Tobacco Smoking Modifies the Association between Hormonal Factors and Lung Cancer Occurrence among Post-Menopausal Chinese Women

Abstract
Inconsistent evidence has been reported on the role of female hormonal factors in the development of lung cancer. This population-based case–control study evaluated the main effect of menstrual/reproductive factors on the risk of lung cancer, and the effect modification by smoking status. Multivariable unconditional logistic regression models were applied adjusted for age, income, education, county of residence, body mass index, smoking status, pack-years of smoking, and family history of lung cancer. Among 680 lung cancer cases and 1,808 controls, later menopause (at \( \geq 54 \text{ vs.} <46 \text{ years old} \)) was associated with increased risk of lung cancer (SBOR, \text{semi-Bayes adjusted odds ratio} = 1.61, 95% PI, posterior interval = 1.10–2.36). More pregnancies (2 or 3 vs. 0 or 1) was associated with decreased risk (SBOR = 0.71, 95% PI = 0.53, 0.95). Ever being a smoker and having two or fewer pregnancies in one's lifetime could jointly increase the odds of lung cancer (RERI, relative excess risk due to interaction = 1.71, 95% CI = 0.03, 3.38). An increased number of ovulatory cycles was associated with increased risk of lung cancer (SBOR for 13 ovulatory cycles = 1.02, 95% CI = 1.00+, 1.04).

Introduction
Lung cancer is the most common and deadly cancer with the highest number of new cases (2,093,876) and deaths (1,761,007) in 2018 according to GLOBOCAN estimation [1]. It was also estimated that 37.0% of new lung cancer cases and 39.2% of lung cancer deaths happened in China in 2018 [2]. According to the CONCORD-3 study [3] and SEER project [4], the 5-year survival rate of lung cancer was 19.8% in China and 21.2% in the USA.
Tobacco smoking is believed to be the leading cause of lung cancer [5], with very high attributable risks (about 90% for men and 60% for women [6]) and odds ratios (23.6 for current male smokers and 7.8 for current female smokers respectively [6,7]. Smoking is more strongly associated with squamous cell carcinoma (SCC) and small cell carcinoma than with adenocarcinoma.

In smokers, pack-years [7–9] of smoking, cumulative tar exposure in tobacco products [10] and smoking intensity [7] are positively associated with the risk of lung cancer. Among non-smokers, environmental tobacco smoking (ETS or second hand smoking) has been shown associated with increased risk of lung cancer for all major histologic types [9].

Men and women differ in terms of lung cancer risk, survival, histology, and genetics. In the US, the lung cancer incidence rate in men peaked around the 1980s and decreased since the 1980s. On the other hand, rates continued to rise among women until 2010 [11] and started to drop in the past couple of years [12]. Women tend to have increased risk of getting lung cancer among non-smokers [13–15] and women have better lung cancer survival than men globally. According to GLOBOCAN estimation in 2018 [1], 34.6% of new lung cancer cases and 32.7% lung cancer deaths happened in women. There is a lower proportion of SCC and a higher proportion of adenocarcinoma in women than in men [14]. In NSCLC (non-small cell lung cancer) patients, EGFR (Epidermal growth factor receptor) mutations are more frequent in female than in male (more than 40% vs. less than 15%) [16]. Among current smokers, female patients had a 3.9-fold median level of CYP1A1 (Cytochrome P450 1A1) mutation compared to males [17]. After all measurable lifestyle and unchangeable factors are accounted for, the occurrences of lung cancer are still imbalanced between men and women [13–15]. These facts suggested that female sex hormones, especially estrogen, may have played a role in the initiation and progression of lung tumors [13,14].

Estrogen has long been associated with cancer development [18,19]. The amount of circulating estrogen and exogenous estrogen determines the effect of ERs (estrogen receptor) on cell proliferation and carcinogenesis. Menstrual and reproductive factors are commonly used as proxies for life-long endogenous estrogen exposure in women, while oral contraceptive use (OC) and hormone replacement therapy (HRT) represent exogenous sources of estrogen.

The results from epidemiologic studies on hormonal factors and the risk of lung cancer [20–27] were generally inconsistent [28]. In order to improve the validity and precision in the investigation of the associations between menstrual/reproductive history and the risk of lung cancer, and to investigate the effect modification by smoking status, we conducted the present analysis using data from the Jiangsu Four Cancers (JFC) Study.

**Materials and Methods**

**Population**

The JFC Study is a large-scale, population-based case–control study of cancers of the lung, liver, stomach, and esophagus conducted in four counties in Jiangsu, China. It was carried out in an effort to obtain high-quality data to investigate the lifestyle, environmental, and genetic factors associated with the four major cancers in China [29].

The incident primary lung cancer cases were reported by CDC-managed local cancer registries between January 1, 2003 and December 31, 2010. Diagnoses were either pathologically or clinically confirmed within 1 year of interview. Cases were required to be female, at least 18 years old, residents of the respective county for at least 5 years, and in a stable medical condition as determined by their physicians. Premenopausal cases were excluded because their estrogen biosynthesis and metabolism are different from postmenopausal women [30,31]. Post-menopausal cases who had undergone induced menopause (menopause due to surgery, radiotherapy, and other reasons) were also excluded [29]. There were a total of 680 cases eligible for the present analysis.

Controls of the JFC Study were females randomly selected from the same county as the cases. The JFC Study individually matched controls to cases by age (+/- 5 years). In the present analysis, all postmenopausal female controls from controls for four cancer sites were combined in order to increase statistical power. The same as cases, controls are those with natural menopause only. A total of 1808 controls were eligible for the present analysis.

A standardized epidemiological questionnaire including demographic characteristics, social economic status and menstrual/ reproductive history, and other risk or protective factors were employed to collect data for both cases and controls by face-to-face interview. The interviewers were trained and the questionnaire was field tested as detailed in a previous study [32]. The quality of participants’ answers was validated logically by repeated or related questions. A variable was coded as missing if the participant had inconsistent answers to the repeated/related questions.

**Exposure Ascertainment**

The standard epidemiological questionnaire collected information on exposures of interest, which were hormonal factors including age at menarche and menopause, parity, gravidity, live birth, outcome of first pregnancy, induced abortion, and oral contraceptive (OC) use. The outliers of ages at menarche and menopause, as well as contradictory answers in reference to current age, were treated as missing data. The normal distributions of ages at menarche and menopause were compared to large observational studies conducted in China among women with similar years of births in the JFC Study [33,34]. Three-sigma rule of thumb [35] was used to identify outliers for these two variables. Reproductive window was calculated as the difference between age at menopause and menarche. Gravidity and parity are the numbers of times a female has been pregnant and carried the pregnancies to a viable gestational age, respectively. Therefore, gravidity was calculated as the sum of numbers of miscarriages, abortions, live births, stillbirths and all other outcomes of pregnancy. Parity was calculated as the total number of live births and stillbirths [36]. Missing data in number of live births, miscarriages, induced abortions and stillbirths were imputed with the medians in the control group. The number of ovulatory cycles was calculated based on the reproductive window, subtracted by the length of time without ovulation due to OC use, live births, stillbirths, miscarriages, induced abortions, and breast feeding [37,38]. It was calculated assuming 36, 28, 12, and 12 weeks of no ovulatory cycles due to a live birth, a stillbirth, a miscarriage, and an induced abortion, respectively. It was also assumed that there was no ovulation during OC use. If the participant’s answer to the breastfeeding question was 'breast feeding only', 'no breast feeding', and 'mixed feeding', the length of time without ovulatory cycles after delivery was assumed to be 24, 6, and 9 months, respectively. Twenty-eight days was applied as the average length of an ovulatory cycle and 52.178 weeks in a year was assumed. In the study of
interaction with smoking status, selected menstrual and reproductive factors were dichotomized in order to calculate RERI (relative excess risk due to interaction) and ROR (ratio of the odds ratios).

The questionnaire also collected information on demography and covariates: age at interview, county of residence, income, education, smoking status, pack-years of smoking, body mass index (BMI), and family history of lung cancer. Age at interview was calculated based on self-reported date of birth. If the participant only remembered the lunar month of birth, the following solar month was used as the proxy. There were no missing data for year of birth. In cases of missing month or date of birth, the median values (July and 15th, respectively) were used as the estimations. BMI cut points were chosen for underweight (<18.5 kg/m²), normal weight (18.5 kg/m² to <24 kg/m²), overweight (24 kg/m² to <28 kg/m²) and obesity (≥ 28 kg/m²) according to the standards for Chinese populations [39]. Ever smoking was defined as having smoked more than 100 cigarettes in one’s lifetime. The missing values for pack-years of smoking (missing rate = 12.6%) were imputed with the county, sex, and age-specific median values of the controls. Family history of lung cancer was defined as lung cancer diagnosis in any family member including parents, grandparents, siblings, children, spouses, and parents’ siblings. Income was measured as the per-capita annual income of the household including wages, bonuses, and allowances on an average over the most recent decade.

**Statistical Analysis**

Cases and controls were compared using the χ² test for categorical variables and the Student t-test for continuous variables. The Mantel trend test was used to assess whether an association was exhibiting a linear trend. The main effect was tested with multivariable unconditional logistic regression models, adjusting for covariates. According to a priori biological rationale [5,27,40,41], age, county, BMI, exposure to tobacco smoking, pack-years of smoking, family history of lung cancer, income, and education were adjusted for. Stratified analysis by smoking status was performed. In addition, age at menarche was adjusted for in the analysis for age at menopause and age at first birth; length of reproductive window was adjusted for in the analysis for parity/gravidity/number of children; age at first birth was adjusted for in the analysis for outcome of first birth. Odds ratios and 95% confidence intervals were calculated for all the possible interactions. Semi-Bayes estimates and 95% posterior intervals were calculated to mitigate the influence from sparse data and multiple associations. Semi-Bayes estimates and 95% posterior intervals were calculated to mitigate the influence from sparse data and multiple comparisons [42]. In the semi-Bayes estimation, the priors of coefficients in the logistic regressions were assigned normal distributions with mean of zero and variance of 0.5 (corresponding to OR = 1, 95% PI = 0.25–4). And the priors were updated with observed values from the study population. Therefore, the SB adjusted ORs have shrunk towards the null. The effect modification by tobacco smoking was calculated by RERI for the additive scale and was by ROR for the multiplicative scale. In these calculations, age, county, BMI, family history of lung cancer, income, and education were adjusted for.

Complete case analyses were applied except where imputation was described in “Exposure ascertainment”. The missing rates of those variables with complete case analyses were all less than 10%. SAS version 9.4 was used for statistical analysis. All p-values were based on α = 0.05 (two-sided). The JFC study was approved by the Human Subject Protection Committees of the Jiangsu Provincial Center for Disease Control (CDC) and the University of California at Los Angeles (UCLA). Informed consent was obtained from all participants upon study enrollment.

**Results**

**Summary Statistics**

The average age in this post-menopausal female population was 67.21 years. Table 1 summarized the distribution of demographic and major risk factors. Cases and controls showed similar average ages and education/income distributions. Ganyu and Tongshan Counties contributed more lung cancer cases while Dafeng and Chuzhou counties contributed more controls. Significantly higher proportions of smokers and larger pack-years were observed in lung cancer cases than in controls. Cases were significantly more likely to be overweight or less likely to be overweight or obese than controls. Family history of lung cancer tended to happen among lung cancer cases rather than controls.

**Menstrual Characteristics**

According to Table 2-1, there was no significant statistics for the relationship between age at menarche and risk of lung cancer in the entire study population. However, we found an 35% increased odds (Table 2-2) among never smoking subpopulation, comparing age at menarche between 16 and 17 years old with that ≤15 years old (SBOR: semi-Bayes adjusted odds ratio = 1.35, 95% PI, posterior interval = 1.01–1.80). The Mantel test P-value was 0.03, indicating a

| Table 1. The Distribution of Demographic and Major Risk Factors in Cases and Controls |
|-----------------------------------|-------------------|-------------------|
| Total Number                      | Lung Cancer       | Controls          | P      |
| Mean(SD)                          | 680 (27.33)       | 1808 (72.67)      | .1641  |
| Age (continuous)                  |                    |                   | .1783  |
| N(column%)                        | 66.79 (9.36)      | 67.37 (9.32)      |        |
| Education                         |                    |                   |        |
| Illiterate                        | 554 (81.71)       | 1431 (79.50)      | .0001  |
| Primary School                    | 106 (15.63)       | 294 (16.33)       |        |
| Middle School +                   | 18 (2.65)         | 75 (4.17)         |        |
| Income                            |                    |                   | .1783  |
| <1000                             | 130 (19.88)       | 416 (23.44)       |        |
| 1000–1499                         | 131 (20.03)       | 382 (21.52)       |        |
| 1500–1999                         | 177 (27.06)       | 459 (25.86)       |        |
| ≥ 2000                            | 216 (33.03)       | 518 (29.18)       |        |
| County                            |                    |                   |        |
| Dafeng                            | 152 (22.35)       | 597 (33.02)       |        |
| Ganyu                             | 185 (27.21)       | 337 (18.64)       |        |
| Chuzhou                           | 120 (17.65)       | 365 (20.19)       |        |
| Tongshan                          | 223 (32.79)       | 509 (28.15)       |        |
| Tobacco smoking                   |                    |                   | .0001  |
| Ever                               | 181 (26.62)       | 344 (19.03)       |        |
| Never                             | 499 (73.38)       | 1464 (80.97)      |        |
| Pack-year                          |                    |                   | .0001  |
| <10                               | 499 (73.38)       | 1464 (80.97)      |        |
| 10–20                             | 25 (3.68)         | 68 (3.76)         |        |
| 20–30                             | 24 (3.53)         | 63 (3.48)         |        |
| 30–40                             | 51 (7.50)         | 91 (5.03)         |        |
| 40–50                             | 19 (2.79)         | 45 (2.49)         |        |
| 50–60                             | 25 (3.68)         | 33 (1.83)         |        |
| ≥ 60                              | 12 (1.76)         | 24 (1.33)         |        |
| BMI                               |                    |                   | .0001  |
| <18.5                             | 115 (17.09)       | 155 (8.61)        |        |
| 18.5 to <24                       | 388 (57.65)       | 998 (55.41)       |        |
| 24 to ≤28                         | 137 (20.36)       | 497 (27.60)       |        |
| > 28                              | 33 (4.90)         | 151 (8.38)        |        |
| Family history of lung cancer     |                    |                   | .0117  |
| Yes                               | 25 (3.68)         | 35 (1.94)         |        |
| No                                | 655 (96.32)       | 1773 (98.06)      |        |
possible dose–response relationship. Compared with ages at meno-
pause between 46 and 54 years, those later than 54 years old were
associated with 1.61 times of odds of lung cancer occurrence in the
whole study population (Table 2-1, SBOR = 1.61, 95% PI =
1.10–2.36, Mantel test P-value = 0.023). Age at menopause also
showed 3% increased odds of lung cancer per one-year increase
(SBOR = 1.03, 95% PI = 1.01–1.06). This aforementioned associ-
ation for age at menopause seemed to exist among never smokers but
could not be found in the ever-smoking subpopulation (Table 2-2).
There was no evidence demonstrating an association between the
length of reproductive window and the risk of lung cancer.

Reproductive History

As reported in Table 2-1, a parity between two and three showed
30% and a parity of four or more showed 26% decreased odds in lung
cancer occurrence, (parity = 2 or 3: SBOR = 0.70, 95% PI =
Table 2. Menstrual and Reproductive Factors in Association with the Risk of Lung Cancer, by Smoking Status

|                      | Never Smokers | Ever Smokers |
|----------------------|---------------|--------------|
|                      | Cases, n = 499 | Cases, n = 181 |
|                      | Ctrls, n = 1464 | Ctrls, n = 344 |
|                      | N % (95% CI) | N % (95% CI) |
|                      | N % (95% CI) | N % (95% CI) |

Menstrual Characteristics

Age at menarche

| Group | Never Smokers | Ever Smokers |
|-------|---------------|--------------|
|       | OR (95% CI)   | OR (95% CI)  |
| <=15  | 1.00 (Ref)    | 1.00 (Ref)   |
| 16-17 | 1.02 (0.91, 1.15) | 1.01 (0.90, 1.14) |
| >18   | 1.02 (0.92, 1.13) | 1.01 (0.90, 1.13) |

Age at menopause

| Group | Never Smokers | Ever Smokers |
|-------|---------------|--------------|
|       | OR (95% CI)   | OR (95% CI)  |
| <=40  | 1.00 (Ref)    | 1.00 (Ref)   |
| >40   | 1.02 (0.92, 1.14) | 1.01 (0.90, 1.13) |

Reproductive window

| Group | Never Smokers | Ever Smokers |
|-------|---------------|--------------|
|       | OR (95% CI)   | OR (95% CI)  |
| <=35  | 1.00 (Ref)    | 1.00 (Ref)   |
| >35   | 1.02 (0.92, 1.14) | 1.01 (0.90, 1.13) |

Reproductive History

Parity

| Group | Never Smokers | Ever Smokers |
|-------|---------------|--------------|
|       | OR (95% CI)   | OR (95% CI)  |
| 0 or 1| 1.00 (Ref)    | 1.00 (Ref)   |
| >1    | 1.02 (0.92, 1.14) | 1.01 (0.90, 1.13) |

Gravidity

| Group | Never Smokers | Ever Smokers |
|-------|---------------|--------------|
|       | OR (95% CI)   | OR (95% CI)  |
| 0 or 1| 1.00 (Ref)    | 1.00 (Ref)   |
| >4    | 1.03 (0.93, 1.14) | 1.02 (0.91, 1.13) |

Number of live birth

| Group | Never Smokers | Ever Smokers |
|-------|---------------|--------------|
|       | OR (95% CI)   | OR (95% CI)  |
| 0 or 1| 1.00 (Ref)    | 1.00 (Ref)   |
| >4    | 1.03 (0.93, 1.14) | 1.02 (0.91, 1.13) |

Life time abortion

| Group | Never Smokers | Ever Smokers |
|-------|---------------|--------------|
|       | OR (95% CI)   | OR (95% CI)  |
| 0     | 1.00 (Ref)    | 1.00 (Ref)   |
| >1    | 1.03 (0.93, 1.14) | 1.02 (0.91, 1.13) |

Outcome of first pregnancy

| Group | Never Smokers | Ever Smokers |
|-------|---------------|--------------|
|       | OR (95% CI)   | OR (95% CI)  |
| VLB   | 1.00 (Ref)    | 1.00 (Ref)   |
| SIB   | 1.02 (0.91, 1.14) | 1.01 (0.90, 1.13) |

Number of Ovulatory Cycles

| Group | Never Smokers | Ever Smokers |
|-------|---------------|--------------|
|       | OR (95% CI)   | OR (95% CI)  |
| <=366 | 1.00 (Ref)    | 1.00 (Ref)   |
| >366  | 1.03 (0.93, 1.14) | 1.02 (0.91, 1.13) |

Exogenous Hormone

Oral Contraceptive use

| Group | Never Smokers | Ever Smokers |
|-------|---------------|--------------|
|       | OR (95% CI)   | OR (95% CI)  |
| 0     | 1.00 (Ref)    | 1.00 (Ref)   |
| >1    | 1.03 (0.93, 1.14) | 1.02 (0.91, 1.13) |

Note:
1. Odds ratios and 95% confidence intervals adjusted for age (as a continuous variable), smoking status (ever or never), pack-years of smoking, family history of lung cancer (yes or no), income, education, county of residence, and BMI.
2. Odds ratios and 95% confidence intervals adjusted for age (as a continuous variable), family history of lung cancer (yes or no), income, education, county of residence, and BMI.
3. Odds ratios and 95% confidence intervals adjusted for age (as a continuous variable), pack-years of smoking, family history of lung cancer (yes or no), income, education, county of residence, and BMI.
4. Mantel trend test.
5. Additional adjustment for age at menarche (as a continuous variable).
6. Additional adjustment for length of reproductive window.
7. Additional adjustment for age at first birth.
8. Absolute number/count as the continuous variable.
Table 3. Interaction with Smoking Status

| Factor                                      | Case/Ctrl | aOR (95%CI)\(^1\) | Interactions (95%CI)\(^1\) |
|---------------------------------------------|-----------|--------------------|-----------------------------|
| Menarche at 17 or later                     | Ever smoking | 275/838 | 1.00 (Ref) |  |
| No                                          | No        | 94/170  | 2.36 (1.70, 3.27) | RERI = -0.43 (-1.36, 0.5) |
| Yes                                         | Yes       | 217/620 | 1.28 (1.02, 1.61) | ROR = 0.73 (0.47, 1.14)  |
| Yes                                         | No        | 86/169  | 2.21 (1.56, 3.12) |  |
| Menopause at 55 or later                    | Ever smoking | 367/1188 | 1.98 (1.51, 2.59) | RERI = 0.87 (-1.39, 3.12) |
| No                                          | No        | 146/308 | 1.60 (1.19, 2.15) | ROR = 1.09 (0.53, 2.25)  |
| Yes                                         | Yes       | 95/185  | 3.45 (1.80, 6.60) |  |
| Parity                                      | Ever smoking | 22/24   | 1.00 (Ref) |  |
| Parity> 3                                   | No        | 335/1069 | 1.86 (1.40, 2.48) | RERI = 0.96 (-0.39, 2.3) |
| Parity> 3                                   | Yes       | 129/276 | 1.32 (1.03, 1.69) | ROR = 1.28 (0.76, 2.15)  |
| Parity = 0, 1 or 2                          | Yes       | 52/68   | 3.13 (2.03, 4.83) |  |
| Gravidity                                   | Ever smoking | 272/856 | 1.00 (Ref) |  |
| Gravidity> 3                                | No        | 352/1111 | 1.77 (1.33, 2.34) | RERI = 1.71 (0.03, 3.38) |
| Gravidity> 3                                | Yes       | 132/290 | 1.26 (0.98, 1.62) | ROR = 1.68 (0.98, 2.89)  |
| Gravidity = 0, 1 or 2                       | No        | 146/353 | 3.73 (2.36, 5.90) |  |
| Gravidity = 0, 1 or 2                       | Yes       | 49/54   | 0.73 (0.47, 1.14) |  |
| #live birth = 0, 1 or 2                     | Ever smoking | 333/1058 | 1.85 (1.39, 2.47) | RERI = 0.95 (-0.36, 2.27) |
| No                                          | No        | 128/273 | 3.0 (1.01, 1.66) | ROR = 1.29 (0.77, 2.16)  |
| Yes                                         | Yes       | 165/408 | 3.10 (2.02, 4.76) |  |

Notation:  
Point estimates and 95% confidence intervals were adjusted for age (as a continuous variable), family history of lung cancer (yes or no), income, education, county of residence, and BMI (categorical).

0.53–0.93, parity = 4 or more: SBOR = 0.74, 95% PI =0.55–0.99, reference group is parity = 1). The dose–response trend and change per unit increase in parity were not statistically significant (P\(_{\text{trend}}\) = .177 and the semi-Bayes posterior interval was across the null). Among never smoking or ever smoking subpopulation (Table 2-2), there seemed to be no association between parity and the risk of lung cancer. A moderate gravidity seemed to decrease risk of lung cancer by 29% (SBOR = 0.71, 95% PI = 0.53–0.95, gravidity = 2–3 compared with gravidity = 0 or 1) in the study population (Table 2-1). Among never smokers, gravidity did not show a significant relationship with the risk of lung cancer. On the other hand, treated as a continuous variable, a one-unit increase in gravidity was significantly associated with 11% decrease in risk of lung cancer (SBOR = 0.89, 95% PI = 0.81–0.99) for ever-smokers. A moderate number of live births was shown associated with 29% decreased risk of lung cancer in all the post-menopausal women (SBOR = 0.71, 95% PI = 0.43–0.94, number of live birth = 2–3 compared with 0–1). However, this statistical significance for number of live births was not observed in either ever- or never- smokers (Table 2-2).

Induced abortion, reported or not in Tables 2-1 and 2-2, did not show any significant associations with lung cancer. A statistically significant association between an increase of 13 ovulatory cycles (about 1 year) and 2% increase in the risk of lung cancer was shown in our post-menopausal study population (SBOR = 1.02, 95% PI = 1.00+, 1.04)) but not observed in subpopulations stratified by smoking status.

Exogenous Hormone Use  
Tables 2-1 and 2-2 didn’t show a relationship between OC use and risk of lung cancer.

Effect Modification by and Interaction with Smoking Status  
Tests for interaction with smoking status on additive and multiplicative scales were performed for menstrual and reproductive factors that showed statistically significant main effects. As reported in Table 3, gravidity at or below two showed an RERI of 1.71 with 95% CI of 0.33–3.38. The ROR of gravidity was 1.68 without statistical significance. These RERIs and RORs were suggesting superadditivity for the interaction between smoking and gravidity (Table 3).

Discussion  
In the present analysis of hormonal factors, after controlling for potential confounders and correction by semi-Bayes shrinkage, later menopause was found to be associated with increased risk of lung cancer. This relationship remained in the never-smoking subpopulation, but disappeared in the ever-smoking subpopulation. A higher parity, gravidity, and number of live births were respectively associated with reduced risk of lung cancer. Increased number of ovulatory cycles was associated with increased risk of lung cancer. Other hormonal factors such as reproductive window, number of abortions, and outcomes of first pregnancy were not associated with lung cancer risk in our study population.

The stratified analyses for never- and ever-smokers indicated that reproductive factors might interact with smoking status in the development of lung cancer. Superadditivity was corroborated by RERIs, showing a considerably greater joint effect of smoking and low gravidity than expected under an additive model without interactions.

Between 1/1/1988 and 7/31/2018, there were a total of 27 epidemiologic studies [20–27] that tested for associations between hormonal factors and risk of lung cancer. A meta-analysis [26] showed a statistically significant decreased risk of lung cancer with older age at menarche among Caucasian-dominated North American female populations. The effect of hormonal factors on lung cancer varies by race/ethnicity and is inconsistent among Asian populations. In our JFC Study, a menarche age greater than 18 years old could be a marker of poor childhood nutritional status, which has long-term adverse influence on health [23,43,44]. This increased risk by later
menopause was also found in six other studies conducted among Chinese women [23,25,27,45–47] although there was one study with potentially decreased risk linked to late menarche [22].

A greater menopausal age was proposed to increase the risk of lung cancer since a greater menopausal age means more exposure to estrogen [18,19]. However, an ILCCO pooled analysis [20] with a considerable number of missing values for exposure variables didn’t show a significant result for this relationship. An Asian (Singapore, SBCSP) cohort study [23] showed null associations without adjusting for smoking intensity or pack-years of smoking, leaving potential residual confounding by smoking. In this present JFC Study, where missing information was minimal and pack-years of smoking was adjusted, significantly increased risk by greater age at menopause was found. This finding was corroborated by two other studies among Chinese women [25,27].

In this present study, higher parity, gravidity, and live births were associated with decreased risk of lung cancer, which is consistent with all other Asian studies [23,25,27]. Parity and gravidity take into consideration the effect by miscarriage, abortion, stillbirth, and live birth. They cause an estrogen surge and accumulation for a period of time and reduce the number of ovulatory cycles. Our study was the first to report an increased risk associated with increased number of ovulatory cycles, supporting the hypothesis that regular dynamics of estrogen during normal ovulatory cycles, rather than accumulative endogenous or exogenous estrogen exposure, might increase the risk of lung cancer.

It has long been believed that lung cancers not related to smoking are different from those related to it [15,48]. This JFC study with post-menopausal women is the first study to report the effect modification by smoking, showing the risk smoking added to those with fewer pregnancies was greater than that added to those with more pregnancies.

Estrogen is thought to have an effect on lung cancer via estrogen receptors (ERs). Estrogen receptors α and β (ERα and ERβ), the two major types of ERs, are found in bronchial and alveolar epithelia and airway smooth muscle [14]. Both ERα and ERβ are ligand-activating transcription factors activated by 17-β estradiol (E2), the activation form of estrogen in human. The binding of E2 to ERs leads to dimerization and nuclear translocation of theses ERs. In the nucleus, ligand-bound ERα/ERβ dimers bind to the estrogen response elements (ERE) in the promoters of target genes to control cell proliferation, differentiation and apoptosis. ERs also associate with and activates EGFR (epidermal growth factor receptor, a receptor tyrosine kinase) thus triggering MAPK/ERK (mitogen-activated protein kinases/extracellular signal-regulated kinases) pathway and up/down regulating the transcription of genes that promote proliferation and invasion of lung cancer cells [49].

Strengths of this current study included homogeneity in terms of race/ethnicity, a large sample size, and a large proportion of non-smokers, which are important to reduce bias. The weak effect of menstrual and reproductive factors on lung cancer could be undetectable in a population dominated by smokers because of the strong association between smoking and lung cancer [27]. This study design also minimized selection bias by having population-based controls instead of hospital-based designs and by having very low missing rates of exposure variables of interest. In previous studies of the same topics, multiple menstrual and reproductive factors have been tested at the same time within one analysis, possibly resulting in false positive findings from multiple comparisons without correction.

In our study, semi-Bayes shrinkage was applied to mitigate false positive results by updating independent null priors for regression coefficients with observed data [50]. Semi-Bayes estimates were calculated also to improve the sparse data problem [42].

The weaknesses of this study included a lack of histologic categorization of lung cancer cases. The suggested association between hormonal factors and lung cancer has been identified most prominently among adenocarcinoma with EGFR mutations [45,51]. In this present study, lung cancer cases were identified from local CDC cancer registries. Due to the small proportion (20.4% cases) of lung cancer patients who had undergone surgeries or endoscope exams, there was not enough pathology or cytology information collected. Among these patients with known histology information, 65% of the NSCLC cases were adenocarcinoma. Among other studies conducted in China, lung cancer histology was largely dominated by adenocarcinoma, accounting for 61% to 73% of all female lung cancer diagnoses [22,51]. East Asian patients show a much higher prevalence of epidermal growth factor receptor (EGFR) mutations, compared with Caucasian patients with NSCLC (approximately 30% vs. 7%, predominantly among patients with adenocarcinoma and never-smokers), thus showing a higher proportion of patients who are responsive to EGFR tyrosine kinase inhibitors (EGFR-TKIs) [28,52].

Secondly, there could be recall biases brought by the case–control study design. However, all lung cancer cases were diagnosed within 1 year of interview. In addition, the study was not initially designed to test the association between hormonal factors and lung cancer and the participants were never told so. Therefore, their recalls for their exposure information were relatively objective representations of their usual life before the diagnosis of lung cancer and the recall bias has been reduced to minimum. Lastly, the sample size in the interaction study was relatively small for some of the levels of the interacting factors. We were not able to verify the causal interaction between smoking status and those menstrual and reproductive factors.

Conclusion

For postmenopausal Asian women, later menopause, more lifetime ovulatory cycles, and fewer pregnancies were associated with increased risk of lung cancer. This incremental risk appeared larger among ever smokers than their never-smoking counterparts. The potential etiologic clues of estrogen in the occurrence of lung cancer need to be further explored by more epidemiologic studies with biomarkers measurements. The identification of relationships between hormonal factors and the risk of lung cancer could inform preventive strategies and therapeutic regimes. Although causal interaction was not verified, the effect modification by smoking status could potentially add rationale to tobacco smoking cessation interventions among certain female populations.

Acknowledgement

The authors thank all study participants for their time and effort. This project was partially funded by the Jiangsu Provincial Health Department (RC 2003090, PI: Dr. Jin-Kou Zhao); the National Institutes of Health, National Institute of Environmental Health Sciences, National Cancer Institute, Department of Health and Human Services (ES06718, ES011667, CA90833, CA077954, CA96134, DA11386, and CA09142); and the Alper Research Center for Environmental Genomics of the University of California, Los Angeles Jonsson Comprehensive Cancer Center. The authors declare no conflict of interest.
Translational Oncology Vol. 12, No. 6, 2019

[44] Rogers IS, Northstone K, Dunger DB, Cooper AR, Ness AR, and Emmett PM (2010). Diet throughout childhood and age at menarche in a contemporary cohort of British girls. Public Health Nutr 13(12), 2052–2063.

[45] Weiss JM, Lacey Jr JV, Shu XO, Ji BT, Hou L, Yang G, Li H, Rothman N, Blair A, and Gao YT, et al (2008). Menstrual and reproductive factors in association with lung cancer in female lifetime nonsmokers. Am J Epidemiol 168(11), 1319–1325.

[46] Seow A, Koh WP, Wang RW, Lee MC, and Yu MC (2009). Reproductive Variables, Soy Intake, and Lung Cancer Risk among Nonsmoking Women in the Singapore Chinese Health Study. Cancer Epidem Biomar 18(3), 821–827.

[47] Wuwilliams AH, Dai XD, Blot W, Xu ZY, Sun XW, Xiao HP, Stone BJ, Yu SF, Feng YP, and Ershow AG, et al (1990). Lung-Cancer among Women in North-East China. Br J Cancer 62(6), 982–987.

[48] Sun S, Schiller JH, and Gazdar AF (2007). Lung cancer in never smokers—a different disease. Nat Rev Cancer 7(10), 778–790.

[49] Chakraborty S, Ganti AK, Marr A, and Batra SK (2010). Lung cancer in women: role of estrogens. Expert Rev Respir Med 4(4), 509–518.

[50] Greenland S (2008). Multiple comparisons and association selection in general epidemiology. Int J Epidemiol 37(3), 430–434.

[51] Gao YT, Blot WJ, Zheng W, Ershow AG, Cheng WH, Levin LI, Rong Z, and Fraumeni JF (1987). Lung-Cancer among Chinese-Women. Int J Cancer 40(5), 604–609.

[52] Qin BM, Chen X, Zhu JD, and Pei DQ (2005). Identification of EGFR kinase domain mutations among lung cancer patients in China: implication for targeted cancer therapy. Cell Res 15(3), 212–217.