Evaluation of Risk Factors for Nasopharyngeal Carcinoma in High-Risk Nasopharyngeal Carcinoma Families in Taiwan

Xiaohong (Rose) Yang,1 Scott Diehl,2 Ruth Pfeiffer,1 Chien-Jen Chen,3 Wan-Lun Hsu,3 Mustafa Dosemeci,1 Yu-Juen Cheng,3 Brenda Sun,2 Alisa M. Goldstein,1 Allan Hildesheim,1 and The Chinese and American Genetic Epidemiology of NPC Study Team

1Division of Cancer Epidemiology and Genetics, National Cancer Institute, NIH, DHHS, Bethesda, Maryland; 2Center for Pharmacogenomics and Complex Disease Research, New Jersey Dental School, University of Medicine and Dentistry of New Jersey, Newark, New Jersey; 3Graduate Institute of Epidemiology, College of Public Health, National Taiwan University, Taipei, Taiwan; 4Westat, Inc., Rockville, Maryland

Abstract

A study of nasopharyngeal carcinoma (NPC) families with two or more affected members was conducted in Taiwan (265 families with 2,444 individuals, 502 affected and 1,942 unaffected) to determine the association between NPC and potential etiologic factors in NPC high-risk families. Similar to results from a previous case-control study in Taiwan, Guangdong salted fish consumption during childhood, exposure to wood, and betel nut consumption were all associated with elevated NPC risk using conditional logistic regression, although these associations were not as strong as in the case-control study possibly due to shared environment among family members. Risk associated with cumulative wood exposure and salted fish consumption before age 10 was stronger in families with early NPC age-onset [odds ratio (ORwood), 5.10; 95% confidence interval (95% CI), 1.50-17.34; ORfish, 3.94; 95% CI, 1.47-10.55] or three or more affected members (ORwood, 4.41; 95% CI, 1.16-12.30; ORfish, 4.27; 95% CI, 1.10-16.47). In contrast, a tendency for elevated risk was noted for betel nut use in late age-onset families (OR, 2.44; 95% CI, 1.16-5.13) and the CYP2E1 c2 allele in families with less than three affected members (OR, 2.06; 95% CI, 1.04-3.35). Risk estimates associated with these exposures were similar when the analyses were restricted to EBV-seropositive subjects. To better adjust for degree of relationship among family members and residual genetic correlations, we also calculated ORs using a variance components model. The results from the two methods were similar indicating that the risk estimates from conditional logistic regression were unbiased. (Cancer Epidemiol Biomarkers Prev 2005;14(4):900–5)

Introduction

Nasopharyngeal cancer (NPC) is a rare disease in most parts of the world where the incidence rate is generally <1 per 100,000 person-years. However, it is one of the most common tumors occurring in individuals in southern China and Southeast Asia, with the incidence of 20 to 40 per 100,000 person-years (1). Over the years, numerous studies have shown that the etiology of NPC is multifactorial, including genetic, environmental, and virological factors. Among these factors, EBV seems to be the most important. It has been well established that there is a strong association between infection of EBV and NPC (2). However, given the universality of EBV infection but the unique geographic distribution of NPC, factors other than EBV are also believed to be important determinants of the risk for NPC.

Familial clustering of NPC has been documented in Chinese populations (3, 4) and even in low-risk populations (5). Family history of NPC has been suggested to be associated with NPC risk (6). Several studies showed that NPC risk was significantly higher in first-degree relatives of cases than in the general population (6-8). These data suggested that genetic factors may also contribute to the development of NPC. Recent studies showed that genetic polymorphisms of some metabolic enzyme genes (CYP2E1 and GSTM1; refs. 9-11) and some DNA repair genes (XRCRC1 and hOGG1; ref. 12) influenced susceptibility to NPC. Certain HLA groups were also reported to be associated with increased NPC risk (13). In addition, three genetic linkage studies carried out in families at high risk of NPC from southeast China, southern China, and Taiwan have mapped major susceptibility loci to chromosomes 4, 3, and 14, respectively (14, 15).5

Epidemiologic studies have suggested that a number of environmental factors may increase the risk of sporadic NPC; these factors include cigarette smoking, occupational exposures to wood and possibly, formaldehyde, and salted fish consumption (2, 16-21). In our previous evaluations from a case-control study conducted in Taiwan, we have shown that increased risks of NPC were associated with occupational exposure to wood dust for longer than 10 years (20); high intakes of nitrosamines and nitrite from meat, fish, and preserved vegetables during childhood and weaning (21); long-term (≥25 years) cigarette smoking (18); homozygosity for the c2 allele of the CYP2E1 gene (9); and several HLA alleles and haplotypes (13). The effects of such environmental and host factors in NPC high-risk families have not been characterized to date.

In this study, we examined environmental and some host exposures in multiplex families with more than two affected people collected in Taiwan and compared the NPC risk

Received 9/15/04; revised 11/16/04; accepted 12/14/04.

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 1754 solely to indicate this fact.

Requests for reprints: Xiaohong (Rose) Yang, Genetic Epidemiology Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, 6120 Executive Boulevard, Room 7014, Bethesda, MD 20852. Phone: 301-594-7804; Fax: 301-402-4489.
E-mail: royang@mail.nih.gov
Copyright © 2005 American Association for Cancer Research.

associated with these exposures to that of sporadic NPC. A better understanding of factors involved in the etiology of NPC is essential for counseling high-risk family members and for guiding future research of familial NPC kindreds.

Materials and Methods

Study Population. This study originated from a family study initiated in 1996 by NIH and National Taiwan University. The design of the family study was previously described. Briefly, 475 potentially eligible families with two or more nonparent-offspring NPC subjects (in first-, second-, and third-degree relatives) were identified after an extensive review of records from the nationwide tumor registry in Taiwan, supplemented by information from listings obtained from 10 tertiary care hospitals throughout the Island and prospective identification at select outpatient clinics that treat NPC. In addition to recruiting all available affected family members, we also recruited up to five unaffected siblings and the parents of affected subjects. If NPC cases had adult children, we recruited their spouses and up to three unaffected offspring. When NPC cases were more distant relatives such as cousins, we recruited their shared second-degree relatives such as unaffected aunts and uncles to genetically connect the NPC cases. All subjects provided informed consent to participate, and the protocol was approved by human subjects review committees in Taiwan and the United States.

Questionnaire. A risk factor questionnaire was given to all eligible individuals, as detailed previously (9). About half \( n = 203, 42.5\% \) of the NPC cases were deceased at the time of interview and proxy interviews were conducted for these individuals with a close family member [spouses (38%), siblings (33%), and children (29%)]. Most unaffected family members (96.6%) were self-respondents. Risk factor interviews obtained information on demographic factors (age, sex, ethnicity, and education), cigarette smoking, betel nut and alcohol consumption, diet, and occupational exposures. A dietary questionnaire included questions about the consumption of Guangdong moldy salted fish, Guangdong firm salted fish, and other types of salted fish during childhood (before age 10) and as an adult (between age 10 and 30). Information on occupational history was obtained by reports of all jobs held for \( \geq 1 \) year since the age of 16. The assessment of the occupational history data was detailed previously (20). Briefly, occupational history data were obtained from each participant and reviewed by an expert industrial hygienist. The intensity of exposure to wood, formaldehyde, and organic solvents was classified on a scale of 0 (not exposed) to 9 (strong) as intensity of exposure to wood, formaldehyde, and organic solvents was classified on a scale of 0 (not exposed) to 9 (strong) as intensity index for each exposure. For each subject, this information was combined with duration data to estimate the cumulative exposure (defined as duration of exposure \( \times \) average intensity). In addition, duration of exposure was also calculated with the exclusion of exposures occurring in the 10 years preceding diagnosis (cases) or interview (controls).

Biological Specimen Collection and Testing. Each study participant was asked to consent to the collection of 30 mL blood for genotyping and EBV antibody testing. Serum was tested for antibody titers to several anti-EBV antibodies known to be associated with NPC, including viral capsid antigen IgA, EBV nuclear antigen 1 IgA, and anti-DNase antibodies, according to the methods described previously (22). Individuals positive for any of these EBV markers were considered seropositive and individuals negative to all were classified as seronegative. For CYP2E1 genotyping, DNA extracted from peripheral blood mononuclear cells was tested for three polymorphisms of CYP2E1 (RSA1, PST1, and DRA1) by PCR-based restriction fragment length polymorphism analysis described previously (9).

Data Analysis. Conditional logistic regression was used to obtain odds ratios (OR) and 95% confidence intervals (95% CI) between NPC risk and each exposure, matching on families. Several measures of substance (cigarette, alcohol, and betel nut) use and occupational exposures (wood, formaldehyde, and organic solvents) were analyzed separately [ever/never, total years of exposure, years of exposure before diagnosis (substance use), years of exposure excluding 10 years before the diagnosis/interview (occupational), age at first exposure, and cumulative exposure]. ORs obtained by using each individual variable resembled one another; thus, we selected years of consumption of substances before diagnosis and cumulative occupational exposures in final analyses. For diet, total consumption of each food group were analyzed separately. The three polymorphisms of CYP2E1 (RSA1, PST1, and DRA1) were strongly correlated and thus only RSA1 data was presented. The numbers of subjects with \( c2-c2 \) genotype were small (3 affected and 31 unaffected); thus, we combined subjects with \( c1-c2 \) and \( c2-c2 \) genotype. Because age and sex were strong determinants of NPC risk, all analyses were adjusted for age (<40, 40-55, and >55 years) and sex. Variables in the final multivariate model also included cigarette smoking, betel nut use, Guangdong salted fish consumption during childhood, and cumulative wood exposure. Additional adjustment for education, ethnicity, other salted fish consumption during childhood and cumulative formaldehyde exposure did not affect risk estimates and thus were removed from the final multivariate model. RSA genotypic data was available for only a subset (27%) of all subjects recruited earlier in the study, and including RSA in the multivariate model did not change the risk estimates for other variables. Therefore, RSA variable was also removed from the final model and presented separately.

Because the family study is still ongoing, occupational data were not available for the most recently recruited subjects (40%). We believe these variables were missing completely at random because the availability of data was determined by the time of recruitment and should not introduce bias in exposure levels. In addition, 10% of the recruited subjects responded “don’t know” to the question about the consumption of salted fish. Missing data for these variables were accounted for in the multivariate model as dummy categories. To avoid the potential bias caused by this approach, we also analyzed the multivariate model for subjects with complete data only. Risk estimates for all variables did not change measurably and thus we included the missing categories in the final model.

Because EBV is a necessary risk factor for NPC and NPC patients consistently show higher antibody titers to various EBV-associated antigens before treatment compared with controls (1), we considered all NPC cases as EBV positives. The effect of EBV seropositivity on risk of NPC associated with other exposures was assessed by analyses restricted to all cases \( n = 502 \) and EBV seropositive controls \( n = 501 \).

To reduce possible reporting bias by surrogate respondents, we also repeated all analyses restricted to self-respondents. Because the results were essentially unchanged, we combined self-respondents and surrogate respondents in the subsequent analyses. We also evaluated risk according to degree of relationship of controls to NPC cases. There were only 29 (1.5%) unaffected individuals who did not have a first-degree relative and 172 (8.9%) who were spouses of a NPC case. ORs between NPC risk and all variables calculated among first-degree relatives were quite similar to those obtained from all individuals. Therefore, we included all
subjects in the final model. All conditional logistic regression analyses were done using the PROC PHREG procedure included in the SAS (version 8.0, SAS Institute, Inc., Cary, NC) software.

Risk estimates associated with cigarette smoking, betel nut use, occupational exposures to wood and formaldehyde, Guangdong fish consumption during childhood, and RSA at risk allele were also calculated by comparing the familial cases to community controls collected from our previously conducted case-control study (9), using unconditional logistic regression (PROC LOGISTIC in SAS).

Previous work indicated that for rare diseases like NPC with a multifactorial/polygenic etiology, use of conditional logistic regression models that condition on family will yield largely unbiased estimates of environmental risk factors if residual genetic correlations among family members are modest (23). To assess if the presented estimates are unbiased, the final choice of variables was also analyzed in a variance components model proposed by Pfeiffer et al. (24), which accounts for different residual correlations among family members. Risk estimates from the variance components model for each variable were compared with those obtained from conditional logistic regression. All statistical tests were two sided.

Results

Risk factor questionnaire data was available from 265 families consisting of a total of 2,444 (502 affected and 1,942 unaffected) individuals. The distribution of the number of affected and unaffected individuals within these families is shown in Table 1. Demographic (age, sex, ethnicity, and education) and substance (cigarette smoking, betel use, and alcohol drinking) consumption data were available for most (>99%) respondents. Occupational and dietary data were available for 59% (65% cases and 57% controls) and 89% (82% cases and 91% controls) subjects, respectively. EBV testing results were available for 1,771 unaffected family members (91.1%). All NPC cases were considered EBV positive. Genotyping results for RSA1 of CYP2E1 were available on 103 affected (21%) and 553 (28%) unaffected individuals.

NPC cases were older (mean age = 51.6 years) than controls (mean age = 46.5 years) at the time of the interview and more often male (71.7%) among NPC cases compared with controls (47.2%). In addition, cases were less educated than controls; 41.9% of cases finished senior high school or higher education versus 52.7% of controls. Cases and controls were similar with regard to ethnicity, diagnosis of other types of cancer, consumption of alcohol and salted fish (moldy and firm) after age 10, and occupational exposure to formaldehyde and solvents (data not shown). In age- and sex-adjusted analysis, the measured variables were not associated with statistically significant elevations in NPC risk, although OR for heavy wood exposure (≥25) was suggestive (OR, 1.90; 95% CI, 0.88-4.13). Results of the multivariable regression analysis of multiple risk factors of NPC are shown in Table 2. The adjusted OR for cumulative wood exposure (≥25) was significant; the ORs for longer betel use (≥20 years) and frequent intake of Guangdong salted fish (≥1/wk) during childhood were suggestive but not significant (Table 2). Interestingly, cigarette smoking, especially with longer duration (≥20 years), seemed inversely related to NPC risk. The patterns observed were similar in analyses restricted to all cases and 501 controls seropositive for antibodies against EBV, except for the attenuated effect of RSA allele (Table 2).

To assess if the presented estimates are unbiased, the final multivariate model was also analyzed in a variance components model. Results from the variance components model were very similar to the OR estimates obtained from conditional models, with the changes of ORs all under 10% (range, ±0.05% to 7.6%). We therefore used conditional logistic regression in all subsequent stratified analyses.

To further determine whether the risk associated with exposure variables was modified by genetic background, we evaluated the subset of families with a higher potential for genetic risk of NPC. We did stratified analyses based on the numbers of affected NPC cases and age-onset of NPC within a family. Families with ≥3 affected cases or families with at least one affected individual diagnosed before age 40 were defined as families with high genetic risk. The risk of NPC associated with heavy wood exposure (≥25) and frequent Guangdong salted fish intake (≥1/wk) during childhood was elevated in families with high genetic risk (Table 2). The test for interaction between genetic risk (≥3 affected cases) and heavy wood exposure was statistically significant (P = 0.018). An interaction between genetic risk (≥3 affected cases) and frequent salted fish intake was not statistically significant (P = 0.23), possibly due to the small number in the exposed group. Betel nut use and RSA allele, on the other hand, seemed to have stronger effects in families with low genetic risk (Table 2). Long duration of cigarette smoking (≥20 years) showed a consistently inversely related effect in all stratified analyses (Table 2).

To examine the possible effect of shared environment among affected and unaffected family members on the risk estimates, we compared NPC risk associated with these exposures between familial cases (n = 502) and community controls (n = 327) ascertained in the previously conducted case-control study in Taiwan (9). Table 3 shows the risk estimates obtained from the comparisons of sporadic cases against community controls, familial cases against family-based controls, and familial cases against community controls, respectively. Compared with the data from the community-based case-control study, the risk of NPC associated with each of these exposures was attenuated in the family-based study. However, when comparing familial cases with the community controls, age- and sex-adjusted ORs associated with betel nut use, formaldehyde exposure, and Guangdong salted fish consumption during childhood showed stronger effects, with ORs even higher than those obtained from the community-based case-control comparisons. Risks associated with smoking, RSA c2 allele, and wood exposure did not change significantly regardless of which control group was used. To match the geographic location of the community controls selected in the Taipei area (9), we restricted the familial cases to Taipei residents and the results were not measurably changed (data not shown).

Discussion

In this study, we evaluated some potential NPC risk factors in NPC multiplex families with two or more affected individuals collected in Taiwan and compared the risk associated with

Table 1. Distribution of affected and unaffected individuals in NPC high-risk families (n = 265)

<table>
<thead>
<tr>
<th>No. NPC subjects in family*</th>
<th>Total no. families</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>69</td>
</tr>
<tr>
<td>2</td>
<td>155</td>
</tr>
<tr>
<td>3-5</td>
<td>39</td>
</tr>
<tr>
<td>&gt;5</td>
<td>2</td>
</tr>
<tr>
<td>No. unaffected in family*</td>
<td></td>
</tr>
<tr>
<td>≤5</td>
<td>107</td>
</tr>
<tr>
<td>6-10</td>
<td>93</td>
</tr>
<tr>
<td>&gt;10</td>
<td>74</td>
</tr>
</tbody>
</table>

*These values reflect number of recruited cases within a family; All families have two or more affected individuals; however, most families had two or less cases recruited in this study.
these exposures to that of sporadic NPC estimated from a previously reported case-control study also conducted in Taiwan. Overall, a similar pattern of associations was observed in the two studies, with the exception of cigarette smoking. Similar to the results obtained from the case-control study, Guangdong salted fish consumption during childhood (≥1/wk), exposure to wood (≥20 years), and occupational and dietary data were available for 95% (65% cases and 95% controls) and 89% (82% cases and 91% controls) subjects, respectively. The interaction was statistically significant between genetic risk (defined as having three or more cases within a family) and heavy wood exposure (P = 0.018), but was not significant with frequent fish consumption (P = 0.23). Interactions were tested by including an interaction term in the model.*At least one affected individual in that family had early age-onset.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Case (%)</th>
<th>Control (%)</th>
<th>OR (95% CI)</th>
<th>OR (95% CI)</th>
<th>OR (95% CI)</th>
<th>OR (95% CI)</th>
<th>OR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Duration of smoking (y)^1</td>
<td>Never 269 (58.3)</td>
<td>1,328 (68.3)</td>
<td>1 (reference)</td>
<td>1 (reference)</td>
<td>1 (reference)</td>
<td>1 (reference)</td>
<td>1 (reference)</td>
</tr>
<tr>
<td></td>
<td>≤100 100 (21.7)</td>
<td>385 (19.8)</td>
<td>0.80 (0.57-1.13)</td>
<td>0.59 (0.35-0.99)</td>
<td>0.57 (0.36-0.91)</td>
<td>0.78 (0.45-1.35)</td>
<td>0.73 (0.45-1.28)</td>
</tr>
<tr>
<td></td>
<td>≥20 92 (20.0)</td>
<td>232 (11.9)</td>
<td>0.66 (0.45-0.96)</td>
<td>0.58 (0.34-1.00)</td>
<td>0.57 (0.32-1.00)</td>
<td>0.73 (0.45-1.28)</td>
<td>0.49 (0.26-0.95)</td>
</tr>
<tr>
<td>Duration of betel nut use (y)^1</td>
<td>Never 429 (87.0)</td>
<td>1,759 (90.0)</td>
<td>1 (reference)</td>
<td>1 (reference)</td>
<td>1 (reference)</td>
<td>1 (reference)</td>
<td>1 (reference)</td>
</tr>
<tr>
<td></td>
<td>≤20 27 (5.5)</td>
<td>137 (7.0)</td>
<td>0.77 (0.46-1.28)</td>
<td>0.62 (0.30-1.25)</td>
<td>1.08 (0.56-2.06)</td>
<td>0.77 (0.35-1.70)</td>
<td>0.70 (0.36-1.35)</td>
</tr>
<tr>
<td></td>
<td>≥25 37 (7.5)</td>
<td>59 (3.0)</td>
<td>1.69 (0.97-2.92)</td>
<td>1.36 (0.65-2.89)</td>
<td>0.71 (0.27-1.84)</td>
<td>2.44 (1.16-5.13)</td>
<td>1.38 (0.50-3.83)</td>
</tr>
<tr>
<td>Cumulative wood^1</td>
<td>None 303 (93.2)</td>
<td>1,062 (94.9)</td>
<td>1 (reference)</td>
<td>1 (reference)</td>
<td>1 (reference)</td>
<td>1 (reference)</td>
<td>1 (reference)</td>
</tr>
<tr>
<td></td>
<td>≤25 7 (2.2)</td>
<td>38 (3.4)</td>
<td>0.54 (0.22-1.36)</td>
<td>0.74 (0.21-2.62)</td>
<td>0.29 (0.16-0.54)</td>
<td>0.70 (0.20-2.49)</td>
<td>0.33 (0.06-1.65)</td>
</tr>
<tr>
<td></td>
<td>≥25 13 (4.6)</td>
<td>19 (1.7)</td>
<td>2.29 (1.04-5.07)</td>
<td>2.98 (0.93-9.51)</td>
<td>5.10 (1.50-17.34)</td>
<td>1.55 (0.47-5.08)</td>
<td>4.41 (1.58-12.30)</td>
</tr>
<tr>
<td>Guangdong salted fish (age &lt;10)^1</td>
<td>Never 342 (93.4)</td>
<td>1,528 (93.4)</td>
<td>1 (reference)</td>
<td>1 (reference)</td>
<td>1 (reference)</td>
<td>1 (reference)</td>
<td>1 (reference)</td>
</tr>
<tr>
<td></td>
<td>≤1/wk 13 (3.6)</td>
<td>72 (4.4)</td>
<td>0.65 (0.31-1.35)</td>
<td>0.64 (0.23-1.74)</td>
<td>0.63 (0.26-1.57)</td>
<td>0.70 (0.26-2.31)</td>
<td>0.33 (0.06-1.65)</td>
</tr>
<tr>
<td></td>
<td>≥1/wk 13 (3.0)</td>
<td>36 (2.2)</td>
<td>1.78 (0.82-3.89)</td>
<td>1.74 (0.82-3.69)</td>
<td>1.74 (0.82-3.69)</td>
<td>1.74 (0.82-3.69)</td>
<td>1.74 (0.82-3.69)</td>
</tr>
<tr>
<td>R^SA^4</td>
<td>c1-c1 57 (55.3)</td>
<td>331 (59.9)</td>
<td>1 (reference)</td>
<td>1 (reference)</td>
<td>1 (reference)</td>
<td>1 (reference)</td>
<td>1 (reference)</td>
</tr>
<tr>
<td></td>
<td>c1/c2 46 (44.7)</td>
<td>222 (40.1)</td>
<td>1.45 (0.79-2.65)</td>
<td>0.81 (0.35-1.85)</td>
<td>1.08 (0.43-2.54)</td>
<td>1.62 (0.47-5.62)</td>
<td>0.42 (0.15-1.21)</td>
</tr>
</tbody>
</table>

^1At least one affected individual in that family had early age-onset.

^2No cases in a family was based on the total number of affected subjects, regardless of how many were recruited in that family.

^3Total years of exposure before diagnosis.

^4Occupational and dietary data were available for 95% (65% cases and 95% controls) and 89% (82% cases and 91% controls) subjects, respectively.

^5The interaction was statistically significant between genetic risk (defined as having three or more cases within a family) and heavy wood exposure (P = 0.018), but was not significant with frequent fish consumption (P = 0.23). Interactions were tested by including an interaction term in the model.

^6Combined c1-c2 and c2-c2 genotype. Genotyping results were available on 103 affected (21%) and 553 (28%) unaffected individuals. Adjusted for age, sex, cigarette smoking, betel nut consumption, wood and formaldehyde exposure, and Guangdong and other salted fish consumption during childhood.
the long duration of betel nut use, frequent intake of Guangdong salted fish during childhood, and heavy formaldehyde exposures (Table 3), indicating that the attenuation of risk estimates with familial case to family control comparisons was at least partly due to the shared environment among affected and unaffected family members.

Results from the epidemiologic studies examining the association between cigarette smoking and the risk of NPC have been inconsistent. Whereas some studies suggested cigarette smoking to be a risk factor for NPC (17, 18, 26, 27), others failed to observe this association (28-31). In our previously reported case-control study conducted in Taiwan, long duration of smoking before diagnosis was significantly associated with NPC (18). In contrast, in our family study, smoking before diagnosis seemed inversely related to NPC risk, especially for long duration (>25 years). Similar results were obtained from analyses using ever/never smoking status, smoking intensity, and pack-years of smoking (data not shown). In addition, in contrast to an earlier report that age at start of smoking was inversely associated with the risk of NPC (32), our results showed no such association (OR_{age 18-21}, 0.92; 95% CI, 0.66-1.28; OR_{age ≥22}, 1.03; 95% CI, 0.75-1.42). The inversely related effect of cigarette smoking was even stronger in familial case to population control comparisons, suggesting that familial cases smoked less compared with either control groups. Because data from about half of the familial cases were collected from proxy interviews, we restricted our analyses to self-respondents to avoid the potential bias from surrogate report. Results from self-respondent data were generally consistent with those obtained from all respondents, suggesting that this is unlikely a strong source of bias in our study. However, because some of these familial cases were not incident cases and there was a delay between time of diagnosis and data collection, cases may have been more aware of the risk factors for NPC, such as smoking, than controls. This potential bias may have, at least partly, explained why familial cases tended to report smoking less. Because this is the first effort to examine the effect of nongenetic factors on familial NPC risk, this is an interesting finding that needs to be further evaluated.

We used conditional logistic regression to account for some dependency within families. However, this approach can sometimes lead to underestimates of exposure effects if genetic correlations are ignored (24). Pfeiffer et al. proposed a variance components or random effects model to further account for common familial effects and for different genetic correlations among family members (24). Subsequently, they investigated the robustness of this model with respect to various mis-specifications of genetic random effects in simulations. Their results showed that if the disease of interest is rare and includes an unmeasured genetic component that is modest, then conditional logistic regression yields nearly unbiased estimates (23). In our study, we also analyzed the NPC family data using this variance components model and showed that the risk estimates were similar to those obtained from conditional logistic regression. Therefore, conditional logistic regression provides a reasonable analytic strategy for our study.

It should be noted that our study has several limitations. As we mentioned above, information on about half of the cases was obtained from proxy respondents, because many cases were deceased at the time of interview. To evaluate whether proxy

### Table 3. Risk estimates from case-control and family studies

<table>
<thead>
<tr>
<th>Variables</th>
<th>Case-control study (n_{case}=525, n_{control}=327), OR* (95% CI)</th>
<th>Family study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Duration of smoking (y)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Never</td>
<td>1 (reference)</td>
<td>1 (reference)</td>
</tr>
<tr>
<td>&lt;25</td>
<td>1.70 (1.09-2.74)</td>
<td>0.75 (0.53-1.07)</td>
</tr>
<tr>
<td>≥25</td>
<td>1.20 (0.80-2.00)</td>
<td>0.88 (0.63-1.22)</td>
</tr>
<tr>
<td>Duration of betel nut use (y)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Never</td>
<td>1 (reference)</td>
<td>1 (reference)</td>
</tr>
<tr>
<td>&lt;20</td>
<td>1.08 (0.57-2.07)</td>
<td>0.68 (0.42-1.10)</td>
</tr>
<tr>
<td>≥20</td>
<td>1.37 (0.61-3.10)</td>
<td>1.36 (0.81-2.27)</td>
</tr>
<tr>
<td>Cumulative wood exposure</td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>1 (reference)</td>
<td>1 (reference)</td>
</tr>
<tr>
<td>&lt;25</td>
<td>1.20 (0.58-2.50)</td>
<td>0.49 (0.20-1.19)</td>
</tr>
<tr>
<td>≥25</td>
<td>2.40 (1.20-5.10)</td>
<td>1.90 (0.88-4.13)</td>
</tr>
<tr>
<td>Cumulative formaldehyde exposure</td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>1 (reference)</td>
<td>1 (reference)</td>
</tr>
<tr>
<td>&lt;25</td>
<td>1.30 (0.70-2.40)</td>
<td>1.03 (0.60-1.76)</td>
</tr>
<tr>
<td>≥25</td>
<td>1.50 (0.88-2.70)</td>
<td>1.31 (0.87-1.97)</td>
</tr>
<tr>
<td>Guangdong salted fish consumption (age &lt;10)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Never</td>
<td>1 (reference)</td>
<td>1 (reference)</td>
</tr>
<tr>
<td>Ever</td>
<td>1.40 (0.47-4.50)</td>
<td>0.92 (0.54-1.57)</td>
</tr>
<tr>
<td>RSA&lt;sup&gt;a&lt;/sup&gt;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c1-c2</td>
<td>0.83 (0.60-1.20)</td>
<td>1.44 (0.82-2.54)</td>
</tr>
<tr>
<td>c2-c2</td>
<td>2.60 (1.20-5.74)</td>
<td>0.68 (0.16-2.86)</td>
</tr>
</tbody>
</table>

*ORs obtained from the case-control study for duration of smoking before diagnosis, cumulative wood and formaldehyde exposure, Guangdong salted fish consumption before age 10, and RSA allele were previously published (9, 18, 20, 21). OR for the duration of betel nut use before diagnosis was calculated using data collected in the same case-control study using unconditional logistic regression. ORs for smoking, betel nut use, Guangdong salted fish consumption, and RSA allele carrier were adjusted for age and sex. ORs for wood and formaldehyde exposure were adjusted for age, sex, education, and ethnicity.

<sup>a</sup> All ORs were calculated using conditional logistic regression and were adjusted for age and sex; Guangdong salted fish consumption was presented as ever/never to compare with that of the case-control study.

<sup>b</sup> All ORs were calculated using unconditional logistic regression and were adjusted for age and sex.

<sup>c</sup> To compare with the published result from the case-control study, RSA c1-c2 and c2-c2 genotype were analyzed separately; however, because only three cases in the family study had c2-c2, the risk estimate in this category from the conditional logistic regression might not be accurate.
data biassed our findings, we did separate analyses among self-
respondents. The findings of our study were unchanged when
the analyses were restricted to the self-respondents indicating
that inclusion of proxy data is unlikely to have introduced a
strong source of bias. The distributions of cases in each exposure
category obtained from self-respondents and all respondents
were quite similar, with a maximal difference of 1.7%. Differences
in ORs obtained from the two analyses were also small
(0.56% for frequent salted fish consumption during childhood,
17% for heavy wood exposure, 9% for long duration of smoking
before diagnosis, and 3.6% for long duration of betel nut use). Of
note, when compared with the ORs obtained from self-
respondents, the direction of association for the risk factors
was invariably towards the null. Another limitation of our study
is that familial cases, and/or surrogates for cases, might have
been more aware of some risk factors for NPC than controls,
which might have caused a reporting bias. Finally, because we
were dealing with family data, we had to face the inevitable
challenge of missing data. We included the missing data in the
multivariate model as dummy categories to avoid unnecessarily
discard of data. To avoid the potential bias caused by this
approach, we also analyzed the multivariable model for subjects
with complete data only. Except for the fish variable, for which
the change of OR was nearly 30%, the risk estimates for all other
variables did not change measurably (range, 0.08%). The more
noticeable change for the fish variable was possibly due to the
extremely small number in the heavy exposure group when
using the complete data.

In conclusion, the results from this study suggested that
NPC risk factors, such as salted fish consumption, wood
exposure, betel nut use, and possibly, CYP2E1 polymorphism,
might be associated with an increased risk of familial NPC.
People at putatively higher genetic risk might be at further
increased risk if they have some of these exposures. In contrast
to what has been suggested in some population-based studies,
smoking was not associated with increased NPC risk in our
family study. Overall, our results confirmed the risk of these
exposures and emphasized the need for additional research
focusing on the interplay of genetic and environmental factors
in the etiology of NPC.

Appendix 1

The Chinese and American Genetic Epidemiology of NPC Study Team includes Czau-Siung Yang and Jen-Yang Chen (Department of Microbiology, National Taiwan University), Mei-Yang Liu (National Taipei College of Nursing), Mow-Ming Hsu and Tzung-Shian Sheen (Department of Otolaryngology, Chung-Shan Medical and Dental College Hospital), Ming-Tsai Tsai (Department of Otolaryngology, China Medical College Hospital), Sen-Tien Tsai and Wen-Yuan Chao (Department of Otolaryngology, National Cheng-Kung University Hospital), Peir-Rong Chen (Department of Otolaryngology, Buddhist Tzu-Chi General Hospital), Ih-Jen Su and Hey-Chi Hsu (Department of Pathology, National Taiwan University Hospital), Chi-Long Chen (Department of Pathology, China Medical College Hospital), Ching-Kuang Tai (Gee-Tien Ear Nose and Throat Hospital), Luo-Ping Ger (Department of Medical Education and Research, Kaohsiung Veteran General Hospital), and I-How Chen and Ji-Hong Hong (Taipei Chang-Gung Memorial Hospital).

References

Evaluation of Risk Factors for Nasopharyngeal Carcinoma in High-Risk Nasopharyngeal Carcinoma Families in Taiwan

Xiaohong (Rose) Yang, Scott Diehl, Ruth Pfeiffer, et al.


Updated version
Access the most recent version of this article at:
http://cebp.aacrjournals.org/content/14/4/900

Cited articles
This article cites 28 articles, 6 of which you can access for free at:
http://cebp.aacrjournals.org/content/14/4/900.full#ref-list-1

Citing articles
This article has been cited by 6 HighWire-hosted articles. Access the articles at:
http://cebp.aacrjournals.org/content/14/4/900.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, use this link http://cebp.aacrjournals.org/content/14/4/900.
Click on "Request Permissions" which will take you to the Copyright Clearance Center’s (CCC) Rightslink site.