Long-term Lung Cancer Risk Associated with Sputum Atypia: A 27-Year Follow-up Study of an Occupational Lung Screening Cohort in Yunnan, China

Yaguang Fan[1], Zheng Su[2], Mengna Wei[3], Hao Liang[4], Yong Jiang[2], Xuebing Li[1], Zhaowei Meng[5], Ying Wang[6], Hongli Pan[1], Jinzhao Song[7], Youlin Qiao[2,8], and Qinghua Zhou[1,4]

ABSTRACT

Background: Sputum cytologic atypia is associated with increased lung cancer risk. However, little is known about the long-term magnitude and temporal trend of this risk.

Methods: An extended follow-up was conducted in a prospective screening cohort among occupational tin miners in Yunnan, China. Sputum samples were collected prospectively at baseline and 7 annual screenings since enrollment. The associations between sputum cytologic results from baseline screening, the first 4 consecutive rounds of sputum screening, and lung cancer risk were analyzed by time-varying covariate Cox regression model.

Results: A moderate or worse cytologic result was associated with a significantly increased lung cancer risk. This relative hazard significantly decreased over time. Compared with negative screening results, the adjusted hazard ratios of baseline-moderate or worse atypia, at least one moderate or worse atypia in the first 4 consecutive screening rounds during the first 10 years of follow-up were 3.11 [95% confidence interval (CI): 2.37–4.07], 3.25 (95% CI: 2.33–4.54) respectively. This association was stronger for persistent atypia (adjusted hazard ratio = 17.55, 95% CI: 8.32–37.03); atypia identified in the recent screening rounds (adjusted HR = 4.14, 95% CI: 2.70–6.35), and those were old in age, had higher level of smoking, occupational radon, and arsenic exposure. In terms of histology, this increased risk was significant for squamous cell carcinoma and small cell lung cancer.

Conclusions: Although decreasing over time, an increased lung cancer risk concerning moderate or worse sputum atypia can continue at least for 10 years.

Impact: Sputum atypia might be helpful for identifying high-risk individuals for screening, surveillance, or chemoprevention of lung cancer.

Introduction

Lung cancer is currently the leading cause of cancer-related death both in China and around the world (1, 2). Results of the National Lung Cancer Screening Trial (NLST) demonstrated a 20% reduced mortality due to lung cancer with low-dose computed tomography (LDCT) screening among heavy smokers compared with chest radiography (3). Subsequently, several randomized controlled trials have reported a mortality reduction for lung cancer with LDCT screening (4–6).

However, LDCT screening still raises some questions. Precisely identifying a population in high-risk of developing lung cancer could improve the ratio between the benefits and harms of lung cancer screening (7, 8), and several studies have reported that an individual's screening history might be helpful for risk stratification (9–11). Furthermore, data from NLST also suggested that the effectiveness of LDCT screening might vary according to histology, and the LDCT benefit was not observed for squamous carcinoma and small cell lung cancer (SCLC), which are mainly located in central airways that were insensitive for LDCT screening (12, 13).

Sputum cytology is a noninvasive test for lung cancer, especially central-airway tumors. Early trials did not observe a mortality reduction of lung cancer from screening with chest X-ray and sputum cytology (14, 15). Accordingly, sputum cytology is currently not recommended for lung cancer screening. However, a combined mortality analysis from these trials showed that a modest benefit might have been present (16). Besides, including our previous study, several cohort studies found that sputum atypia detected 4 or more years before diagnosis was associated with increased lung cancer risk (17–19), and it was also used for identifying a candidate for lung cancer screening or chemoprevention trials (20, 21). Usually, sputum cytologic screening was conducted periodically, and previous study found that the association between sputum atypia and lung cancer incidence was greatest for those sputum samples collected 5 months or less before the diagnosis of lung cancer (17). However, few studies report the magnitude and temporal change of lung cancer risk according to different sputum cytologic results in previous screening rounds, especially in occupational population. Besides, there was still

[1]Tianjin Key Laboratory of Lung Cancer Metastasis and Tumor Microenvironment, Tianjin Medical University General Hospital, Tianjin, China. [2]Department of Cancer Epidemiology, Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China. [3]Breast Cancer Center, Sun Yat-Sen Memorial Hospital, Sun Yat-Sen University, Guangzhou, Guangdong, China. [4]Lung Cancer Center, Lung Cancer Institute, West China Hospital of Sichuan University, Chengdu, Sichuan, China. [5]Department of Nuclear Medicine, Tianjin Medical University General Hospital, Tianjin, China. [6]Department of Radiology, Tianjin Medical University General Hospital, Tianjin, China. [7]Department of Mechanical Engineering & Applied Mechanics, School of Engineering and Applied Science, University of Pennsylvania, Philadelphia, Pennsylvania. [8]Center of Global Health, School of Population Medicine and Public Health, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China.

Note: Supplementary data for this article are available at Cancer Epidemiology, Biomarkers & Prevention Online (http://cebp.aacrjournals.org/).

Y. Fan and Z. Su contributed equally to this article.

Corresponding Authors: Youlin Qiao, Center of Global Health, School of Population Medicine and Public Health, Chinese Academy of Medical Sciences and Peking Union Medical College, No. 9 Dongdan Santiao, Dongcheng District, Beijing 100730, China. E-mail: qiaoy@cicams.ac.cn; and Qinghua Zhou, Lung Cancer Center, Lung Cancer Institute, West China Hospital of Sichuan University, No. 37 Guoxue Alley, Wuhou District, Chengdu City, Sichuan Province, China. E-mail: zhouqh135@163.com

Cancer Epidemiol Biomarkers Prev 2021;30:2122–9

doi: 10.1158/1055-9965.EPI-21-0339

controversy about this risk in relation to histology (17, 19). In this study, we focus on the long-term lung cancer risk of sputum atypia based on the extended follow-up of an occupational screening cohort in Yunnan, China.

Materials and Methods

Study design and population

In 1992, to establish a biological specimen bank for the identification and validation of markers of early lung cancer, a dynamic prospective cohort study among tin miners was initiated in the Yunnan Tin Corporation (YTC). A total of 9,295 tin miners ≥ 40 years old who had 10 or more years of underground radon exposure and/or arsenic exposure were enrolled from 1992 to 1998. Detailed information on inclusion criteria was described previously (19, 22). Written informed consent was obtained for each participant. This study was conducted in accordance with Declaration of Helsinki, and was approved by the institutional review board of Cancer Hospital/Institute of Chinese Academy of Medical Sciences (201812190401002).

Lung cancer screening and sputum cytology

From 1992 to 1999, 8 rounds of annual lung cancer screening were conducted with chest radiograph and sputum cytology in YTC (no participant was enrolled in 1999). Of 9,295 participants, 9,094 received at least 1 sputum screening. Detailed information on chest radiograph and sputum cytology can be seen elsewhere (22). For sputum screening, induced or 3-day pooled spontaneous sputum specimens were collected. Sputum specimens were stored at room temperature in Saccomanno solution and were smeared on glass slides and stained with Pap for cytologic examination. Slides were reviewed independently by 2 well-trained cytopathologists. Sputum cytologic screening result was classified into degrees as follows: (i) 1: negative, (ii) 2–1: slight atypical metaplasia, which was considered to result from inflammatory or irritative process, but not from cancer, (iii) 2–2: moderate atypical metaplasia, which indicated a cytologic abnormality sufficient to repeat examination and surveillance, (iv) 2–3: grave atypical metaplasia, which implied that changes were suspicious for cancer, but not diagnostic, either because the abnormal cells were too few or the abnormalities were not sufficiently well defined, (v) 3: suspicious for cancer, (vi) 4: highly suspicious for cancer, (vii) 5–1: squamous carcinoma, (viii) 5–2: adenocarcinoma, (ix) 5–3: undifferentiated carcinoma (includes both large cell and small cell), and (x) other cancers (19, 23). In this study, we used moderate or worse atypia as a primary measure for long-term lung cancer risk.

Exposure information

At the baseline interview, data were collected with a standardized questionnaire, including demographic characteristics, work history, tobacco consumption, and previous medical history. Individuals who had smoked cigarettes and/or pipes (water or long-stem pipes) regularly for 6 months or longer were defined as smokers, while those who had a smoking duration of less than 6 months were considered nonsmokers (24). Pack-years were calculated as the cumulative exposure to cigarettes, pipes (1 g pipes = 1 cigarette), and total tobacco exposure. The cumulative radon and arsenic exposure for each participant were calculated by summing across the estimated working-level months (WLM) and index of arsenic-exposure months (IAEM) for each job held at the YTC before the date of enrollment, respectively (22). Moreover, exposure information was annually updated from 1992 to 1996. Thus, the cumulative radon and arsenic exposure were calculated to the end of 1996 for participants who entered into study before 1996. In this study, occupational radon and arsenic exposure were grouped into four quartiles (Q1 to Q4) based on each individual's cumulative radon or arsenic levels, respectively. For tobacco smoking, we presumed that the smoking status since 1996 was unchanged, since the rate of smoking cessation were as low as to 5.8%.

Follow-up and case ascertainment

From 1992 to 1999, annual follow-up was conducted combined with screening. During the postscreening period after 1999, the first follow up was performed in 2005 and 2006. In 2019, an extended follow up was conducted, and the end date of this follow up was December 31, 2018. By the end of this extended follow up, 204 participants (2.2%) were lost to follow up, with a follow-up rate of 97.8%.

Lung cancer cases were confirmed via the following ways: (i) screen-detected cases with chest radiograph and/or sputum cytologic examination, (ii) interval case with negative screening results identified by hospital due to symptoms during the time from 1992 to 1999, and (iii) cases confirmed by the YTC cancer registry system, which was established in 1973 and received information of all YTC cancers from medical record system and the local hospital.

Statistical analysis

Two sets of statistical analyses were performed (**Fig. 1**). First, the analysis was based on the baseline sputum cytologic results, and person-years of follow up were calculated from the date of enrollment to the date of lung cancer diagnosis or censoring as of December 31, 2018 (whichever came first). Lung cancer incidence and incidence-rate ratios according to personal characteristics and sputum cytologic results were also calculated. The joint effects of sputum atypia and other factors including age, smoking, radon, and arsenic exposure were also analyzed.

Second, since a large proportion of sputum atypias were detected in annual screening rounds after baseline screening, we restricted the analysis in participants who received baseline (T0) and 3 subsequent annual (T1-T3) sputum screenings with the aim to further explore the magnitude and temporal change of lung cancer risk according to different sputum cytologic results in previous screening rounds, and person-years of follow up were calculated from the date of T3 to the date of lung cancer diagnosis or censoring as of December 31, 2018. The lung cancer risks associated with moderate or worse cytologic result for at least once, only in the first 2 rounds (T0 or T1), only in the last 2 rounds (T2 or T3), or in both T0-T1 and T2-T3 were compared with those with all negative results in T0-T4 rounds. The association between sputum atypia and lung cancer-risk was analyzed with time-varying covariate Cox regression model since the proportional hazards assumption was violated based on the results of the Schoenfeld residuals test. In the time-varying covariate Cox regression model, a sputum cytologic result*log of time, i.e., ln (t), was added. To control the confounding effect from the changes in various kinds of exposure during the long-term follow up, age, cumulative exposure of radon, arsenic and smoking (for current smokers), years since last exposure of radon, arsenic and smoking (for former smokers) were adjusted as time-varying covariate.

The effect of sputum cytologic result on lung cancer risk was also analyzed according to the different intervals of the follow-up period. In consideration of the increased risk of death from a cause other than lung cancer accompanied by aging, competing-risks regression analysis was conducted as a sensitivity analysis (25, 26). Finally, E value was calculated to assess the magnitude of the potential residual

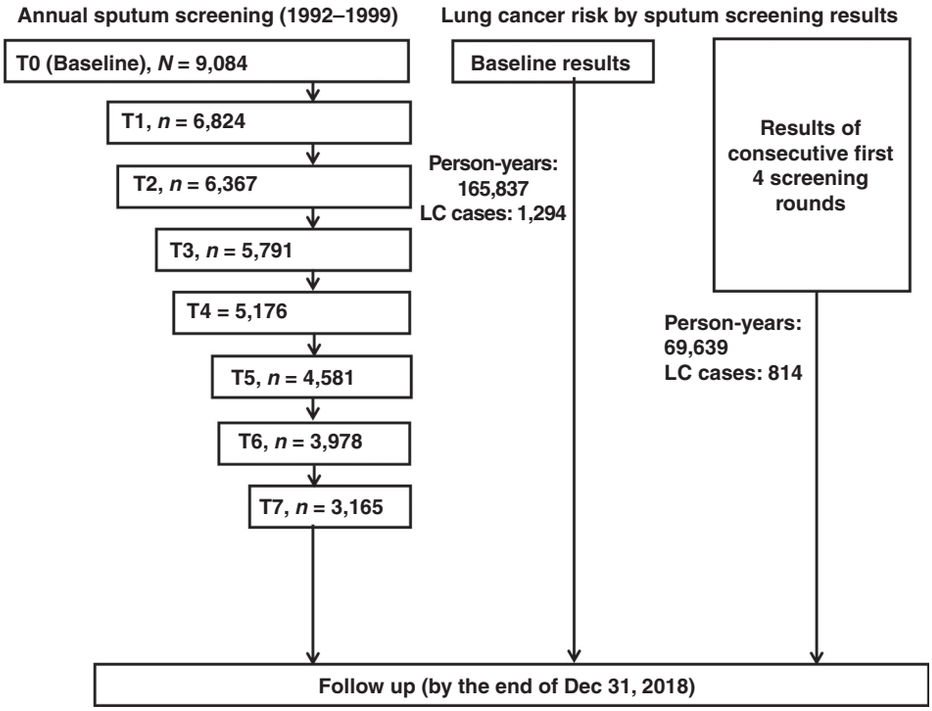

Figure 1.
Flowchart of annual sputum cytologic screening and statistical analysis.

confounding (27). E value was defined as the minimum strength of association on the risk-ratio scale that an unmeasured confounder would need to have with both the exposure and the outcome, conditional on the measured covariates, to fully explain away a specific exposure–outcome association. Statistical analysis was performed using Stata 14.0 software.

Results

The personal and occupational characteristic–specific lung cancer incidence of 9,084 participants who received at least 1 sputum cytological screening is shown in **Table 1**. Most participants were male (93.3%), and over 50% were 50 years old or older. Most of participants were smokers (84.9%), and had radon and arsenic exposure. After an extended median follow-up of 18.3 years, 1,326 lung cancer cases were confirmed. Of these lung cancers, 1,294 had a definite diagnosis date, with an overall lung cancer incidence of $780.28/10^5$. Significant higher lung cancer incidence-rate ratios were observed in male smokers, and those who had lower education levels, higher occupational radon/arsenic exposure, and previous lung disease. The incidence-rate ratio of participants who had moderate or worse atypia compared with those with negative results was 4.52 [95% confidence interval (CI): 3.42–5.88].

Table 2 showed the association of baseline sputum cytologic results and lung cancer risk. After adjusting for other potential risk factors listed in **Table 1**, moderate or worse baseline cytologic result was associated with a significantly increased lung cancer risk, with adjusted hazard ratio HR of 6.62 (95% CI: 4.66–9.42), and this relative hazard significantly decreased with time. The association between baseline moderate or worse atypia and risk of lung cancer by tumor histology was also analyzed. As shown in **Table 2**, moderate or worse atypia was associated with a moderately and continuously increased risk of SCLC (adjusted HR = 3.04, 95% CI: 1.74–5.33), but not with adenocarcinoma. The association was substantially stronger for squamous cell lung cancer (adjusted HR = 13.68, 95% CI, 8.86–21.13) and was decreased significantly with time.

The joint effects of baseline sputum atypia and other risk factors on lung cancer risk are presented in **Table 3**. Compared with participants ≤ 60 years old who had negative cytologic results, the adjusted HRs for lung cancer associated with moderate or worse atypia were 13.68 (95% CI: 7.04–26.58) and 18.62 (95% CI: 10.78–32.15) for participants ≤ 60 or > 60 years old, respectively. Both of them demonstrated a significantly decreasing trend with time. Similarly, the lung cancer risks associated with moderate or worse sputum atypia in participants with higher level smoking, radon, and arsenic exposure were a little higher than in those participants with lower exposures.

A total of 4,269 participants received the first 4 consecutive rounds of sputum screening (T0-T3). Compared with negative screening results, the adjusted HR for at least one moderate or worse cytologic atypia was 7.27 (95% CI: 4.67–11.32). Besides, the adjusted HRs for moderate or worse cytologic atypia only in T0-T1 and only in T2-T3 were 3.23 (95% CI: 1.55–6.73) and 9.48 (95% CI: 5.68–15.82) respectively, while the adjusted HR for moderate or worse cytologic atypia in both T0-T1 and T2-T3 was as high as 37.44 (95% CI: 15.00–93.47). These results suggested that more recent or consistent cytologic results were associated with higher lung cancer risk, and also showed a significant decreasing trend over time.

To directly demonstrate the temporal pattern of lung cancer risk related to sputum atypia, the adjusted HRs in different time intervals during the follow up were estimated (**Table 5**). Significantly increased lung cancer risks concerning at least one moderate or worse cytologic atypia were observed in the first 10 years since follow up, with the adjusted HRs of 3.11 (95% CI: 2.37–4.07) and 3.25 (95% CI: 2.33–4.54) for baseline and the first 4 consecutive screening rounds respectively. Besides, the association was stronger among those who had recent or consistent moderate or worse atypia in the first 4 consecutive rounds of sputum screening with adjusted HRs of 4.14 (95% CI: 2.70–6.35) and 17.55 (95% CI: 8.32–37.03) respectively.

Table 1. Lung cancer incidence among the YTC sputum-screening cohort.

Characteristic	Participants	Person-years	Cases	Incidence ($1/10^5$)	Incidence rate ratio (95% CI)
All	9,084	16,5837.1	1,294	780.28	–
Gender					
Female	553 (6.7)	12,190.0	50	410.2	Reference
Male	8,531 (93.3)	153,647.1	1,244	809.7	1.97 (1.49–2.67)
Age group					
40–49 y	3,853 (42.4)	85,102.2	291	341.9	Reference
50–59 y	2,320 (25.5)	43,816.7	392	894.6	2.62 (2.24–3.06)
60–69 y	2,341 (25.8)	31,904.6	507	1,589.1	4.65 (4.02–5.39)
>70	570 (6.3)	5,013.5	104	2,074.4	6.07 (4.80–7.61)
Education					
No	2,174 (23.9)	31,449.8	437	1,389.5	Reference
≤6 y	4,376 (48.2)	80,958.6	633	781.9	0.56 (0.50–0.64)
>6 y	2,534 (27.9)	53,428.7	224	419.3	0.30 (0.26–0.36)
Smoking status					
Never	1,370 (15.1)	28,990.6	117	403.6	Reference
Former	880 (9.7)	14,172.2	131	924.4	2.29 (1.77–2.96)
Current	6,384 (75.2)	122,674.3	1,046	852.7	2.11 (1.74–2.58)
Arsenic level					
Q1 (0–1,390.3)	2,271 (25.0)	48,726.3	167	342.7	Reference
Q2 (1,390.3–6,915.0)	2,271 (25.0)	42,734.3	312	730.1	2.13 (1.76–2.59)
Q3 (6,915.0–16,982.3)	2,271 (25.0)	33,752.9	497	1,472.5	4.30 (3.60–5.15)
Q4 (16,982.3)	2,271 (25.0)	40,623.5	318	782.8	2.28 (1.89–2.77)
Radon level					
No exposure	1,808 (19.9)	37,908.3	170	448.5	Reference
Q1 (0.1–151.7)	1,819 (20.0)	37,769.9	147	389.2	0.87 (0.69–1.09)
Q2 (151.7–284.6)	1,819 (20.0)	34,624.4	233	672.9	1.50 (1.23–1.84)
Q3 (284.6–614.4)	1,819 (20.0)	30,721.7	319	1,038.4	2.32 (1.91–2.81)
Q4 (614.4+)	1,819 (20.0)	24,812.8	425	1,712.8	3.82 (3.19–4.59)
Asthma					
No	8,417 (92.7)	156,040.3	1,160	743.4	Reference
Yes	667 (7.3)	9,796.8	134	1,367.8	1.84 (1.53–2.20)
Chronic bronchitis					
No	6,682 (73.6)	126,639.4	836	660.1	Reference
Yes	2,402 (26.4)	39,197.6	458	1,168.4	1.77 (1.58–1.99)
Silicosis					
No	8,633 (95.0)	160,198.7	1,200	749.1	Reference
Yes	451 (5.0)	5,638.4	94	1,667.1	2.23 (1.78–2.75)
Tuberculosis					
No	8,821 (97.1)	161,400.2	1,258	779.4	Reference
Yes	263 (2.9)	4,436.9	36	811.4	1.04 (0.73–1.45)
Sputum cytology					
Normal	8,069 (88.8)	150,370.8	1,068	710.2	Reference
Slight	847 (9.3)	13,630.3	167	1,225.2	1.73 (1.46–2.03)
Moderate	124 (1.4)	1,632.8	28	1,714.8	2.41 (1.60–3.50)
Worse than moderate	44 (0.5)	203.2	31	15,257.5	21.48 (14.52–30.68)
Moderate or worse	168 (1.9)	1,836.0	59	3,213.5	4.52 (3.42–5.88)

Abbreviation: y, years.

We also used competing-risks regression model to evaluate the association between sputum atypia and lung cancer risk, and the results are shown in Supplementary Table S1. No significant differences were observed between the results from the time-varying covariate Cox model (**Tables 2** and **4**) and those from the competing-risks regression model.

In this study, the E values were 5.67 (CI: 4.17) and 5.95 (CI: 4.09) for the 10-year lung cancer risk following moderate or worse sputum atypia at baseline screening or at least one moderate or worse sputum atypia in the first 4 consecutive rounds respectively. These results suggested that no significant residual confounding exist in this study.

Discussion

In this prospective study, although there was a decreasing trend, an up to 10 years increase in the risk of lung cancer associated with moderate or worse sputum atypia was observed. This association was stronger for recent and persistent atypia, and in terms of histology, for squamous carcinoma and SCLC.

Similar to our previous study and other studies, this extended follow-up study continues to confirm the association between increased lung cancer risk and sputum atypia (17–19 28, 29). Three hypotheses might explain this result. First, long-term, widespread exposure to smoking, occupational radon and arsenic, or other risk factors for lung cancer might lead to field cancerization (30). However,

Table 2. Lung cancer risk by baseline sputum cytologic screening results.

Cell type	Sputum results	Participants	Cases	Crude HR (95% CI)	Adjusted HR (95% CI)[a]	
All	Normal	8,069	1,068	Reference	Reference	Interaction with time
	Slight atypical	847	167	1.74 (1.48–2.05)	1.35 (0.92–2.00)	0.94 (0.79–1.13)
	Moderate	124	28	2.53 (1.74–3.68)	2.27 (1.20–4.30)	0.77 (0.57–1.05)
	Worse than moderate	44	31	24.77 (17.29–35.50)	22.02 (14.27–33.97)	0.50 (0.38–0.65)
	Moderate or worse	168	59	4.78 (3.67–6.21)	6.62 (4.66–9.42)	0.52 (0.43–0.63)
Squamous	Normal	8,069	290	Reference	Reference	Interaction with time
	Slight atypical	847	54	2.02 (1.51–2.70)	1.81 (1.16–2.84)	0.80 (0.64–0.99)
	Moderate	124	17	5.20 (3.18–8.48)	6.03 (3.14–11.61)	0.67 (0.47–0.95)
	Worse than moderate	44	26	40.22 (24.87–65.05)	46.20 (26.94–79.23)	0.41 (0.29–0.54)
	Moderate or worse	168	43	9.37 (6.59–13.33)	13.68 (8.86–21.13)	0.47 (0.37–0.60)
Adenocarcinoma	Normal	8,069	117	Reference	Reference	Interaction with time
	Slight atypical	847	25	2.39 (1.55–3.68)	1.39 (0.89–2.16)	–
	Moderate or worse	168	2	1.44 (0.36–5.84)	1.25 (0.62–2.53)	–
Small cell	Normal	8,069	88	Reference	Reference	Interaction with time
	Slight atypical	847	13	1.61 (0.90–2.89)	0.90 (0.61–1.31)	–
	Moderate or worse	168	6	5.47 (2.39–12.52)	3.04 (1.74–5.33)	–

[a]Age, smoking, and occupational radon and arsenic were adjusted as time-varying covariates; gender, prior lung disease, and education were also adjusted.

significantly increased lung cancer in relation to risk of baseline moderate or worse sputum atypia was also observed in participants who had relatively young age, lower exposure of smoking, radon and arsenic, which implied that field cancerization could not be fully responsible for the association between sputum atypia and increased lung cancer risk. Second, the high lung cancer risk among older individuals, those who had high levels of smoking, radon, and arsenic exposure implied a positive interaction between sputum atypia and other carcinogens. Finally, in addition to sputum atypia as an independent risk factor, exfoliated abnormal cells of lung cancer might also contribute to the stronger association between sputum atypia detected in more recent screening rounds and lung cancer risk (17).

Stepwise progression to invasive cancer was reported for both adenocarcinoma and squamous carcinoma. However, we did not find significantly increased risk of adenocarcinoma following sputum atypia. The main reason might be that abnormal cells that are exfoliated into the sputum are located more commonly in the central airways than in the periphery of the lung. Accordingly, preinvasive lesions of squamous carcinoma might be more easily to be detected by sputum cytologic examination, which resulted in a stronger association between sputum atypia and risk of squamous carcinoma compared with that for adenocarcinoma. Several studies have indicated that squamous cell carcinoma is developed in a step-wise pattern where the epithelium changes from normal to hyperplasia, metaplasia, mild, moderate, and severe dysplasia and then carcinoma *in situ* (31, 32). Besides, it is generally accepted that high-grade lesions are more likely to progress to invasive cancer than low-grade lesions (33). Similarly, in this study, 70.5% (31/44) of baseline sputum atypia worse than moderate progressed to lung cancer with a significantly higher frequency than that of moderate atypia. High-grade lesions might also

Table 3. Lung cancer risk according to baseline sputum-screening results and other exposure.

Exposure	Participants	Cases	Sputum atypia	Crude HR (95% CI)	Adjusted HR (95% CI)[a]	
Age at baseline						
≤60	5,679	606	Negative	Reference	Reference	Interaction with time
>60	2,390	485	Negative	3.49 (3.09–3.95)	4.08 (2.59–6.42)	0.80 (0.67–0.95)
≤60	52	14	Moderate or worse	3.43 (2.02–5.84)	13.68 (7.04–26.58)	0.38 (0.28–0.51)
>60	116	49	Moderate or worse	10.95 (8.06–14.87)	18.62 (10.78–32.15)	0.51 (0.40–0.66)
Cumulative smoking						
≤25	3,504	420	Negative	Reference	Reference	Interaction with time
>25	3,273	560	Negative	1.61 (1.42–1.83)	0.97 (0.68–1.38)	1.01 (0.89–1.15)
≤25	62	22	Moderate or worse	4.28 (2.73–6.72)	5.08 (2.74–9.42)	0.62 (0.46–0.83)
>25	98	39	Moderate or worse	7.12 (5.08–9.98)	6.52 (4.08–10.44)	0.51 (0.40–0.64)
Cumulative radon						
Quartile 1–2	3,368	351	Negative	Reference	Reference	Interaction with time
Quartile 3–4	2,945	577	Negative	2.50 (2.19–2.86)	1.81 (1.22–2.69)	0.91 (0.89–1.06)
Quartile 1–2	32	7	Moderate or worse	2.50 (1.12–5.61)	8.55 (4.02–18.17)	0.40 (0.25–0.64)
Quartile 3–4	129	55	Moderate or worse	9.50 (7.08–12.74)	9.53 (5.81–15.62)	0.55 (0.43–0.70)
Cumulative arsenic						
Quartile 1–2	4,132	428	Negative	Reference	Reference	Interaction with time
Quartile 3–4	3,937	663	Negative	2.02 (1.79–2.28)	2.81 (1.87–4.22)	0.83 (0.71–0.97)
Quartile 1–2	38	10	Moderate or worse	3.68 (1.96–6.90)	13.26 (6.23–28.24)	0.43 (0.30–0.61)
Quartile 3–4	130	53	Moderate or worse	8.51 (6.32–11.46)	14.17 (8.48–23.67)	0.48 (0.37–0.61)

[a]Age, smoking, and occupational radon and arsenic were adjusted as time-varying covariates; gender, prior lung disease, and education were also adjusted.

Table 4. Lung cancer risk by sputum cytologic results of first 4 consecutive screening rounds.

Sputum results	Participants	Cases	Crude HR (95% CI)	Adjusted HR (95% CI)[a]	
The first 4 consecutive screening rounds					
Normal	3,410	453	Reference	Reference	Interaction with time
At least one moderate or worse	198	62	1.70 (1.55–1.87)	7.27 (4.67–11.32)	0.80 (0.65–0.99)
M/M+ at least once in T0-T1, not in T2-T3	93	19	2.28 (1.34–3.65)	3.23 (1.55–6.73)	0.62 (0.43–0.89)
M/M+ at least once in T2-T3, not in T0-T1	92	34	4.47 (3.04–6.37)	9.48 (5.68–15.82)	0.40 (0.31–0.52)
M/M+ in both T0-T1 and T2-T3	13	9	26.45 (13.65–51.23)	37.44 (15.00–93.47)	0.34 (0.23–0.48)

Abbreviations: M, moderate; M+, worse than moderate.
[a]Age, smoking, and occupational radon and arsenic were adjusted as time-varying covariates; gender, prior lung disease, and education were also adjusted.

progress more rapidly to invasive cancer compared with low-grade lesions. Compared with moderate atypia, more high-grade atypias might progress to squamous carcinoma in a relative shorter period, meanwhile, fewer high-grade atypias would regress to normal or low-grade level. Thus, more lung cancers would appear during the first few years of follow up. On the other hand, relatively fewer lung cancers would be found over time due to the regression of some moderate atypia during the later years of follow up. In an earlier long-term follow up study, of 14 participants with severe atypia, 6 (42.9%) were found to have lung cancer during the 9-year follow up, and most were found during the first year. In contrast, of 169 participants with moderate atypia, only 18 (10.6%) developed lung cancer during the follow up, and most of these occurred during the 6 to 9 years of follow up (31). High, rapid progression rate of high-grade atypia and low, slow progression rate of moderate atypia might be the reason for the decreasing trend of the association between sputum atypia and increased lung cancer risk during the long-term follow up. However, other studies found no significant difference in progression rate between individuals with or without severe dysplasia, and progression of metaplasia to invasive cancer within a relatively short time has also been reported, which challenged the concept of step-wise progression of preinvasive bronchial lesions to invasive cancer (34, 35).

Unlike, adenocarcinoma and squamous carcinoma, SCLC is believed to arise from severely molecularly-damaged epithelium without going through recognizable preneoplastic changes (36). However, in line with our previous study (19), we still observed a persistently increased risk of SCLC concerning sputum atypia. In contrast, the results of the Colorado study demonstrated a nonsignificant reduced risk of SCLC associated with sputum atypia (17). Molecular studies might help to elucidate the true relationship between sputum atypia and the risk of SCLC.

Sputum cytology was not recommended for lung cancer screening in most counties. However, controversy still exists for its effectiveness in lung cancer screening. In Japan, sputum cytology was still recommended for lung cancer screening (37). In clinical practice, sputum cytology is also the routine examination for other lung diseases, such as chronic cough, asthma in China (38). Molecular analysis of histopathologic grading might be helpful for a better understanding of the natural history of preinvasive cancers (39). For example, progression of bronchial dysplasia was reported to be associated with specific immune alterations (40). Characterizing the immune microenvironment of bronchial dysplasia will allow for a better understanding of progressive lesions of the central airway and advance the field of precision chemoprevention and lung cancer risk stratification.

To our knowledge, this study is the longest long-term evaluation of lung cancer risk in relation to sputum atypia. It is the first to demonstrate a decreasing trend of lung cancer risk associated with sputum atypia over time. However, significantly increased lung cancer risk could still be observed for up to 10 years since follow-up. The dynamic changes of lung cancer risk following sputum atypia might contribute to a refinement of selection criteria for lung cancer screening or chemoprevention trials (21). Besides, based on the results of NLST, noncalcified nodules identified on LDCT screening were predictive of lung cancer risk up to 10 or more years following the screen (41). In theory, LDCT and sputum are complementary, since the former is more efficient to detect peripheral lesion, while the latter is relatively more efficient for central lesions. Based on the results of these two studies, it is natural that noncalcified nodule and sputum atypia might also be complementary in high-risk identification and surveillance, especially given the time-varying effect of these two kinds of precursors of lung cancer risk. Additionally, biomarkers in the sputum would allow further refinement of screening selection criteria and discrimination of indeterminate pulmonary nodules (42).

Table 5. Piece-wise lung cancer risk according to previous sputum cytologic screening results.

Sputum result	Participants	Cases	Adjusted HR[a]		
			<10 years	10–15 years	>15 years
Baseline (T0)					
Normal	8,069	1,068	Reference	–	–
Moderate or worse	168	59	3.11 (2.37–4.07)	1.00 (0.50–2.02)	1.28 (0.72–2.13)
First 4 consecutive rounds (T0-T4)					
Normal	3,410	453	Reference	–	–
At least one moderate or worse	198	62	3.25 (2.33–4.54)	1.57 (0.68–3.60)	0.91 (0.40–2.07)
M/M+ at least once in T0-T1, not in T2-T3	93	19	1.02 (0.25–4.18)	0.54 (0.13–2.28)	–
M/M+ at least once in T2-T3, not in T0-T1	92	34	4.14 (2.70–6.35)	1.65 (0.52–5.22)	1.26 (0.46–3.43)
M/M+ in both T0-T1 and T2-T3	13	9	17.55 (8.32–37.03)	6.14 (0.85–44.71)	–

Abbreviations: M, moderate; M+, worse than moderate.
[a]Age, smoking, and occupational radon and arsenic were adjusted as time-varying covariates; gender, prior lung disease and education were also adjusted.

There were some strengths to this study. Firstly, for the prospective cohort design, information about other risk factors was obtained before lung cancer occurrence. Secondly, through an extended follow up with a low rate of loss to follow up, the number of lung cancer cases doubled compared with our previous study, increasing the power of analysis. Meanwhile, this extended follow up allowed us to assess not only the magnitude but also the temporal pattern of the lung cancer risk in relation to sputum atypia. Finally, the robustness of the association between sputum atypia and lung cancer risk was confirmed by the sensitivity analysis including competing-risks regression model, joint effect analysis and the assessment of impact of residual confounding. However, limitations of this study should also be considered. The first one was that most participants were exposed to smoking, occupational radon, and arsenic exposure. Therefore, residual confounding might not have been fully eliminated. The second limitation was that the histology information is lacking for nearly half of lung cancer cases, which would decrease the statistical power when the analysis was conducted according to histology. In addition, the number of lung cancer cases in some subgroups according to other exposures was small, which also led to a reduced statistical power.

In conclusion, this study confirmed the long-term increase of lung cancer risk in individuals with moderate or worse sputum atypia. Sputum atypia might play a complementary role in identifying high-risk individuals for lung cancer screening, surveillance, or lung cancer chemoprevention. Besides, the risk decreased over time, especially for squamous cell lung cancer. Molecular sputum analysis is warranted to gain insight into the natural history of bronchial preinvasive lesions and to further quantify lung cancer risk.

Disclaimer

The sponsor had no role in the design of the study, the collection and analysis of the 301 data, or the preparation of the manuscript.

Authors' Disclosures

No disclosures were reported.

Authors' Contributions

Y. Fan: Conceptualization, formal analysis, funding acquisition, writing–original draft. **Z. Su:** Funding acquisition, investigation, methodology, writing–review and editing. **M. Wei:** Investigation, writing–review and editing. **H. Liang:** Funding acquisition, writing–review and editing. **Y. Jiang:** Project administration, writing–review and editing. **X. Li:** Funding acquisition, writing–review and editing. **Z. Meng:** Funding acquisition, writing–review and editing. **Y. Wang:** Writing–review and editing. **H. Pan:** Writing–review and editing. **J. Song:** Funding acquisition, writing–review and editing. **Y. Qiao:** Conceptualization, supervision, project administration, writing–review and editing. **Q. Zhou:** Conceptualization, supervision, writing–review and editing.

Acknowledgments

This work was supported by Cancer Foundation of China (grant number: CFC2020KYXM001, CFC2020KYXM002 to Y. Fan). This study was also partly funded from Tianjin Natural Science Foundation (grant number: 17JCYBJC25400 to Y. Fan, 18JCYBJC92100 to X. Li), NIH (grant number K01 1K01TW011190–01A1 to J. Song), Key R & D projects of Science and Technology Department of Sichuan (grant number 2020YFS0212 to H. Liang); National Natural Science Foundation of China (grant number 81971650 to Z. Meng). We gratefully acknowledge all participants who have participated in this study. We thank the staff of Office of Gejiu Municipal Leading Group for Cancer Prevention and Control, Gejiu City, Yunnan, China for their assistance in the collection of the follow-up data of the YTC screening cohort. We thank Dr. Nasra Mohamoud Ali (a native English speaker) for providing assistance in editing this manuscript.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received March 15, 2021; revised May 12, 2021; accepted August 18, 2021; published first August 26, 2021.

Reference

1. Cao M, Chen W. Epidemiology of lung cancer in China. Thorac Cancer 2019;10:3–7.
2. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2018;68:394–424.
3. National Lung Screening Trial Research, Aberle DR, Adams AM, Berg CD, Black WC, Clapp JD, et al. Reduced lung cancer mortality with low-dose computed tomographic screening. N Engl J Med 2011;365:395–409.
4. de Koning HJ, van der Aalst CM, de Jong PA, Scholten ET, Nackaerts K, Heuvelmans MA, et al. Reduced lung cancer mortality with volume CT screening in a randomized trial. N Engl J Med 2020;382:503–13.
5. Becker N, Motsch E, Trotter A, Heussel CP, Dienemann H, Schnabel PA, et al. Lung cancer mortality reduction by LDCT screening-Results from the randomized German LUSI trial. Int J Cancer 2020;146:1503–13.
6. Pastorino U, Silva M, Sestini S, Sabia F, Boeri M, Cantarutti A, et al. Prolonged lung cancer screening reduced 10-year mortality in the MILD trial: new confirmation of lung cancer screening efficacy. Ann Oncol 2019;30:1162–9.
7. Tammemagi MC, Katki HA, Hocking WG, Church TR, Caporaso N, Kvale PA, et al. Selection criteria for lung cancer screening. N Engl J Med 2013;368:728–36.
8. Katki HA, Kovalchik SA, Berg CD, Cheung LC, Chaturvedi AK. Development and validation of risk models to select ever-smokers for CT lung cancer screening. JAMA 2016;315:2300–11.
9. Walter JE, Heuvelmans MA, de Jong PA, Vliegenthart R, van Ooijen PMA, Peters RB, et al. Occurrence and lung cancer probability of new solid nodules at incidence screening with low-dose CT: analysis of data from the randomised, controlled NELSON trial. Lancet Oncol 2016;17:907–16.
10. Yousaf-Khan U, van der Aalst C, de Jong PA, Heuvelmans M, Scholten E, Walter J, et al. Risk stratification based on screening history: the NELSON lung cancer screening study. Thorax 2017;72:819–24.
11. Patz EF Jr, Greco E, Gatsonis C, Pinsky P, Kramer BS, Aberle DR. Lung cancer incidence and mortality in National Lung Screening Trial participants who underwent low-dose CT prevalence screening: a retrospective cohort analysis of a randomised, multicentre, diagnostic screening trial. Lancet Oncol 2016;17:590–9.
12. Pinsky PF, Church TR, Izmirlian G, Kramer BS. The National Lung Screening Trial: results stratified by demographics, smoking history, and lung cancer histology. Cancer 2013;119:3976–83.
13. McWilliams A, Mayo J, MacDonald S, leRiche JC, Palcic B, Szabo E, et al. Lung cancer screening: a different paradigm. Am J Respir Crit Care Med 2003;168:1167–73.
14. Melamed MR, Flehinger BJ, Zaman MB, Heelan RT, Perchick WA, Martini N. Screening for early lung cancer. results of the Memorial Sloan-Kettering study in New York. Chest 1984;86:44–53.
15. Frost JK, Ball WC Jr, Levin ML, Tockman MS, Baker RR, Carter D, et al. Early lung cancer detection: results of the initial (prevalence) radiologic and cytologic screening in the Johns Hopkins study. Am Rev Respir Dis 1984;130:549–54.
16. Doria-Rose VP, Marcus PM, Szabo E, Tockman MS, Melamed MR, Prorok PC. Randomized controlled trials of the efficacy of lung cancer screening by sputum cytology revisited: a combined mortality analysis from the Johns Hopkins Lung Project and the Memorial Sloan-Kettering Lung Study. Cancer 2009;115:5007–17.
17. Byers T, Wolf HJ, Franklin WA, Braudrick S, Merrick DT, Shroyer KR, et al. Sputum cytologic atypia predicts incident lung cancer: defining latency and histologic specificity. Cancer Epidemiol Biomarkers Prev 2008;17:158–62.
18. Prindiville SA, Byers T, Hirsch FR, Franklin WA, Miller YE, Vu KO, et al. Sputum cytological atypia as a predictor of incident lung cancer in a cohort of heavy smokers with airflow obstruction. Cancer Epidemiol Biomarkers Prev 2003;12:987–93.

19. Fan YG, Hu P, Jiang Y, Chang RS, Yao SX, Wang W, et al. Association between sputum atypia and lung cancer risk in an occupational cohort in Yunnan, China. Chest 2009;135:778–85.
20. Spiro SG, Hackshaw A, Lung SCG. Research in progress–LungSEARCH: a randomised controlled trial of surveillance for the early detection of lung cancer in a high-risk group. Thorax 2016;71:91–3.
21. Szabo E, Mao JT, Lam S, Reid ME, Keith RL. Chemoprevention of lung cancer: diagnosis and management of lung cancer, 3rd ed: American College of Chest Physicians evidence-based clinical practice guidelines. Chest 2013; 143:e40S–60S.
22. Qiao YL, Taylor PR, Yao SX, Erozan YS, Luo XC, Barrett MJ, et al. Risk factors and early detection of lung cancer in a cohort of Chinese tin miners. Ann Epidemiol 1997;7:533–41.
23. Berlin NI, Buncher CR, Fontana RS, Frost JK, Melamed MR. The National Cancer Institute Cooperative Early Lung Cancer Detection Program. Results of the initial screen (prevalence). Early lung cancer detection: Introduction. Am Rev Respir Dis 1984;130:545–9.
24. World Health Organization. Guidelines for controlling and monitoring the tobacco epidemic. World Health Organization; 1998. Available from: https://apps.who.int/iris/handle/10665/42049.
25. Thabane L, Mbuagbaw L, Zhang S, Samaan Z, Marcucci M, Ye C, et al. A tutorial on sensitivity analyses in clinical trials: the what, why, when and how. BMC Med Res Method 2013;13:92.
26. Fine JP, Gray RJ. A proportional hazards model for the subdistribution of a competing risk. J Am Stat Assoc 1999;94:496–509.
27. VanderWeele TJ, Ding P. Sensitivity analysis in observational research: introducing the E-value. Ann Intern Med 2017;167:268–74.
28. Merrick DT, Gao D, Miller YE, Keith RL, Baron AE, Feser W, et al. Persistence of bronchial dysplasia is associated with development of invasive squamous cell carcinoma. Cancer Prev Res 2016;9:96–104.
29. van Boerdonk RA, Smesseim I, Heideman DA, Coupe VM, Tio D, Grunberg K, et al. Close surveillance with long-term follow-up of subjects with preinvasive endobronchial lesions. Am J Respir Crit Care Med 2015;192: 1483–9.
30. Robles AI, Harris CC. Lung cancer field cancerization: implications for screening by low-dose computed tomography. J Natl Cancer Inst 2017;109:djw328.
31. Frost JK, Ball WC Jr, Levin ML, Tockman MS, Erozan YS, Gupta PK, et al. Sputum cytopathology: use and potential in monitoring the workplace environment by screening for biological effects of exposure. J Occup Med 1986;28: 692–703.
32. Saccomanno G, Archer VE, Auerbach O, Saunders RP, Brennan LM. Development of carcinoma of the lung as reflected in exfoliated cells. Cancer 1974;33: 256–70.
33. Jeremy George P, Banerjee AK, Read CA, O'Sullivan C, Falzon M, Pezzella F, et al. Surveillance for the detection of early lung cancer in patients with bronchial dysplasia. Thorax 2007;62:43–50.
34. Breuer RH, Pasic A, Smit EF, van Vliet E, Vonk Noordegraaf A, Risse EJ, et al. The natural course of preneoplastic lesions in bronchial epithelium. Clin Cancer Res 2005;11:537–43.
35. Bota S, Auliac JB, Paris C, Metayer J, Sesboue R, Nouvet G, et al. Follow-up of bronchial precancerous lesions and carcinoma in situ using fluorescence endoscopy. Am J Respir Crit Care Med 2001;164:1688–93.
36. Gazdar AF, Brambilla E. Preneoplasia of lung cancer. Cancer Biomark 2010;9: 385–96.
37. Hamashima C. Cancer screening guidelines and policy making: 15 years of experience in cancer screening guideline development in Japan. Jpn J Clin Oncol 2018;48:278–86.
38. Ding H, Xu X, Wen S, Yu Y, Pan J, Shi C, et al. Changing etiological frequency of chronic cough in a tertiary hospital in Shanghai, China. J Thorac Dis 2019;11: 3482–9.
39. Denisov EV, Schegoleva AA, Gervas PA, Ponomaryova AA, Tashireva LA, Boyarko VV, et al. Premalignant lesions of squamous cell carcinoma of the lung: the molecular make-up and factors affecting their progression. Lung Cancer 2019;135:21–8.
40. Beane JE, Mazzilli SA, Campbell JD, Duclos G, Krysan K, Moy C, et al. Molecular subtyping reveals immune alterations associated with progression of bronchial premalignant lesions. Nat Commun 2019;10:1856.
41. Pinsky P, Gierada DS. Long-term cancer risk associated with lung nodules observed on low-dose screening CT scans. Lung Cancer 2020;139:179–84.
42. Seijo LM, Peled N, Ajona D, Boeri M, Field JK, Sozzi G, et al. Biomarkers in lung cancer screening: achievements, promises, and challenges. J Thorac Oncol 2019; 14:343–57.