Original Article

Case–Spouse Control Design in Practice: An Experience in Estimating Smoking and Chronic Obstructive Pulmonary Disease Deaths in Chinese Adults

Jingmei Jiang,1 Boqi Liu,2* Freddy Sitas,3 Xianjia Zeng,1 Junshi Chen,4 Wei Han,1 Xiaonong Zou,2 Yanping Wu,2 Ping Zhao,2 Junyao Li2

Background/Purpose: We assessed the effect of smoking on death from chronic obstructive pulmonary disease (COPD) in China by employing a large population-based, case–spouse control study design using data from a nationwide survey of mortality.

Methods: During 1989–1991, a nationwide retrospective survey of mortality was conducted in China. For approximately 1,000,000 adults dying from all causes during 1986–1988, their surviving spouses or other informants provided detailed information about their own as well as the deceased person’s smoking history. For this study, 183,393 individuals who died of COPD at age ≥ 40 years were taken as cases, while 272,984 sex-matched surviving spouses of subjects who died from any cause were taken as controls.

Results: COPD death rates for smokers were more than twice as high as those of non-smokers, with a dose–response risk pattern, despite the fact that COPD death rates varied widely by region and age. Tobacco accounted for 41.4% of COPD deaths in men, but only 13.5% of those in women, who had a lower rate of smoking.

Conclusion: A case–spouse control study, as an alternative design, is valid and feasible in utilizing information from population-based, retrospective mortality survey data for an analytical epidemiological study of disease etiology.

Key Words: case-control studies, China, chronic obstructive pulmonary diseases, mortality, smoking

In Western countries, the main increase in cigarette consumption took place 40 years before that in China, thus the peak number of studies on tobacco-related deaths were conducted about 40 years earlier than in China. China has 20% of the world’s population and accounts for 30% of its cigarette consumption.1 It is impossible to use classical epidemiological studies to assess the

©2010 Elsevier & Formosan Medical Association

1Department of Epidemiology and Biostatistics, Institute of Basic Medical Sciences Chinese Academy of Medical Sciences, and 2The Cancer Institute, Chinese Academy of Medical Sciences, 3The Institute of Nutrition and Food Safety, Chinese Center for Disease Control and Prevention, Beijing, China; 4Cancer Epidemiology Research Unit, The Cancer Council NSW, Australia.

Received: September 4, 2008 *Correspondence to: Dr Boqi Liu, National Cancer Institute, Chinese Academy of Medical Sciences, 17 Pan Jia Yuan Nan Li, Beijing 100021, China.
Revised: June 26, 2009 E-mail: jingmeijiang@yahoo.com
Accepted: August 31, 2009
patterns of all mortality related to tobacco in a large developing population, where the majority of the people do not realize the hazards of tobacco, because these hazards are insidious. For this reason, innovative and robust epidemiological methods to assess the health risks of smoking need to be explored. Chronic obstructive pulmonary disease (COPD) is one of the most common public health problems worldwide and its major risk factor, cigarette smoking, has been well described. In China, COPD is currently the second most common cause of death. It has been estimated that the prevalence of COPD among Chinese adults is around 5.9%. Determination of the impact of smoking on COPD in Chinese adults is an important aspect of reducing mortality from this disease.

The purpose of the present study was to present an alternative control group selection for a large population-based, case–control study using data from a nationwide retrospective survey of mortality conducted in China from 1989 to 1991. We assessed the effect of smoking on deaths from COPD in Chinese adults, using spouses of subjects who died from any cause as controls, to determine the validity of this study design.

Subjects and Methods

National mortality survey and the study base
From 1989 to 1991, a nationwide, retrospective mortality survey was conducted in China, which involved 103 study areas and approximately 1,000,000 adults who died from all causes during the years 1986–1988. We defined the total population from which the mortality survey was conducted as the study base. The study base included 24 major cities, which were chosen non-randomly to represent a wide geographical spread, and 77 rural counties were also selected through stratified random sampling from the 2000 counties whose cancer rates in 1973–1975 were recorded in the Chinese Cancer Atlas. We obtained all the information on the base population from the local Population Administrative Office. In China, the administrative system consists of five strata: central government, provincial/municipal government, district/county government, street/town government, and administrative villages. At each level, there is a corresponding department charged with responsibility for collecting and managing the demographic records of individuals who reside in the area. The records include name, address, sex, date and place of birth, as well as other items of personal information. The average total population of the study base was about 69 million (72.6% in urban areas and 27.4% in rural areas). All deaths within the study base were identified primarily from the same local administrative records, which were supplemented by a review of the medical records. When medical records were unavailable, discussions were conducted (a few years after death) with local health workers, community leaders, and family members concerning the circumstances surrounding the deaths. A standard death certificate was located for 90% of all cases. The underlying cause of each death was coded according to the World Health Organization International Classification of Diseases, 9th revision (ICD-9). Further details concerning the study design, sampling methods, and field survey methods have been described elsewhere.

Recruitment of cases and controls
The current study used data from the above survey. However, we employed an alternative control selection for our large population-based, case–control study using the same data set. In this study, 183,393 individuals aged ≥ 40 years (52.7% male and 47.3% female) who died from COPD during 1986–1988 within the study base were taken as cases [COPD deaths were defined as those who died from bronchitis or chronic bronchitis (ICD-9: 490-1), emphysema (492), chronic pulmonary heart disease, or other diseases of the pulmonary circulation (416-7)], and 272,984 sex-matched surviving spouses (33.5% male and 66.5% female) of those who died from any cause during 1986–1988 were taken as controls. The age range of the controls was the same as for the cases when his/her spouse died during those
same years. The control selection procedure was based on the assumption that individuals in the control group had smoking habits in 1980 that were similar to those of the study base.

**Verification of smoking information**

To obtain information on smoking, we interviewed informants for all adult deaths from 1989 to 1991 according to the address provided in the administrative records. We first selected the surviving spouse for interview (28.1% of total interviewees in urban areas; 20.7% in rural areas). If the surviving spouses were unavailable or they were unable to be interviewed because of their physical or mental condition, another family member (35.6% in urban areas; 27.1% in rural areas), or other relatives and local informants (36.3% in urban areas; 52.2% in rural areas) were interviewed. The fieldwork involved over 500 interviewers, who usually worked in teams of two in urban areas and four in rural areas, and at least one team member was medically trained. Consistency was checked by repeat interviews (about 6% for all interviews) by team leaders in randomly selected households. A short structured questionnaire was used during interview, and was divided into two sections: one for the deceased person and one for a living person. The two sections contained the same questions with regard to smoking history. The information on smoking history was provided by living informants who described their own smoking habits as well as those of their dead partner. The smoking-related data included age at which smoking began, average number of cigarettes smoked per day, number of years smoked, and types of tobacco consumed (cigarettes, hand-rolled cigarettes, or other forms of tobacco). These data were used to determine whether people had ever smoked before 1980. To ask about smoking habits about 8 years prior to death was an attempt to avoid reverse causation, that is, measuring changes in smoking habits as a result of disease. Furthermore, a non-smoker was defined as a person who had never smoked during his/her life or had only smoked infrequently at a young age.

**Statistical methods**

The main analyses involved comparison between COPD cases and sex-matched spouse controls. We used unconditional logistic regression to obtain death rate ratios for smokers to non-smokers adjusted for age (5-year age groups) and locality (i.e. by county or city district of residence).

If a proportion of the COPD deaths involved smokers, then the fraction attributed to smoking was $P \times (1 - 1/RR)$, where $P$ was the proportion of smokers among COPD deaths, and $RR$ was the adjusted risk ratio.

To describe the distribution of COPD deaths by smoking status nationwide, two data sets—national mortality survey data and data from a nationwide case-control comparison—were pooled to calculate the area- and age-specific COPD death rates. We used the formula $P_0 = M/(1 + (RR - 1) \times P_c)$ to calculate COPD death rates for non-smokers, and $P_1 = RR \times P_0$ for smokers, where $M$ denoted the age-standardized COPD death rates for the whole of the study base, irrespective of whether the smoking habits were ascertained (standardized rate at age ≥ 40 years was defined as the mean of the nine 5-yearly rates in this age range). Death registration records were only available for about 90% of all cases in our survey, therefore, it was assumed that this percentage would also apply to COPD deaths included in the study. Thus, when calculating the death rate $M$, we made a 10% reduction in the local population estimates. $RR$ represents the adjusted risk ratio, and $P_c$ denotes the prevalence of smoking among the controls. We obtained area- and age-specific COPD death rates when the calculation was divided by region or by age group, respectively.

**Results**

There was a total of 183,393 COPD cases (urban 62.2%; rural 37.8%) and 272,984 living spouse controls (urban 76.0%; rural 24.0%) aged ≥ 40 years, with information on their smoking status. The cases and controls had similar birthplaces and socioeconomic levels. The mean (standard
deviation) age for cases and controls was 71.7 (9.4) and 63.4 (10.7) years, respectively, for men, and 74.0 (10.2) and 61.0 (9.5) years for women.

Most cases (56.4% of urban cases and 6.1% of rural cases) were diagnosed in hospitals at provincial or district levels. A total of 36.2% of urban and 66.0% of rural deaths were diagnosed in hospitals at county level; 4.9% of urban and 26.1% of rural deaths were diagnosed by a general practitioner; and an even smaller proportion (urban, 2.5%; rural, 1.8%) were diagnosed by inference by a qualified physician after death.

The prevalence of smoking among cases and controls in men was higher in rural than in urban areas, but in women, the prevalence was lower in rural than urban areas (Table). After standardization for age and area of residence, the standardized risk ratios [95% confidence interval (CI)] for COPD deaths were 2.40 (2.34–2.45) for men, and 2.13 (2.08–2.18) for women. Tobacco caused 41.4% of COPD deaths in men, but only 13.5% of those in women, as the proportion of women who smoked was small.

### Demographic characteristics of COPD mortality in smokers and non-smokers

The age-standardized COPD death rate (M) for the whole population of each geographical area was calculated. This local rate, combined with the local prevalence of smoking (in the reference group) and the smoker-to-non-smoker risk ratio was used to calculate the local COPD death rates for smokers and non-smokers. The counties have smaller populations, therefore, for statistical stability, they were combined into two groups: 33 counties in coastal provinces and 44 in inland provinces, and we also combined some cities in the same province for the same reason (Figure 1). Throughout China, even among non-smokers, the COPD death rates differed markedly between cities. For both sexes, for example, the death rate in non-smokers was 2.1% in Jilin (North-eastern city), whereas it was 14.1% in Zigong (South-western city). Thus, although the risk ratios were almost the same in both cities, the absolute excess of COPD produced by smoking in men and women was much greater in Zigong than in Jilin. The very high COPD death rate of 29.1% in female smokers in Zigong was five times that (5.6%) of female smokers in Jilin.

Although the COPD death rates in non-smokers varied markedly in different regions, rates in smokers were consistently around twice those of non-smokers ($p < 0.001$), and the results produced the striking pattern shown in Figure 1. The populations in different cities are large; therefore, the general pattern of absolute risk obtained is reliable. COPD death rates in urban China were 5.4% for male and female non-smokers, and 10.7% for male and 10.4% for female smokers. In rural areas, the COPD death rates were about 1.5 times higher than in urban areas. They were 8.2% for male and

### Table. Smoking-attributable chronic obstructive pulmonary disease deaths in Chinese adults aged ≥40 years, 1986–1988

<table>
<thead>
<tr>
<th></th>
<th>Cases</th>
<th>Controls</th>
<th>RR *</th>
<th>95% CI</th>
<th>Smoking-associated deaths (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Men</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urban</td>
<td>60,095 (69.3)</td>
<td>58,723 (50.9)</td>
<td>2.63</td>
<td>2.66–2.70</td>
<td>43.0</td>
</tr>
<tr>
<td>Rural</td>
<td>36,483 (73.6)</td>
<td>26,959 (61.8)</td>
<td>1.97</td>
<td>1.90–2.05</td>
<td>36.2</td>
</tr>
<tr>
<td>Total</td>
<td>96,578 (70.9)</td>
<td>85,682 (54.3)</td>
<td>2.40</td>
<td>2.34–2.45</td>
<td>41.4</td>
</tr>
<tr>
<td><strong>Women</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urban</td>
<td>53,996 (30.1)</td>
<td>148,693 (17.0)</td>
<td>2.11</td>
<td>2.05–2.17</td>
<td>15.8</td>
</tr>
<tr>
<td>Rural</td>
<td>32,819 (17.8)</td>
<td>38,609 (8.2)</td>
<td>2.36</td>
<td>2.24–2.50</td>
<td>10.3</td>
</tr>
<tr>
<td>Total</td>
<td>86,815 (25.5)</td>
<td>187,302 (15.1)</td>
<td>2.13</td>
<td>2.08–2.18</td>
<td>13.5</td>
</tr>
</tbody>
</table>

*Adjusted risk ratios for chronic obstructive pulmonary disease deaths by age group (5-year intervals) and area of residence. RR = Risk ratio; CI = confidence interval.
female non-smokers, and 14.5% for male and 17.1% for female smokers.

**Age characteristics of COPD mortality in smokers and non-smokers**

The age-specific COPD death rates for smokers and non-smokers were also plotted by sex and region (Figure 2). These plots showed a similar pattern in different subgroups. Although the overall COPD death rates were greater in rural than urban areas, and also in women than in men, in general, the differences in COPD death rates by smoking status were not obvious before the age of 55 years. However, these differences increased markedly with age at >55 years, which indicates a significant cumulative impact of smoking on COPD mortality in both sexes in urban and rural areas.

**Dose–response relationship between COPD mortality and smoking index**

A dose–response association between COPD and smoking index was still evident after adjusting for age and area of residence. We excluded smokers who smoked only other forms of tobacco (chiefly
Chinese long pipes or local handmade cigarettes) because these were less easily quantified. Using cutoff points of <10, 10–19, 20–29, 30–39, 40–49 and ≥50 pack–years to define the smoking index categories, the risk ratios of death as a result of smoking increased with smoking index, and this trend also showed sex and regional variability (Figure 3). Taking urban and rural areas together, the RR of COPD death for male smokers with different smoking index levels ranged from 1.33 (95% CI: 1.26–1.41) in those who smoked <10 pack–years, to 3.03 (2.91–3.16) in those who smoked ≥50 pack–years. For female smokers, the RR of death from COPD ranged from 1.20 (1.14–1.28) in those who smoked <10 pack–years to 3.73 (3.43–4.05) in those who smoked ≥50 pack–years (all trend tests, \( p < 0.001 \)).

**Discussion**

We observed a substantial association between smoking and COPD mortality in Chinese adults. This association was consistent among different ages and segments of the population. Our study involved uniquely large numbers of cases and spouse controls, which provided statistically stable associations despite the fact that COPD death rates varied widely by area and age.

In this case–spouse control design, the total population from which the mortality survey was conducted was defined as the study base. Cases were all deaths from COPD, and sex-matched living spouses of those who died from any cause were selected as controls. The merit of this design is its practicality and feasibility, which allows us to examine relationships between one form of exposure and the risk of various causes of death simultaneously on a national scale. This approach is of particular relevance to developing countries, where vital registration systems are not yet sufficiently developed to document disease trends and cannot support large-scale studies of the avoidable causes of disease. Prospective studies take several years to mature, whereas this novel retrospective method requires much less time and money.

Smoking, atmospheric pollution, and respiratory tract infection are the three main extrinsic causes of COPD. In China and other developing countries, the problem of atmospheric pollution, including environmental, domestic and occupational pollution, has not yet been resolved, and acute respiratory tract infection is not treated properly or promptly, especially in rural areas. Hence, although smoking is one of the main causes of COPD in China, the fraction attributable to smoking is lower when compared with developed countries. Our findings are broadly compatible
with the limited evidence available from other Chinese studies.\textsuperscript{16–20} The wide geographical variation in COPD death rates among non-smokers, and the size of the absolute excess among smokers show that there remain important risk factors for this disease other than that caused by smoking. One study\textsuperscript{18} from Xuanwei County, Yunnan Province, where the rates of COPD were over twice the national average, has reported that the incidence of COPD decreased markedly after completion of a program to install chimneys in household coal stoves. Another recent survey in Guangdong\textsuperscript{19} has shown that a combination of biomass fuels and cigarette smoking had a strong synergistic impact on the prevalence of COPD. Also, a study in Beijing has shown that a significant decrease in the prevalence of COPD can be achieved by comprehensive interventions in the community to increase knowledge of COPD prevention, by improving the living environment and decreasing smoking rates.\textsuperscript{20}

Similar to some other studies,\textsuperscript{21–24} our study showed that female smokers had a higher risk of death from COPD. For example, a report from Canada\textsuperscript{21} has shown that the risk of COPD in female smokers is higher than that in male smokers, which suggests that female smokers are more susceptible to COPD than their male counterparts. In China, biomass/traditional fuels (coal, oil, firewood, and straw) are widely used in rural areas, and have been linked with diseases including COPD. Generally in China, women do more housework, and a recent report\textsuperscript{23} on smoking and COPD has shown that, after adjustment for potential confounding factors (biomass/traditional fuels, educational attainment), no male to female differences in risk were found. Further studies are warranted to explore whether sex is an independent risk factor of COPD.

This study employed data from a nationwide retrospective survey of mortality in China. As with any other mortality surveys, one of the limitations is that we used the noted cause of death, and this may have underestimated COPD mortality, because the duration of COPD is usually several decades, and many patients die as a consequence of comorbidity after years of dyspnea and disability. One study\textsuperscript{25} has used the Multiple-Cause Mortality Files compiled by the National Center for Health Statistics in the United States and has found that 8.2% of all deaths had a diagnosis of COPD, but COPD was recorded as the underlying cause of death in only 43% of these deaths. This indicates that mortality due to COPD might be severely underestimated when noting only the underlying cause of death. Another unique feature of our design was that there was no significant evidence of bias from differences in the source (proxy or self) of the information on tobacco use because: (1) as mentioned previously, the prevalence of smoking in the controls was almost the same as that estimated from national studies;\textsuperscript{26,27} (2) our findings revealed a similar result to that observed from a proportional mortality analysis on the same data set;\textsuperscript{5} and (3) other studies\textsuperscript{7,9,28} using this design have obtained valid results. If there was a strong association between smoking habits in surviving spouses and those of the deceased, then the risks may have been attenuated. In our study, the kappa coefficient of agreement on smoking habits of couples was 0.065 in urban areas, and 0.160 in rural areas, which indicated a very weak association between the smoking habits of couples.

This study had several limitations. First, there was no ex-smoker information, which could have limited the interpretation. However, there are very few former smokers in China.\textsuperscript{5} It also takes a long time for a person with COPD to die, therefore, many people quit smoking upon onset of symptoms. Thus the risk of smoking might have been underestimated despite our having asked for information on smoking about 8 years before death in an attempt to avoid this bias. Second, 7.4% of urban and 27.9% of rural deaths were diagnosed by a general practitioner or by inference after dying, which could have resulted in misclassification of the cause of death, particularly in rural areas, although our design included a higher urban than rural population because we were of the difference in accuracy of death certificates. Third, COPD deaths are more common in those of a lower socioeconomic status, which results in inherent bias.
when comparing deaths from COPD with controls, although we matched controls from the same area and adjusted for the locality during the analysis to try to limit socioeconomic confounders. Fourth, the impact of passive smoking was not considered in this study, and could also have resulted in the risk of smoking being underestimated, because the control group could have been affected by passive smoking. Finally, COPD is very common in China, therefore, many people with mild or moderate COPD might not be diagnosed, and as this would also be the case for the control population, it could have had an effect on estimation of the risk of tobacco smoking. Further work should be carried out to estimate the risks of smoking in the development of COPD.

In summary, our data confirm that, in China, smoking is associated with a risk of dying from COPD. A case–spouse control study, as an alternative design, is valid and feasible in utilizing information from a population-based, retrospective mortality survey for analytical epidemiological study of the etiology of the disease.

Acknowledgments

We thank Cancer Research UK, the United Kingdom Medical Research Council, the United States National Institutes of Health, the Chinese Ministry of Health, and the Chinese Academy of Medical Sciences who supported the original survey. We are grateful to Former Minister Chen Min Zhang for his encouragement, and for the cooperation of local governments. We thank Professor Richard Peto who gave us great support for the project, and doctors, nurses and other field workers who conducted the surveys. Finally, we would also like to thank the interviewees.

References


