Familial Lung Cancer and Aggregation of Smoking Habits: A Simulation of the Effect of Shared Environmental Factors on the Familial Risk of Cancer

Justo Lorenzo Bermejo1 and Kari Hemminki1,2

1Division of Molecular Genetic Epidemiology, German Cancer Research Centre, Heidelberg, Germany and 2Department of Biosciences at Novum, Karolinska Institute, Huddinge, Sweden

Abstract

Background: Tobacco smoking is the principal cause of lung cancer. The risk of lung cancer in the offspring of lung cancer patients is about twice higher than the risk in the general population. The present study investigated the contribution of shared smoking habits to the familial clustering of lung cancer.

Methods: We estimated the relative risk of lung cancer attributable to smoking according to the extent to which smokers transmit their smoking habits to the offspring (heritability of smoking), the prevalence of smoking in the general population, and the risk of lung cancer for smokers compared with nonsmokers.

Findings: The relative risk of lung cancer for the offspring of lung cancer patients attributable to smoking was 1.19 when published data on smoking practice were modeled (i.e., assuming that the heritability of smoking was 0.5, the smoking prevalence 40%, and the odds ratio of lung cancer for smokers versus nonsmokers was 2).

Interpretation: Most familial cases of lung cancer cannot be attributed to shared smoking habits. The example of smoking can be used for other familial cancers, for which no strong environmental risk factors are usually known, to infer the primary role for heritable genes.

Introduction

Practically all cancers show a familial aggregation, which may give useful clues about the underlying etiologic factors (1). For lung cancer, tobacco smoking is the overwhelming environmental cause and it is probable that familial susceptibility to lung cancer is, at least in part, mediated by heritable genes that govern the smoking behavior and the individual susceptibility to carcinogenesis. The familial aggregation of lung cancer has been described in numerous studies, including segregation analyses (2); the estimated relative risk of lung cancer is ~1.8 for 0- to 68-year-old offspring of parents with lung cancer (3). The population-attributable fraction for a parental history of lung cancer has been around 3% in Sweden (3). The heritability of lung cancer estimated in family studies has been ~0.1, and a higher concordance for monozygotic than for dizygotic twins has been noted (3-5). Furthermore, linkage analysis has been able to map a dominant locus to chromosome 6 in lung cancer (6). Although a recent family study showed that the aggregation of lung cancer is more pronounced than the familial clustering of tobacco smoking (7), the apparently high heritability of smoking habits, ~0.5 in twin studies, raises concern about the possibility of discriminating the heritable basis of smoking-related behavior from that of cancer susceptibility (8-10). Passive smoking is another complication in the assortment of the causes for familial aggregation of cancer (11).

In the present study, we modeled the familial aggregation of lung cancer depending on the smoking prevalence, the odds ratio of lung cancer for smokers versus nonsmokers, and the heritability of the smoking behavior. The assumed parameter values were based on empirical data from Sweden on smoking prevalence (e.g., 40% of the men in Sweden smoked in 1970; ref. 12) and the heritability of smoking habits (13). The aim of the exercise was to explore the proportion of familial cases of lung cancer attributable to smoking, according to reasonable values for the assumed parameters. The relevance of common environmental factors in the familial clustering of other cancers was also discussed.

Materials and Methods

The heritability of smoking liability ($h^2$) describes the extent to which smokers transmit their smoking habits to the offspring. Details on the estimation of $h^2$ can be found, e.g., in ref. (14). Briefly, $h^2 = (x_e - x_r) / ra$, where $x_e$ and $a$ depend on the smoking prevalence in the general population, $x_r$ can be inferred from the smoking frequency in offspring of smokers, and $r$ is the coefficient of relationship (0.5 for first-degree relatives). For example, if the smoking prevalence in the general population was 40% and the smoking frequency in offspring of smokers was 49.6%, $h^2$ would be 0.5. In this study, data on the smoking prevalence in the general population and $h^2$ were used to calculate the smoking frequency in offspring of smokers.

The familial incidence ratio (IR) of lung cancer (i.e., the incidence of lung cancer among the offspring of lung cancer patients divided by the incidence in the general population) represents the familial clustering of lung cancer. The effect of smoking on the IR of lung cancer depends on the smoking frequency in offspring of smokers and nonsmokers. Under certain assumptions, e.g., random mating and similar number of descendants in smokers and nonsmokers, the calculation of the smoking frequency in offspring of nonsmokers is simple. Following our example, if the smoking prevalence in the general population was 40% and the smoking frequency in offspring of smokers was 49.6%, the smoking frequency in offspring of nonsmokers would be $0.4 (1 - 0.496) / (1 - 0.4) = 33.6%$. If individuals with two smoking parents show similar
smoking frequencies than individuals with one smoking parent (15), the assumption on random mating can be relaxed. Let us represent the smoking prevalence in the general population by \( p \) and the odds ratio of lung cancer for smokers versus nonsmokers by \( OR \). If the proportion of nonsmokers with cancer is \( f_0 \), the prevalence of lung cancer in the general population is:

\[ \text{Prevalence of lung cancer} = f_0 + \frac{p}{1-p} \times f_0 \times OR \]

Under random mating, the distribution of smoking status in parents and offspring is given by:

\[
\begin{align*}
    p & \quad (\text{parent smoker, offspring smoker}) = pt \\
    p & \quad (\text{parent smoker, offspring nonsmoker}) = p(1-t) \\
    p & \quad (\text{parent nonsmoker, offspring smoker}) = (1-p)u \\
    p & \quad (\text{parent nonsmoker, offspring nonsmoker}) = (1-p)(1-u)
\end{align*}
\]

The probability that both the parent and offspring are affected by lung cancer is:

\[ \kappa_{PO} = pt \times OR \times f_0^2 + p(1-t) \times OR \times f_0^2 + (1-p)u \times OR \times f_0^2 + (1-p)(1-u) \times f_0^2 \]

Finally, the IR of lung cancer attributable to smoking is:

\[ \text{IR} = \kappa_{PO}/\kappa^2 \]

Please note that the above calculations are independent of the prevalence of lung cancer among nonsmokers \( (f_0) \). In our example, the smoking prevalence in the general population was 40% and \( h^2 \) was 0.5. Assuming an OR of lung cancer for smokers versus nonsmokers equal to 20, the IR of lung cancer attributable to smoking would be 1.19.

### Results

Figure 1 shows the relationship between the smoking prevalence in the general population and the smoking frequency in offspring of smokers, according to smoking heritabilities from 0.1 to 0.8. If the smoking prevalence was 10%, the frequency of smoking in offspring of smokers ranged from 11.6% \((h^2 = 0.1)\) to 28.1% \((h^2 = 0.8)\). For higher population prevalences, the dependence of the smoking frequency on \( h^2 \) weakened. For example, if the population prevalence of smoking was 40%, the difference between the smoking frequency for \( h^2 = 0.8 \) versus \( h^2 = 0.1 \) was 55.3% – 41.9% = 13.4%.

Table 1 shows the offspring smoking frequencies by the heritability of smoking, assuming that 40% of the population smoked. The largest difference in smoking frequency between offspring of smokers and nonsmokers was found when \( h^2 = 0.8 \); the frequencies were 55.3% and 29.8%, respectively. Table 1 also presents the offspring smoking frequencies by the prevalence of smoking in the general population, considering that \( h^2 = 0.5 \). Under this assumption, the largest difference in smoking frequency between offspring of smokers and nonsmokers was found when 10% of the population smoked; the frequencies were 20% and 8.9%, respectively.

Figure 2 represents the familial IR attributable to smoking, depending on \( h^2 \) and the OR of lung cancer for smokers versus nonsmokers; the assumed smoking prevalence in the population was 40%. The OR had a minor effect on the IR, particularly for low heritabilities. If overestimated parameters were considered \((OR = 30; h^2 = 0.8)\), the familial risk due to smoking was limited to 1.32.

Figure 3 represents the familial IR attributable to smoking, depending on the smoking prevalence and the OR for smokers.
versus nonsmokers; $r^2$ was fixed to 0.5. The OR had little effect on the IR for high smoking prevalences. The smoking prevalence in Sweden is ~40% and, assuming an OR of 20, a reasonable estimate of the IR attributable to shared smoking habits would be around 1.19. For low smoking prevalences (≤10%), the familial IR attributable to smoking may be higher than 1.55.

**Discussion**

The current great achievements in human genomics have raised hopes of using genomic tools in epidemiology, and genetic association studies on all major diseases have become commonplace. In cancer research, such studies are carried out without much consideration about the basic tenets of such studies (16): Genetic association studies assess heritable effects, which may not be overwhelming for most cancer types (17, 18). For the common cancers, such as breast and prostate cancers, the incidence has increased some 2- to 4-fold in most Western countries in the past 50 years, during which the genetic background in the population has remained unchanged (19). Obviously, the reasons of these increases are environmental, offering a major challenge to genetic association approaches (20). Although a part of the increased incidence can be ascribed to diagnostic and screening practices, a large part of it is unknown, and, if gene-environment interactions are involved, they may be numerous and difficult to disentangle. The gene-environment approach seems more amenable to lung cancer because of the shared environmental factor, tobacco, is well defined and scientific hypothesis can be formulated on its carcinogenic mechanisms. However, there has been uncertainty about the degree to which the familial aggregation of lung cancer can be explained by heritability of cancer susceptibility compared with inheritance of smoking behavior and other factors, such as passive smoking (9). Relevant to this point, the spouse correlation for lung cancer is higher than for any other cancer (~1.23; ref. 21, 22). This question was addressed in the present paper by simulation, taking into account other factors responsible of familial lung cancer than susceptibility genes.

We showed that for common smoking prevalences in the population (30-50%), the frequency of smoking offspring mainly depended on the population prevalence of smoking and to a minor degree on $h^2$ (Fig. 1). At a constant OR for lung cancer, the familial risk was directly dependent on $h^2$ and inversely on smoking prevalence. If $h^2 = 0.5$, as estimated in twin studies, and considering typical values for Swedish men (OR = 20 and 40% smoking prevalence in 1970; ref. 12), smoking would result in a familial risk of 1.19. Even when $h^2$ was fixed to 0.8, the familial risk was smaller than 1.30. These results are in agreement with previous data suggesting that familial risk factors must be of very large effect to induce strong familial aggregation of the disease (23). It thus seems reasonable to conclude that the familial risk of lung cancer noted between offspring and parents cannot be explained by the aggregation of the smoking habit (24).

It would be instructive to speculate about the extent to what the familial sharing of environmental exposures contributes to other familial cancers, which typically show familial risks of ~2.0 (22). No common risk factors of cancer exert effects as high as that of tobacco, with the possible exception of human papilloma virus (25, 26). Also, their familial sharing is probably much lower than that for smoking. Therefore, it is not possible to devise a mechanism for them to appreciably contribute to familial risks of cancer. The contribution of common environments to the risk of breast and colorectal cancers has been assessed and found to be not measurable (1, 27). The information from the present data is that, in the absence of strong risk factors of cancer, familial risks are likely to be due to heritable genes.

**References**


Figure 3. Dependence of familial IR on OR of smoking at various smoking prevalences in the population, assuming heritability of 0.5.

Familial Lung Cancer and Aggregation of Smoking Habits: A Simulation of the Effect of Shared Environmental Factors on the Familial Risk of Cancer

Justo Lorenzo Bermejo and Kari Hemminki


Updated version  Access the most recent version of this article at: [http://cebp.aacrjournals.org/content/14/7/1738](http://cebp.aacrjournals.org/content/14/7/1738)

Cited articles  This article cites 22 articles, 3 of which you can access for free at: [http://cebp.aacrjournals.org/content/14/7/1738.full#ref-list-1](http://cebp.aacrjournals.org/content/14/7/1738.full#ref-list-1)

Citing articles  This article has been cited by 9 HighWire-hosted articles. Access the articles at: [http://cebp.aacrjournals.org/content/14/7/1738.full#related-urls](http://cebp.aacrjournals.org/content/14/7/1738.full#related-urls)

E-mail alerts  Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions  To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions  To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.