.28 to 1.39), and 1.49 (95% CI, 1.41 to 1.57), respectively.

Consuming a meat-based diet (HR, 1.08; 95% CI, 1.01 to 1.15) was a significant independent risk factor for the development of lung cancer (Table 2). In a Cox regression analysis to validate the confounding effect of age, the risk of developing lung cancer in subjects consuming a meat-based diet was significantly higher in those aged ≥ 70 years (HR, 1.28; 95% CI, 1.12 to 1.46) than in those aged < 40 years (HR, 0.91; 95% CI, 0.72 to 1.15) (Table 3).

DISCUSSION

The aim of the present study was to investigate the risk factors for lung cancer among never-smoking Korean females. This population-based study showed that never-smoking Korean females with an older age, lower BMI, less exercise, more frequent consumption of alcohol, a meat-based diet, rural residence, or previous cancer history had a significantly increased risk of incidental lung cancer. In addition, the negative association with BMI and positive association with alcohol consumption remained significant even after controlling for the effect of age.

Table 1. General characteristics and behaviors of the study population according to newly diagnosed lung-cancer status

Characteristic	Lung cancer			p value
	No	Yes	Total	
Follow-up period, yr	11.5 (1.0–12.8)	6.8 (1.0–12.6)	11.4 (1.0–12.8)	< 0.001
Age, yr	46.7 ± 14.5	61.1 ± 11.4	46.8 ± 14.5	< 0.001
< 40	1,857,398 (99.9)	1,583 (0.1)	1,858,981 (31.7)	< 0.001
40–49	1,582,663 (99.6)	5,884 (0.4)	1,588,546 (27.1)	
50–59	1,135,001 (99.1)	9,789 (0.9)	1,144790 (19.5)	
60–69	863,645 (98.2)	15,573 (1.8)	879,218 (15.0)	
≥ 70	378,742 (97.3)	10,644 (2.7)	389,386 (6.7)	
Body mass index, kg/m²	23.4 ± 3.2	23.6 ± 3.1	23.4 ± 3.2	< 0.001
< 18.5	273,476 (99.4)	1,738 (0.6)	275,214 (4.7)	
18.5–22.9	2,446,185 (99.3)	16,586 (0.7)	2,462,771 (42.0)	
23–24.9	1,383,413 (99.2)	11,230 (0.8)	1,394,643 (23.8)	
25–29.9	1,554,019 (99.2)	11,230 (0.8)	1,566,748 (26.7)	
≥ 30	160,356 (99.3)	1,189 (0.7)	161,545 (2.8)	
Exercise, times/wk				< 0.001
< 3–4	4,764,048 (99.3)	34,785 (0.7)	4,798,832 (81.9)	
≥ 3–4	1,053,401 (99.2)	8,688 (0.8)	1,062,089 (18.1)	
Alcohol, times/wk				0.3093
< 1–2	4,891,410 (99.2)	36,475 (0.7)	4,927,885 (84.1)	
≥ 1–2	9,260,39 (99.2)	6,998 (0.8)	933,036 (15.9)	
Diet				< 0.001
Vegetables	1,391,877 (99.2)	11,428 (0.8)	1,403,305 (23.9)	
Vegetables + meat	4,199,277 (99.3)	31,008 (0.7)	4,230,284 (72.2)	
Meat	226,295 (99.5)	1,037 (0.5)	230,284 (3.9)	
Residence				< 0.001
Rural	1,211,421 (98.9)	12,782 (1.1)	1,224,203 (20.9)	
Urban	4,606,028 (99.3)	30,691 (0.7)	4,636,718 (79.1)	
Previous cancer history				< 0.001
None	5,732,229 (99.3)	41,572 (0.7)	5,773,800 (98.5)	
Uterine, ovarian, cervical cancer	7,508 (98.3)	126 (1.7)	7,634 (0.1)	
Breast cancer	9,063 (98.2)	160 (1.8)	9,223 (0.2)	
Other cancer	68,649 (97.7)	1,615 (2.3)	70,264 (1.2)	
Total	5,817,449 (100.0)	43,473 (100.0)	5,860,922 (100.0)	

Values are presented as median (range), mean ± SD, or number (%).

Several epidemiologic studies have supported an inverse relationship between BMI and lung- cancer development [7-11]. A recent meta-analysis of 29 studies investigating BMI and lung-cancer risk showed that such an inverse relationship was more prominent among current and ex-smokers; however, the effects of BMI were attenuated when the analysis was restricted to non-smokers [7]. The present study showed a sta-

Ko YH, et al. Lung cancer and never-smoking women

KJIM

Table 2. Cox proportional hazards ratios for risk factors for newly diagnosed lung cancer

Characteristic	Univariate Cox regression		Multivariate Cox regression
	Crude	Age-adjusted	
Age, yr			
< 40	1		1
40–49	4.32 (4.09–4.57)		4.61 (4.35–4.87)
50–59	10.04 (9.23–10.59)		10.95 (10.38–11.56)
60–69	21.31 (20.24–22.45)		23.06 (21.88–24.30)
≥ 70	38.09 (36.13–40.16)		39.87 (37.79–42.06)
BMI, kg/m^2			
< 18.5	0.95 (0.91–1.00)	1.34 (1.27–1.41)	1.33 (1.27–1.40)
18.5–22.9	1	1	1
23–24.9	1.19 (1.16–1.22)	0.88 (0.86–0.90)	0.88 (0.86–0.90)
25–29.9	1.20 (1.18–1.23)	0.80 (0.78–0.82)	0.80 (0.78–0.82)
≥ 30	1.09 (1.03–1.15)	0.74 (0.70–0.78)	0.75 (0.70–0.79)
Exercise, times/wk			
< 3–4	0.89 (0.87–0.91)	1.02 (1.00–1.04)	1.02 (1.00–1.05)
≥ 3–4	1	1	1
Alcohol, times/wk			
< 1–2	1	1	1
≥ 1–2	1.02 (0.99–1.04)	1.24 (1.20–1.27)	1.25 (1.21–1.28)
Diet			
Vegetables	1	1	1
Vegetables + meat	0.90 (0.88–0.92)	1.04 (1.02–1.07)	1.04 (1.02–1.06)
Meat	0.56 (0.52–0.59)	1.09 (1.02–1.16)	1.08 (1.01–1.15)
Residence			
Rural	1	1	1
Urban	0.63 (0.61–0.64)	0.93 (0.91–0.95)	0.94 (0.92–0.96)
Previous cancer history			
None	1	1	1
Uterine, ovarian, cervical cancer	2.37 (1.99–2.82)	1.91 (1.60–2.27)	1.98 (1.66–2.36)
Breast cancer	2.50 (2.14–2.92)	2.16 (1.85–2.52)	2.25 (1.92–2.62)
Other cancer	3.61 (3.43–3.79)	2.06 (1.96–2.17)	2.03 (1.93–2.13)

Values are presented as hazard ratio (95% confidence interval). Multivariate Cox regression: age, BMI, exercise, alcohol intake, diet, residence, and previous cancer history were adjusted.
BMI, body mass index.

tistically significant inverse relationship between BMI and lung-cancer risk. This discordance compared with previous studies could be caused by several differences among the analyses, such as study design, differences in
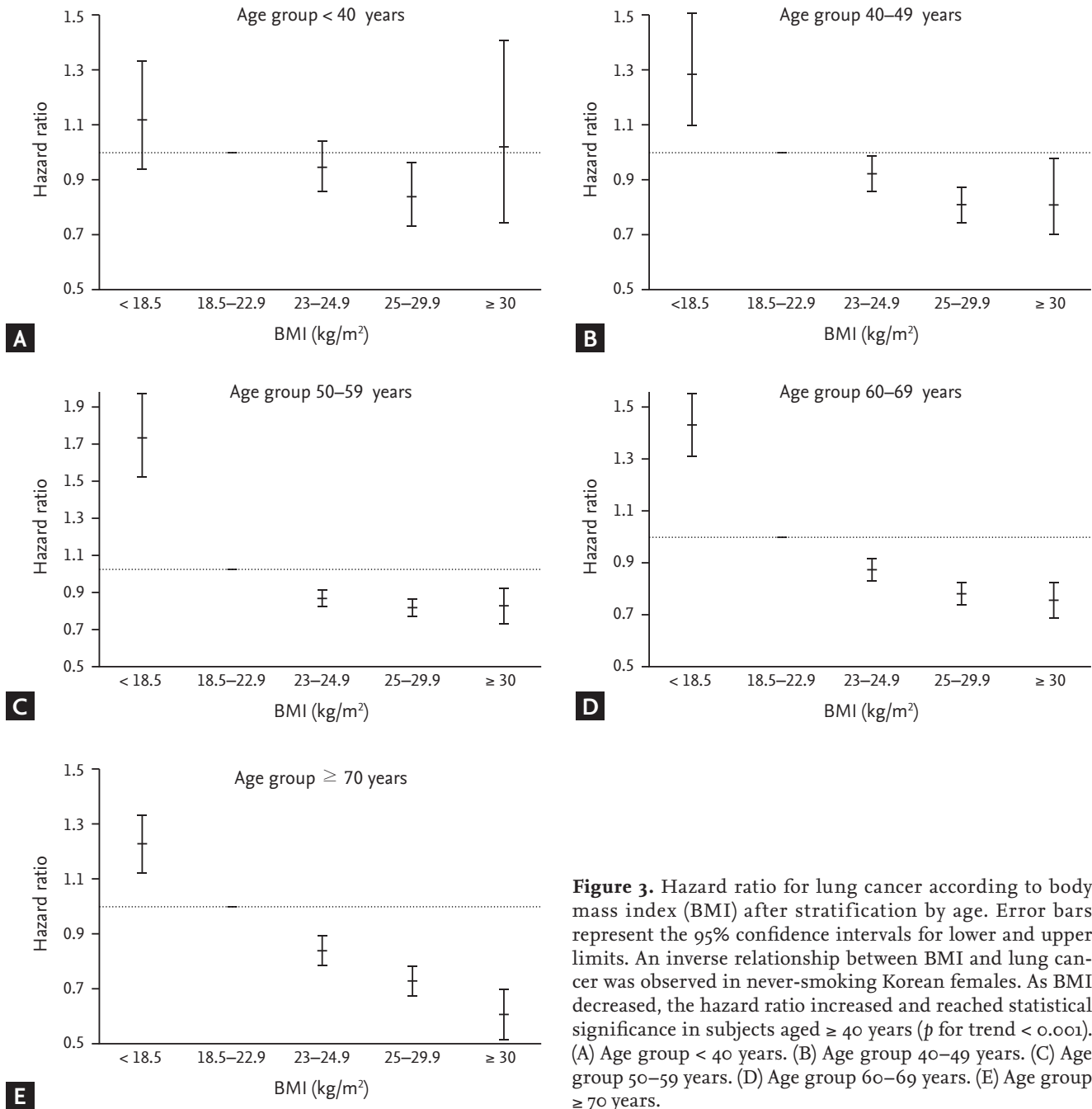
Figure 3. Hazard ratio for lung cancer according to body mass index (BMI) after stratification by age. Error bars represent the 95% confidence intervals for lower and upper limits. An inverse relationship between BMI and lung cancer was observed in never-smoking Korean females. As BMI decreased, the hazard ratio increased and reached statistical significance in subjects aged ≥ 40 years (p for trend < 0.001). (A) Age group < 40 years. (B) Age group 40–49 years. (C) Age group 50–59 years. (D) Age group 60–69 years. (E) Age group ≥ 70 years.

the methods used to assess BMI and analyze the data, and high heterogeneity across studies. However, this study enrolled a large number of never-smoking Korean females, nearly 6 million; thus, the 95% CIs were narrower than those in previous pooled studies.

BMI itself does not take into account specific body composition metrics such as muscle mass, visceral fat mass, and subcutaneous fat mass. Moreover, as body composition and BMI differ considerably among different ethnic groups [12], patients with the exact same BMI can have significantly different body compositions and different clinical outcomes. For prostate cancer, although gaining weight was positively associated with prostate cancer risk in Western populations, there was

Ko YH, et al. Lung cancer and never-smoking women

KJIM

Table 3. Multivariate Cox regression analysis of dietary habits after adjusting for all dependent variables

Diet	Age, yr				
	< 40	40–49	50–59	60–69	≥ 70
Vegetables	1	1	1	1	1
Vegetables + meat	0.94 (0.83–1.06)	0.95 (0.90–1.01)	1.04 (0.99–1.08)	1.05 (1.02–1.09)	1.12 (1.02–1.17)
Meat	0.91 (0.72–1.15)	0.93 (0.79–1.10)	1.03 (0.89–1.18)	1.19 (1.06–1.33)	1.28 (1.12–1.46)

Values are presented as hazard ratio (95% confidence interval).

Figure 4. Hazard ratios for lung cancer according to alcohol consumption after stratification by age.

an inverse association between weight gain and disease risk in Asian populations [13]. In the present study, a multivariate Cox regression model with age stratification also showed a robust inverse relationship between BMI and lung-cancer development in never-smoking Korean females aged ≥ 40 years.

There are several possible explanations for the inverse relationship between BMI and lung-cancer development. BMI is a strong surrogate for adipose tissue volume. First, adipose tissue modulates the storage of extrinsic potential carcinogens such as benzo(a)pyrene, which induces DNA adduct formation and prevents the accumulation of carcinogen-DNA adducts in target organs [14-17]. Second, the fat mass- and obesity-associated genotype, which accounts for the greatest genetic variance in obesity traits over a lifespan, is associated with a reduced risk of lung cancer [18]. Third, smoking-induced weight loss may be associated with lung cancer [19,20]. However, the negative association between BMI and lung cancer is not clearly understood. Therefore, a future

study is needed to evaluate the causal relationship.

Alcohol consumption has been suggested as a potential carcinogen involved in lung cancer [21]. Acetaldehyde, an alcohol metabolite, forms DNA adducts *in vitro* [22], and alcohol enables potential carcinogens to dissolve in the upper aerodigestive mucosal layers [23]. In addition, the frequency of alcohol consumption potentially functions as a surrogate for the frequency of secondhand smoking, which indicates the amount of exposure to environmental tobacco smoke [24]. Although the Korean government established a National Health Promotion Act in 1995, smoking in public places such as restaurants, pubs, and bars was not banned until 2013. Thus, individuals who frequented such establishments in this study might have been exposed to more secondhand tobacco smoke. In addition, drinking alcohol at least one to two times per week increased the risk of lung cancer in older compared with younger never-smoking females, which might reflect higher exposure to secondhand tobacco smoke in previous decades.

The HR for dietary habits indicated a positive association with lung-cancer development after adjusting for baseline variables including age and BMI. Cox regression analysis with stratification by age showed that consuming a meat-based diet increased the risk of lung-cancer development in the elderly (aged 50 or older) compared with the younger group (aged 50 or younger) (Table 3). Dietary habits may be closely associated with age, and elderly subjects may have had greater exposure to cooking exhaust over several decades. Levels of particulate matter in cooking exhaust from fuels and meats, which contain carcinogens such as benzo(a)pyrene [25], can reach 1,000 μg/m^3 in 20 minutes [26]. Thus, never-smoking females who prefer meat and who cook indoors may be exposed to high levels of carcinogens during cooking.

The HR for diets high in meat was negatively associated with lung-cancer in the unadjusted model. However,

KJIM

The Korean Journal of Internal Medicine. Vol. 35, No. 3, May 2020

an alternative result was obtained after adjustment for all variables. Age was the most important factor for the incidence of lung cancer in the present study. Therefore, adjusting for age may play an important role in changing the direction of the association between dietary habits and lung cancer. Univariate Cox regression analysis to validate the effects of dietary habit with stratification according to age group, as outlined in Table 3, showed a different direction of association between a diet high in meat and lung cancer incidence according to age group (the results were similar to those in Table 3; data not shown).

With regard to residence area, the percentages of subjects residing in a rural area were 38.3% for those aged ≥ 70 years, 31.2% for those aged 60 to 69 years, and 16.2% for those aged < 40 years. This implies that living in a rural versus urban area is associated with age. However, additional analyses stratified by age implied that the relative risk between urban and rural residence was not a differential misclassification bias. Individuals residing in rural locations may have greater exposure to potential carcinogens such as arsenics in pesticides [27], asbestos, and herbicides [28], which increase the risk of lung-cancer development.

A previous history of cancer other than lung cancer was also a risk factor for developing lung cancer. Genetic predisposition and environmental exposure to common carcinogens may be related to the development of multiple cancers. In addition, secondary lung cancer after cancer at other sites might be associated with previous chemotherapy or radiation therapy [29,30].

There are several limitations to this study. First, several potential information biases might exist. Baseline lifestyle habits (exercise, alcohol consumption, and diet) were assessed by a previous questionnaire, which used a limited number of response options. Therefore, detailed information regarding lifestyle, such as type and amount of alcohol consumption, exercise type, etc., was not assessed because of the use of a questionnaire prepared by the NHIS for health screening in the general population. In addition, a family history of lung cancer is also an important factor. However, clustering of familial lung cancer was not fully assessed in the questionnaire. Further studies with reduction of these sources of bias are required. In addition to these biases, there may also have been additional bias associated with validation of

secondhand smoking (SHS) status. The questionnaire for the NHIS health examination does not include SHS, so its effect could not be evaluated in this study. Second, in comparison to the Korea Central Cancer Registry, in which cases were confirmed pathologically, the operational definition of lung cancer was derived from the Disease Code dataset (T40) of the NHIS. To overcome this information bias, a multistep approach was performed to identify newly diagnosed lung cancer (C34 criteria 2 in Fig. 2). Furthermore, in cases of specific severe disease (cancer or rare intractable diseases), the NHIS dataset covers more than 90% of the population, because medical care expenses are reimbursed up to 90% to 95% of the total amount due. A previous study that compared medical claim data in the Health Insurance and Review Agency with chart reviews for rare and intractable diseases indicated 97% to 98% and 92% to 93% sensitivity and specificity, respectively [31]. Thus, a misclassification bias of lung cancer using the NHIS claim dataset is less likely.

This study also has several strengths. First, the main strengths are the large sample size of nearly 6 million subjects and the long median follow-up period of 11.5 years. Second, the effect of smoking was excluded using strict inclusion/exclusion criteria so that other potential risks for lung cancer could be investigated in never-smoking females. Third, BMI and lung cancer showed a statistically significant inverse dose–response relationship. However, a higher BMI *per se* is a potential risk for other cancers such as breast, ovarian, and endometrial cancer; therefore, maintaining an adequate BMI may be useful for cancer prevention.

In conclusion, modifiable baseline characteristics such as BMI, exercise, alcohol consumption, and diet are risk factors for lung-cancer development among never-smoking females. Thus, lifestyle modification may prevent lung cancer.

KEY MESSAGE

1. The 0.74% participants among non-smoking females were newly diagnosed with lung cancer.
2. Low body mass index (BMI), alcohol consumption, and a meat-based diet were risk factors for lung-cancer development among non-smoking females.

Ko YH, et al. Lung cancer and never-smoking women

KJIM

3. Therefore, lifestyle modifications of baseline characteristics, such as BMI, exercise, alcohol consumption, and diet, may help prevent lung cancer among non-smoking females.

Conflict of interest

No potential conflict of interest relevant to this article was reported.

REFERENCES

1. Global Burden of Disease Cancer Collaboration, Fitzmaurice C, Akinyemiju TF, et al. Global, regional, and national cancer incidence, mortality, years of life lost, years lived with disability, and disability-adjusted life-years for 29 cancer groups, 1990 to 2016: a systematic analysis for the global burden of disease study. JAMA Oncol 2018;4:1553-1568.

2. Jung CY, Antonia SJ. Tumor immunology and immune checkpoint inhibitors in non-small cell lung cancer. Tuberc Respir Dis (Seoul) 2018;81:29-41.

3. In KH, Kwon YS, Oh IJ, et al. Lung cancer patients who are asymptomatic at diagnosis show favorable prognosis: a Korean Lung Cancer Registry Study. Lung Cancer 2009;64:232-237.

4. Park JY, Jang SH. Epidemiology of lung cancer in Korea: recent trends. Tuberc Respir Dis (Seoul) 2016;79:58-69.

5. Choi S, Kim Y, Park S, Lee J, Oh K. Trends in cigarette smoking among adolescents and adults in South Korea. Epidemiol Health 2014;36:e2014023.

6. Sun S, Schiller JH, Gazdar AF. Lung cancer in never smokers: a different disease. Nat Rev Cancer 2007;7:778-790.

7. Duan P, Hu C, Quan C, et al. Body mass index and risk of lung cancer: systematic review and dose-response meta-analysis. Sci Rep 2015;5:16938.

8. Song X, Pukkala E, Dyba T, et al. Body mass index and cancer incidence: the FINRISK study. Eur J Epidemiol 2014;29:477-487.

9. Smith L, Brinton LA, Spitz MR, et al. Body mass index and risk of lung cancer among never, former, and current smokers. J Natl Cancer Inst 2012;104:778-789.

10. Kabat GC, Kim M, Hunt JR, Chlebowski RT, Rohan TE. Body mass index and waist circumference in relation to

lung cancer risk in the Women's Health Initiative. Am J Epidemiol 2008;168:158-169.

11. Kabat GC, Miller AB, Rohan TE. Body mass index and lung cancer risk in women. Epidemiology 2007;18:607-612.

12. Lim U, Ernst T, Buchthal SD, et al. Asian women have greater abdominal and visceral adiposity than Caucasian women with similar body mass index. Nutr Diabetes 2011;1:e6.

13. Hernandez BY, Park SY, Wilkens LR, Henderson BE, Kolonel LN. Relationship of body mass, height, and weight gain to prostate cancer risk in the multiethnic cohort. Cancer Epidemiol Biomarkers Prev 2009;18:2413-2421.

14. Rundle A, Madsen A, Orjuela M, et al. The association between benzo[a]pyrene-DNA adducts and body mass index, calorie intake and physical activity. Biomarkers 2007;12:123-132.

15. Godschalk RW, Feldker DE, Borm PJ, Wouters EF, van Schooten FJ. Body mass index modulates aromatic DNA adduct levels and their persistence in smokers. Cancer Epidemiol Biomarkers Prev 2002;11:790-793.

16. Godschalk R, Nair J, van Schooten FJ, et al. Comparison of multiple DNA adduct types in tumor adjacent human lung tissue: effect of cigarette smoking. Carcinogenesis 2002;23:2081-2086.

17. Garcon G, Garry S, Gosset P, et al. Benzo(a)pyrene-coated onto Fe(2)O(3) particles-induced lung tissue injury: role of free radicals. Cancer Lett 2001;167:7-15.

18. Brennan P, McKay J, Moore L, et al. Obesity and cancer: Mendelian randomization approach utilizing the FTO genotype. Int J Epidemiol 2009;38:971-975.

19. Ruige JB, Dekker JM, Blum WF, et al. Leptin and variables of body adiposity, energy balance, and insulin resistance in a population-based study. The Hoorn Study. Diabetes Care 1999;22:1097-1104.

20. Hodge AM, Westerman RA, de Courten MP, Collier GR, Zimmet PZ, Alberti KG. Is leptin sensitivity the link between smoking cessation and weight gain? Int J Obes Relat Metab Disord 1997;21:50-53.

21. Garcia Lavandeira JA, Ruano-Ravina A, Kelsey KT, et al. Alcohol consumption and lung cancer risk in never smokers: a pooled analysis of case-control studies. Eur J Public Health 2018;28:521-527.

22. Vaca CE, Nilsson JA, Fang JL, Grafstrom RC. Formation of DNA adducts in human buccal epithelial cells exposed to acetaldehyde and methylglyoxal in vitro. Chem Biol

Interact 1998;108:197-208.

23. Wight AJ, Ogden GR. Possible mechanisms by which alcohol may influence the development of oral cancer: a review. Oral Oncol 1998;34:441-447.

24. Reijula J, Kjaerheim K, Lynge E, et al. Cancer incidence among waiters: 45 years of follow-up in five Nordic countries. Scand J Public Health 2015;43:204-211.

25. He LY, Hu M, Huang XF, Yu BD, Zhang YH, Liu DQ. Measurement of emissions of fine particulate organic matter from Chinese cooking. Atmos Environ 2004;38:6557-6564.

26. Schiavon M, Rada EC, Ragazzi M, Antolini D. Indoor measurements of particulate matter during steak cooking under different conditions. In: Brebbia CA, Marinov AM, Safta CA, eds. Energy and Sustainability IV. Southampton (UK): WIT Press, 2013;255-265.

27. Bencko V, Yan Li Foong F. The history of arsenical pes-ticides and health risks related to the use of Agent Blue. Ann Agric Environ Med 2017;24:312-316.

28. Mostafalou S, Abdollahi M. Pesticides: an update of human exposure and toxicity. Arch Toxicol 2017;91:549-599.

29. Travis LB, Gospodarowicz M, Curtis RE, et al. Lung cancer following chemotherapy and radiotherapy for Hodgkin's disease. J Natl Cancer Inst 2002;94:182-192.

30. Huang YJ, Huang TW, Lin FH, Chung CH, Tsao CH, Chien WC. Radiation therapy for invasive breast cancer increases the risk of second primary lung cancer: a nationwide population-based cohort analysis. J Thorac Oncol 2017;12:782-790.

31. Kim HJ, Hann HJ, Hong SN, et al. Incidence and natural course of inflammatory bowel disease in Korea, 2006-2012: a nationwide population-based study. Inflamm Bowel Dis 2015;21:623-630.