The evolving epidemiology of nasopharyngeal carcinoma

Ellen T. Chang\textsuperscript{1,2}, Weimin Ye\textsuperscript{3}, Yi-Xin Zeng\textsuperscript{4,5}, Hans-Olov Adami\textsuperscript{6,7}

Author affiliations:

\textsuperscript{1}Center for Health Sciences, Exponent, Inc., Menlo Park, California
\textsuperscript{2}Department of Cancer Prevention Center, Sun Yat-sen University Cancer Center, Guangzhou, China
\textsuperscript{3}Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden
\textsuperscript{4}State Key Laboratory of Oncology in South China, Collaborative Innovation Center for Cancer Medicine, and Guangdong Key Laboratory of Nasopharyngeal Carcinoma Diagnosis and Therapy, Sun Yat-sen University Cancer Center, Guangzhou, China
\textsuperscript{5}Beijing Hospital, Beijing, China
\textsuperscript{6}Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA, USA
\textsuperscript{7}Clinical Effectiveness Group, Institute of Health and Society, University of Oslo, Oslo, Norway

Running title: Epidemiology of nasopharyngeal carcinoma

Keywords: nasopharyngeal cancer; epidemiology; risk factors; environment; Epstein-Barr virus; genetics

Corresponding author: Ellen T. Chang, Sc.D., Center for Health Sciences, Exponent, Inc., 149 Commonwealth Dr, Menlo Park, CA 94025 USA; echang@exponent.com

Financial support: Funding support for this work was provided by the Sun Yat-sen University Cancer Center; and by the Foundation for Chemistry Research and Initiatives (FCRI), a non-profit organization established by the American Chemistry Council, as an honorarium to Exponent. The FCRI was not involved in the conception, writing, or submission of this article, and did not review any portion of the manuscript prior to journal submission or publication.

Conflict of interest: Ellen Chang is an employee of Exponent, an international science and engineering consulting company that provides consulting support to clients in industry,
government, academia, and the private sector on some of the environmental exposures discussed in this review. The other authors declare no financial conflicts of interest.

**Word count:** 5,939

**Tables:** 2

**Figures:** 3
Abstract

Background: The epidemiology of nasopharyngeal carcinoma (NPC) has long been a source of fascination due to the malignancy’s striking geographic distribution, the involvement of the oncogenic Epstein-Barr virus (EBV), the unique association with intake of Chinese-style salt-preserved fish, and etiologic heterogeneity by histological subtype.

Methods: This review summarizes the current epidemiological literature on NPC, highlighting recent results from our population-based case-control study in southern China.

Results: Findings from our case-control study provide new insight into the epidemiology of NPC, including a diminished role of Chinese-style salt-preserved fish, a profound impact of EBV genetic sequence variation, modest positive associations with passive smoking and household air pollution, and possible effects of oral health and the oral microbiome. Recent findings from other studies include a protective association with infectious mononucleosis, suggesting a causal role of early EBV infection; familial risk conferred by shared genetic variation in the host antibody-mediated immune response to EBV infection; and an unclear association with occupational exposure to formaldehyde.

Conclusions: To shed further light on the interplay of environmental, genetic, and viral causes of NPC, large pooled studies must accumulate sufficient cases with detailed exposure data.

Impact: New epidemiological findings have reshaped the causal model for NPC.
Introduction

For decades, nasopharyngeal carcinoma (NPC) has been endemic in indigenous populations in East and Southeast Asia, the Arctic, North Africa, and the Middle East [1]. The persistence of this malignancy in certain geographic regions suggests that genetic and/or stable environmental risk factors contribute substantially to its development. The involvement of the Epstein-Barr virus (EBV) in NPC pathogenesis adds another layer of complexity to its etiology. Although our understanding of the causes and prevention of NPC remains incomplete, important epidemiological advances have recently been made. Between 2010 and 2014, we and our colleagues in China, Sweden, and the U.S. conducted a large, rigorously designed population-based case-control study of NPC in Guangdong Province and Guangxi Autonomous Region in southern China [2]. New and emerging results from this study and other recent investigations have helped to shed light on the genetic, environmental, and infectious causes of NPC, and how the role of some risk factors may have shifted over time.

Descriptive epidemiology

NPC is a rare malignancy throughout most of the world, with age-standardized rates generally below 1 per 100,000 person-years. Substantially higher rates, however, have long been observed in the Cantonese population of southern China, and intermediate rates are found in indigenous populations in Southeast Asia, the Arctic region, North Africa, and the Middle East (Figures 1 and 2) [3]. Based on Chinese cancer registry data for 2008–2012, age-standardized incidence rates of NPC varied by up to 50-fold, from 0.5 per 100,000 among males in northern China to 25 per 100,000 among males in southern China (Figure 3) [4]. In 2018, an estimated 129,079 incident cases of nasopharyngeal cancer were diagnosed internationally (85% in Asia) and an estimated 72,987 deaths occurred from nasopharyngeal cancer, making it the 23rd most common incident cancer in the world and the 21st most common cause of cancer death worldwide [3]. In Southeast Asia, nasopharyngeal cancer ranked 9th among incident cancers and 8th among cancer deaths [3].

NPC is classified histologically as keratinizing squamous cell carcinoma (type I); differentiated nonkeratinizing carcinoma (type II); undifferentiated nonkeratinizing carcinoma (type III); or basaloïd squamous cell carcinoma, a rare subtype [5]. In high-incidence areas, nearly all NPC is the undifferentiated nonkeratinizing subtype, and most of the remainder is the differentiated nonkeratinizing subtype; in low-incidence areas, a sizeable proportion of NPC is the keratinizing squamous cell subtype [6].

NPC incidence among males is double or triple that among females in most populations (Figures 1 and 2) [4]. In low-risk populations, NPC incidence reaches a modest peak in young adulthood (approximately ages 15–24 years), plateaus or declines slightly until ages 35–39 years, and then rises to a second, higher peak at approximately ages 65–79 years [4]. In high-risk populations, by contrast, NPC incidence exhibits a single peak at approximately ages 45–59 years, followed by a plateau or a modest decline.
Within geographic regions, NPC incidence can vary substantially by race and ethnicity. In the high-risk southeastern Chinese province of Guangdong, the Cantonese-speaking population—especially the traditionally boat-dwelling Tanka ethnic group—historically had double the risk of NPC compared with other dialect groups such as the Hakka, Hokkien, and Chiu Chau [7]. In Southeast Asia, NPC risk appears to be correlated with the extent of racial and social admixture with southern Chinese, particularly the ancestors of the Tanka [8]. Past and recent data from Singapore show a gradient in NPC incidence by degree of intermingling with southern Chinese: rates are lowest among Indians, who have the least admixture with the Singapore Chinese population; intermediate among Malays, who have a history of intermarriage with Chinese; and highest among the Chinese [4]. Consistent with the ethnic variation within China, NPC incidence remains low in Japan, where interaction with China has predominantly involved northern populations. In the U.S., rates are highest among Chinese, followed by Filipinos, Blacks, Koreans, non-Hispanic whites, Japanese, and Hispanic whites [4].

Migration patterns and incidence trends over time also offer hints regarding the interplay of genetic and environmental risk factors in NPC development. When high- or intermediate-risk populations migrate to lower-incidence regions, their incidence of NPC declines, but remains higher than that of other races [9-12], suggesting an important role of genetic and/or early-life (perhaps viral) exposures. Conversely, risk of NPC has been shown to rise following migration of low-risk racial/ethnic groups to higher-incidence geographic areas [13, 14], indicating an etiologic contribution of environmental factors.

NPC incidence in southern China has declined steadily since the 1970s [15], in Taiwan since the 1980s [16], and in Singapore Chinese since the 1990s [17]. The temporal pattern of declines in NPC incidence may be related to the relative timing of rapid economic development, which occurred first in Hong Kong, followed by Taiwan and then Singapore. In Southeast China, where industrialization has taken place more recently, NPC incidence had not decreased as of the early 21st century [18, 19]. Stratification by histological subtype of NPC reveals that the recent decline in Hong Kong was limited primarily to type I squamous cell NPC, whereas the incidence of nonkeratinizing types II and III NPC remained relatively stable [20]. In the U.S., although overall NPC incidence has been falling since the 1990s [21], the incidence of type II differentiated nonkeratinizing NPC has been increasing in women and men across major racial groups [22].

**Environmental risk factors**

**Infections**

*Epstein-Barr virus*

EBV, formally designated as human herpesvirus 4, is a ubiquitous B-lymphotropic virus that is carried latently by almost all humans. Primary infection typically occurs in childhood, when EBV usually causes no or mild symptoms [23]. EBV is classified by the International Agency for Research on Cancer (IARC) as an established cause of several malignancies in humans, namely,
NPC, Burkitt lymphoma, immunosuppression-related non-Hodgkin lymphoma, extranodal natural killer/T-cell lymphoma (nasal type), and Hodgkin lymphoma [24], and EBV also appears to cause a subset of gastric cancers [25]. In NPC-endemic areas, cultural norms favor relatively early infection with EBV in infants and young children [26], suggesting that younger age at EBV infection may contribute to NPC development. Given the highly dissimilar international distributions of EBV, which infects nearly all people, and NPC, which is rare throughout most of the world, co-factors must mediate the effect of EBV on NPC.

Age at primary infection with EBV, for instance, appears to affect NPC risk. Two U.S. studies found that a history of infectious mononucleosis was associated with a statistically nonsignificant 60% decrease in risk of NPC [6, 27], and a registry-based study in Sweden showed that number of older siblings (an indicator of earlier age at infection with common childhood pathogens) was positively associated with NPC risk and inversely associated with infectious mononucleosis risk, supporting the early-infection model for NPC [28]. However, whether infectious mononucleosis, which results from delayed primary infection with EBV in adolescence or young adulthood, is associated with NPC risk in high-incidence areas is unknown, perhaps because late infection with EBV is rare in such populations.

EBV sequence variants may influence NPC risk by modulating the virus’s oncogenic potential. A 30-base-pair deletion and the loss of a XhoI site in EBV’s oncogenic latent membrane protein 1 have been associated with an increased risk of NPC [29]; otherwise, few findings have been replicated. In a large-scale whole-genome sequencing study of 270 EBV isolates from EBV-associated cancer patients and controls, combined with a validation analysis in a separate set of NPC cases and controls from our population-based study in southern China, we identified two non-synonymous EBV variants in the viral BALF2 gene that were strongly and highly significantly associated with an increased risk of NPC (OR = 8.69, p = 9.69 \times 10^{-25} for single nucleotide polymorphism [SNP] 162476_C, detected in 94% of cases and 65% of controls; OR = 6.14, p = 2.40 \times 10^{-32} for SNP 163364_T, detected in 85% of cases and 46% of controls) [30]. In light of the high prevalence among NPC cases, the population attributable risk of NPC due to these two variants, if causal, was 83%. A phylogenetic analysis revealed that these two EBV variants most likely originated and expanded clonally in southern China. High-risk EBV variants could potentially be used as a basis for future NPC screening, for example, being assayed in saliva [31].

Genetic variation in the host immune response to EBV infection is also anticipated to affect EBV-related NPC risk. The latter hypothesis is supported by the finding of greater EBV IgA seroreactivity—a putative indicator of chronic active EBV replication—in first-degree relatives of NPC cases, especially those carrying genetic variants conferring NPC susceptibility [32], than in general community members [33-35]. NPC cases, compared with controls, were also found to have a deficient CD8+ T-cell response specific to EBV nuclear antigen-1 (EBNA-1) epitopes presented by HLA class I alleles [36].
IgA antibodies against the EBV viral capsid antigen (VCA) and early antigen (EA), as well as neutralizing antibodies against EBV-specific DNase, are abnormally elevated several years before NPC onset [37, 38]. Accordingly, these antibodies have long served as the basis for large-scale NPC screening programs in high-risk populations [37, 39], although a mortality benefit has not been proven in a randomized controlled trial. Anti-EBV IgA antibodies, as well as circulating cell-free EBV DNA in plasma or serum, are also correlated with NPC tumor burden, remission, and recurrence, making them useful markers for early NPC detection and disease management [40-42]. Among NPC multiplex families, but not in a general population setting, EBV-neutralizing antibodies and antibodies against the EBV glycoprotein 350, which enables EBV to enter B cells, were found to be lower in NPC cases than controls, suggesting a possible protective effect [43, 44].

Other infections

Several infectious agents other than EBV, including high-risk human papillomavirus [45, 46] and hepatitis B and C viruses [47, 48], have been studied with respect to NPC, but no causal links are established. Incidence of both type I squamous cell and nonkeratinizing types II and III NPC is approximately doubled among individuals with HIV/AIDS, as are other EBV-related malignancies [49]. NPC risk appears not to be increased among immunosuppressed solid organ transplant recipients [50, 51], perhaps mirroring the higher risk of EBV-related lymphomas in HIV/AIDS than after solid organ transplantation [52], although studies in NPC-endemic regions are few.

Tobacco and other smoke

IARC classifies tobacco smoking as an established cause of NPC in humans [53]. Although tobacco smoking is more strongly associated with type I squamous cell than types II and III nonkeratinizing NPC, several studies in high-incidence regions, where the vast majority of NPC is nonkeratinizing, also demonstrate a modestly increased risk of NPC associated with tobacco smoking, with two meta-analyses of more than 20 studies showing an approximately 60% greater risk among ever than never smokers [54, 55]. Smoking may increase NPC risk in part through reactivation of latent EBV infection, as suggested by studies that reported positive associations between cigarette smoking and elevated levels of anti-EBV immunoglobulin A (IgA) antibodies among subjects without NPC [56-58].

In line with prior findings, we found in our population-based case-control study in southern China that NPC risk was significantly increased among men who currently smoked (odds ratio [OR] = 1.34; 95% confidence interval [CI]: 1.15–1.57), but not those who quit smoking at least four years earlier (OR = 0.92; 95% CI: 0.72–1.18), compared with never smokers [59]. Risk increased with intensity, duration, and cumulative pack-years, and decreased with later age at smoking initiation. Fewer studies have evaluated secondhand smoke as a risk factor for NPC, with inconsistent but mostly positive results [53-55]. We observed a 24–30% elevation of NPC...
risk in association with passive smoking during childhood or spousal passive smoking during adulthood [59].

Betel nut chewing is classified by IARC as an established cause of oral and esophageal cancers in humans [53]. Three case-control studies of NPC, however, found no association [60-62], whereas a positive association was reported for familial NPC after age 40 years in Taiwan [60]. Smokeless tobacco use was not associated with NPC in two studies in North Africa [63, 64], one of which found a positive association with cannabis smoking [63].

IARC classifies indoor emissions from household combustion of coal as a cause of lung cancer, but whether household air pollution causes NPC is not yet clear. Epidemiological findings are inconsistent, with up to a five-fold excess risk of NPC associated with domestic wood fire exposure in three studies in China [65-67] and one in Nigeria [68]. Case-control studies in North Africa also reported increased NPC risk in association with poor household ventilation [69] and use of a traditional compact charcoal oven for cooking in childhood, but not adulthood [63]. Other studies in southern China, however, found no association of NPC risk with exposure to indoor combustion, and studies examining burning incense or antimosquito coils have also yielded mixed results [61, 67, 70-75].

In our population-based case-control study in southern China, results generally supported an adverse impact of household air pollution on NPC risk. We observed positive associations of NPC risk with use of wood, coal, or kerosene as cooking fuel (OR for wood = 1.34; 95% CI = 1.03–1.75; OR for coal = 1.70; 95% CI = 1.17–2.47; and OR for kerosene = 3.58; 95% CI = 1.75–7.36, comparing ever vs. never use); having less ventilation in the kitchen, hallway, or bedroom (OR for kitchen = 1.67; 95% CI = 1.34–2.08; OR for hallway = 1.89; 95% CI = 1.55–2.31; OR for bedroom = 3.08; 95% CI = 2.46–3.86, comparing smaller vs. larger windows); and residential exposure to cooking smoke or burning incense (OR for smoke = 1.53; 95% CI = 1.20–1.94; OR for incense = 1.59; 95% CI = 1.31–1.95, comparing highest vs. lowest category of exposure), whereas burning antimosquito coils was associated with lower NPC risk [76].

Two studies that evaluated outdoor air pollution in Taiwan reported positive associations between ambient air levels of fine or coarse particulate matter and NPC risk [77, 78], whereas a study of 10 Chinese cities found no significant association with fine and coarse particulate matter combined [79]. In both settings, ambient air levels of nitrogen dioxide were associated with greater NPC risk [77, 79].

**Diet**

*Salt-preserved fish and other foods*

Dietary consumption of Chinese-style salted fish is classified by IARC as a known cause of NPC [53], and consumption of traditional Asian pickled vegetables is classified by IARC as “possibly carcinogenic to humans,” based on “limited evidence” from epidemiological studies of NPC, stomach cancer, and esophageal cancer [80]. Traditional consumption of salt-preserved foods in
southern China, Southeast Asia, North Africa, the Middle East, and the Arctic may explain part of the distinctive international distribution of NPC. In past studies, the relative risk (RR) of NPC associated with weekly consumption, compared with no or rare consumption, of Chinese-style salt-preserved fish generally ranged from 1.1 to 4, and the RR for daily consumption ranged from 1.8 to 20 [53, 80]. The carcinogenic mechanism of Chinese-style salt-preserved fish may be mediated through dietary exposure to certain N-nitrosamines, bacterial mutagens, direct genotoxins, and/or EBV-reactivating substances [53, 80].

Studies in NPC-endemic regions have also found positive associations, albeit less consistently, with consumption of other preserved food items, including meats, eggs, fruits, and vegetables [38, 58, 62, 69, 81-83]. A meta-analysis of six studies in China, Southeast Asia, and North Africa found that relatively high versus low intake of various preserved vegetables was associated with a doubling of NPC risk [84]. Preserved food consumption may also increase NPC risk in low-incidence regions, including northern China and North America [85, 86].

More recently, however, we found that in southern China in 2010–2014, consumption of Chinese-style salted fish in adulthood, including either hard salted fish (which is prepared by directly salting and then drying) or soft salted fish (which is salted and dried after initially undergoing natural decomposition), was not associated with increased NPC risk [87]. Although intake of hard salted fish during adolescence and overall intake of salted fish during childhood were associated with significantly elevated risk of NPC, the strongest ORs were more modest than in prior studies, on the order of 1.19 (95% CI = 1.03–1.39) to 1.56 (95% CI = 1.24–1.97 for weekly vs. none), respectively. We also found ORs of approximately 1.2 to 1.5 in association with the highest category of consumption of salted eggs in adulthood and adolescence, and pickled vegetables at all three time periods, whereas fermented black beans or bean paste (but not bean curds) exhibited an inverse association with NPC risk. The weak associations in our study could be due in part to lower absolute intake of Chinese-style salted fish in the general population of southern China in recent years [87], leading to a narrower range of exposure and a smaller attributable risk. Residual confounding by socioeconomic status—for instance, correlations of higher socioeconomic status with greater intake of soft salted fish (which is typically more expensive than hard salted fish) and lower risk of NPC—also could have skewed our results. Nevertheless, our equivocal findings for Chinese-style salted fish and other preserved foods suggest that dietary factors may play a lesser role in NPC development than in the past.

**Fresh fruits and vegetables**

A protective effect of fresh fruits and/or vegetables on NPC is suggested by several case-control studies [81-83, 88-90] and a meta-analysis [84]. Results are most likely confounded to some extent by other environmental factors, however, and the impact of any particular food item is difficult to identify. Potential antioxidant, anti-inflammatory, and/or anti-nitrosation effects of fruits and vegetables are suggested by inverse associations of NPC risk with dietary intake of carrots and other red/orange/yellow vegetables [81, 85, 89, 90], green leafy vegetables [65, 89],
citrus fruit, oranges, or tangerines [81, 82, 86, 89, 91], vitamin E [92], vitamin C [86], and carotenoids [93].

**Dietary patterns**

In studies that classified dietary patterns, rather than individual food items, diets with high inflammatory potential (broadly, high in saturated fat and carbohydrates, and low in polyunsaturated fatty acids and flavonoids) were associated with greater NPC risk [94-96], as were diets high in animal products and starches [97]. Conversely, adherence to a healthy diet, evaluated in accordance with international guidelines, was associated with lower NPC risk in southern China [98], while a diet rich in fruits, vegetables, milk, fresh fish, eggs, and tea was inversely associated with NPC risk in Taiwan [99]. Conflicting results were observed for a traditional Mediterranean diet [98, 100].

In our own population-based case-control study, a predominantly plant-based diet in adulthood was associated with a 50% reduction in NPC risk (OR = 0.48; 95% CI: 0.38–0.59 for highest vs. lowest quartile), whereas a diet heavy in animal products was associated with a doubling of risk (OR = 2.26; 95% CI: 1.85–2.77); results were similar but attenuated for dietary patterns in adolescence [101]. Although residual confounding remains possible, these findings may point to a shift away from traditional southern Chinese foods and toward more Western foods as influences on NPC risk in the modern era.

**Traditional herbal medicines**

Use of Asian traditional herbal medicines was associated with a 2- to 4-fold excess risk of NPC in some case-control studies [61, 65, 70, 102], but not others [72-74, 88]. Two studies, including ours, found inverse associations of NPC risk with consumption of Cantonese-style slow-cooked herbal soup [83, 103], including nine specific plant species commonly used in such soups [103]; however, these studies found conflicting inverse and null results, respectively, for consumption of Cantonese-style herbal teas. Associations with use of traditional herbal medicines are susceptible to confounding by indication and, in case-control studies, reverse causation. Confounding by other aspects of a traditional lifestyle, such as diet and use of solid fuels for cooking and heating, is also a concern. Nevertheless, traditional herbal plants may plausibly contribute to NPC development by promoting or inhibiting EBV reactivation, as demonstrated in vitro [104].

**Alcohol and tea**

The reported association between alcohol intake and NPC risk is inconsistent, with statistically null results in several studies, especially in high-incidence areas, but positive associations in other studies in high- and low-incidence populations [105, 106]. Two meta-analyses found a positive meta-RR for the highest versus the lowest category of alcohol intake in combined case-control studies, with a J-shaped exposure-response curve that, if real, may contribute to the inconsistency of results across study populations with different alcohol drinking habits [105,
The lack of association between alcohol and NPC risk in two prospective cohort studies in Singapore [107] and Shanghai [108], however, suggests that bias may contribute to the positive findings in case-control studies. We, too, observed no apparent association between alcohol intake and NPC risk in southern China [109].

Three case-control studies found an inverse association between tea consumption and NPC risk; one study reported this result only for green tea [110], another reported inverse associations with all six specific types of tea evaluated [111], and a third did not distinguish among types of tea [58]. In our study, current intake of tea, including black or pu’er and green or jasmine tea, but not oolong tea, was associated with significantly lower risk of NPC [109]. The lack of a monotonic inverse exposure-response trend, however, may indicate a non-causal explanation.

**Medical conditions**

**Oral and respiratory tract conditions**

Most studies investigating prior chronic ear, nose, throat, and lower respiratory tract conditions found positive associations with NPC risk, with RRs usually around 2.0 or higher [73, 74, 112-119]. This apparent association could be explained by a promoting effect of chronic inflammation in the nasopharyngeal mucosa, or perhaps by the formation of carcinogenic N-nitroso compounds from nitrites generated from nitrates by certain bacteria. Alternatively, the association between chronic respiratory tract infections and NPC risk could be explained by reverse causation or various types of bias, including recall bias. The latter interpretation is corroborated by our own study, where positive associations between NPC risk and a history of chronic ear, nose, and throat diseases and use of related medications mostly disappeared after exclusion of the five years prior to diagnosis or interview [120]; and by a population-based registry linkage study in Taiwan, where chronic rhinosinusitis was associated with greater incidence of NPC within one year, but not after longer follow-up [121].

In our population-based case-control study in southern China, we found positive associations between some indicators of poor oral health and NPC risk (e.g., for having > 3 filled teeth versus none, OR =1.55; 95% CI: 1.13–2.12) and an inverse association with brushing teeth at least twice per day versus less (OR = 0.62; 95% CI 0.55–0.70) [122]. Using rRNA sequencing, we also found significantly lower oral microbial diversity in NPC cases than controls, as well as a pair of closely related co-excluding *Granulicatella adiacens* sequence variants that were associated with both NPC risk and microbial community diversity [123]. These findings suggest that subspecies niche specialization in the oral microbiome may affect NPC risk.

Increased NPC risk was observed in national registry-based studies of patients with diabetes mellitus [124] and women with gestational diabetes mellitus in Taiwan [125]. In addition, a hospital-based case-control study in Italy reported that differentiated NPC, but not undifferentiated NPC, exhibited a significant positive association with metabolic syndrome, and nonsignificant positive associations with type 2 diabetes mellitus, hypercholesterolemia, and obesity [126]. Another registry-based study in Taiwan, however, reported a significant inverse
association between diabetes mellitus and NPC risk [127], and two other investigations of diabetics in Taiwan showed that treatment with metformin was associated with lower risk of subsequent NPC [128, 129].

**Occupation**

IARC classifies formaldehyde as an established cause of NPC and leukemia [130], and wood dust [131] as an established cause of NPC and cancer of the nasal cavity and paranasal sinuses, based on studies of occupational exposures in worker populations. The European Chemicals Agency (ECHA), by contrast, does not classify formaldehyde as a known human carcinogen [132], but rather as a substance that is “presumed to have carcinogenic potential for humans,” largely based on animal evidence.

**Formaldehyde**

Most occupational epidemiological studies of formaldehyde exposure and NPC risk have been conducted in low-incidence regions, where findings may be driven by type I squamous cell carcinoma—a possibility supported by the development of squamous cell carcinomas of the nasal cavity in rodents chronically exposed to high levels of formaldehyde vapors [133, 134]. Few studies have been conducted in East and Southeast Asia, with mixed null [135, 136] and positive findings [61, 137]. A 2010 meta-analysis of occupational formaldehyde exposure and NPC risk found no significant adjusted association based on six case-control studies, and no association based on seven cohort studies after excluding a single plant from a U.S. study of 10 industrial plants (combined standardized mortality ratio = 0.72; 95% CI: 0.40–1.28) [138]. Epidemiological studies post-dating the most recent IARC review of formaldehyde [130], including cohort studies of occupational exposure to formaldehyde in Finland [139], the U.S. [140], and Italy [141], as well as case-control studies in Hong Kong [136] and Finland, Sweden, Norway, and Iceland combined [142], also found no excess risk of NPC. Equivocal results were reported in a retrospective cohort study of laminated plastic workers in Italy, which did not estimate an RR for NPC [143]; and our own population-based case-control study in southern China, where we found a positive association with self-reported ever occupational exposure to formaldehyde, but no trend with duration of use [144].

The pivotal epidemiological evidence cited in support of an effect of formaldehyde on NPC comes from a retrospective cohort study of over 25,000 workers employed before 1966 at 10 U.S. plants that produced or used formaldehyde [145-147]. Although initial analyses revealed an overall increase in NPC risk in this cohort, no significant overall excess of NPC mortality was observed with the most recent follow-up through 2004, nor were significant positive exposure-response trends observed when exposed workers were compared with unexposed workers as the reference group [147]. Positive trends in NPC risk were seen for peak exposure, average exposure intensity, and cumulative exposure only when unexposed workers were excluded from the analysis. The positive findings, however, were driven by a single plant where six of the 11 observed NPC deaths occurred, whereas the remaining nine plants exhibited a nonsignificant
deficit of NPC mortality [148, 149]. An unusually small proportion of deaths coded as unspecified pharyngeal cancer at that single site suggests that diagnostic or death certificate coding bias, along with shared exposures encountered elsewhere, probably contributed to the apparent excess of NPC [150, 151]. Thus, the experience at that location may be the exception rather than the rule.

*Dust inhalation*

Studies of NPC risk among wood workers and others with potential occupational exposure to wood dust have yielded mixed results, with positive associations in several studies, but no association in others, including ours [144]. Two recent meta-analyses found statistically significant positive associations between occupational exposure to wood dust and NPC risk, but results were highly heterogeneous across studies [152, 153], raising questions of whether the apparent association depends on NPC histological subtype, type of wood dust (e.g., hardwood or softwood), or level of exposure, as well as whether misclassification with sinonasal adenocarcinoma might contribute to the observed findings.

Occupational exposure to other types of dust, especially cotton dust, have also been studied with respect to NPC risk. Three studies from China found a significant increase in NPC risk among textile workers, with positive exposure-response trends [108, 154, 155], whereas another study from southern China found a significant deficit of NPC in association with occupational exposure to cotton dust [74]. Our study in the same geographic region found positive associations of NPC risk with self-reported occupational exposure to dust from textiles, metals, cement, or coal, including positive gradients in risk with longer duration of exposure, whereas we found no significant association with self-reported occupational exposure to dust from soil, leather, or chalk [144].

*Genetic and molecular risk factors*

**Inherited susceptibility**

Familial aggregation of NPC has long been documented across high-, intermediate-, and low-incidence populations [156, 157]. A first-degree family history of NPC has been found to confer a 4- to 20-fold increase in risk of NPC [156, 158, 159], but not other malignancies [160]. In our population-based case-control study in southern China, the cumulative risk of NPC was 5.0% (95% CI: 3.6%–7.0%) for males and 1.9% (95% CI 1.1%–1.2%) for females with an affected relative [161]. Risk was highest in siblings, with a lifetime cumulative risk of 6.3% in brothers, 3.5% in sisters, 3.4% in fathers, and 2.5% in mothers of NPC cases.

A complex segregation analysis of familial NPC in southern China found that multiple genetic variants and shared environmental factors, rather than a single major susceptibility gene, most likely explain the observed pattern of inheritance [162]. Consistent with this hypothesis, disparate results were reported in four familial linkage studies, which identified four different regions of the genome shared by southern Chinese family members with NPC: a locus at 6p22
that conferred an RR of 20.9 [163]; a locus at 4p12–p15 with a logarithm of odds for linkage (LOD) score of 4.2 [164]; a locus at 3p21.31–21.2 with an LOD score of 4.18 [165]; and a locus at 5p13.1 with an LOD score of 2.1 [166]. Focusing on SNPs included in standard commercial genotyping arrays, a study using genome-wide data from southern Chinese NPC cases and controls estimated that 10% of the heritability of NPC was explained by these common genetic variants [167].

**Genome-wide association studies and the human leukocyte antigen region**

Table 1 summarizes findings from six genome-wide association studies (GWAS) of NPC conducted in Chinese populations, each evaluating 400,000–600,000 SNPs in hundreds or thousands of subjects [168-173]. Putative susceptibility loci in the human leukocyte antigen (HLA) region were identified across multiple GWAS, including SNPs in HLA-A, -B, -C, and -DQ/DR [168, 170-172].

These findings corroborate links detected in the 1970s between NPC risk and certain alleles in the HLA region, which encodes proteins required for antigen presentation and other essential immune system functions [174, 175]. NPCs risk thus appears to be influenced by genetic variation in the ability of antigen-presenting cells to bind and present intracellular antigens, including those from EBV, to cytotoxic T cells—that is, the function of class I HLA molecules (encoded by HLA-A, -B, and -C). Although the identification of specific causal alleles is complicated by the high density of genes and strong linkage disequilibrium in the HLA region, consistent positive associations of NPC risk have been detected with HLA-A*0207 (in linkage disequilibrium with HLA-B*4601) and HLA-B*5801 (in linkage disequilibrium with HLA-A*3303), and an inverse association has consistently been found with HLA-A*1101 (in linkage disequilibrium with HLA-B*13) [157, 176, 177]. A contributing role of class II HLA molecules (encoded by HLA-DR, -DQ, -DM, and -DP), which present extracellular antigens to helper T cells that in turn stimulate antibody-producing B cells, is also suggested by studies in Chinese, Tunisian, and Caucasian populations [178], as well as findings from one GWAS [168, 179], but not another GWAS [180], which found associations with the class I region but not the class II region.

**Candidate-gene studies**

Numerous epidemiological studies of genetic susceptibility to NPC have evaluated various SNPs in genes involved in immune function, inflammation, DNA repair, biotransformation, and other pathways hypothesized to be involved in NPC development. Although many statistical associations have been reported, few have been detected consistently across multiple study populations, making it difficult to distinguish causal effects from chance findings. Given the substantial uncertainty concerning the impact of common genetic variants on NPC risk, candidate-gene studies are summarized here only briefly, focusing on SNPs with a sufficient number of published studies to prompt meta-analysis (despite the questionable validity and value of such meta-analyses themselves [181]) and selecting genes with a broad range of functions.
The cytochrome P450 (CYP) and glutathione-S-transferase (GST) gene families encode phase I and II xenobiotic metabolism enzymes involved in the biotransformation of chemicals such as N-nitrosamines, components of tobacco smoke, and other toxins. Meta-analyses including up to 12 case-control studies of NPC in relation to genetic polymorphisms in GSTM1 and GSTT1 found overall meta-ORs of approximately 1.4–1.5 for the GSTM1 null genotype and approximately 1.3–1.4 for the GSTT1 null genotype, including after restriction to Asians [182, 183]. A meta-analysis of associations with genetic polymorphisms in CYP2E1 found significant positive associations of NPC risk with the Rsal/PstI variant (meta-OR = 2.7 under the recessive model), based on seven studies (six in Asians), and the DraI variant (meta-OR = 1.7 under the recessive model), based on three studies [184].

Genetic variants in DNA repair genes are also of natural interest as potential cancer risk factors. A meta-analysis of polymorphisms in the X-ray repair cross-complementing protein 1 gene (XRCC1) found a positive association of NPC risk with the Arg399Gln variant (meta-OR = 1.3 under the recessive model), based on 10 studies, but no association with the Arg194Trp and Arg280Gln variants [185]. Results restricted to Asians were similar. A separate evaluation of the Thr241Met polymorphism in XRCC3 reported a meta-OR of 3.1 under the recessive model, based on four NPC case-control studies in Asians [186].

In a meta-analysis including 10 NPC case-control studies of the Arg72Pro polymorphism in the cell-cycle regulation/tumor suppressor gene TP53, the overall meta-OR was 1.5 under the recessive model, with a stronger association in Caucasians than Asians, although the meta-OR was significantly positive in both populations [187]. The IG>2G polymorphism at position -1607 in the matrix metalloproteinase-1 (MMP1) gene, which is involved in the degradation of interstitial collagens and can enhance tumor growth and invasion, exhibited a meta-OR of 0.61 under the homozygous model (meta-OR = 0.91 under the recessive model) based on six studies, with similar results based on the five studies of Asians [188]. The meta-OR for the 8473 T>C polymorphism in the pro-inflammatory cyclooxygenase-2 (COX2) gene was 0.65 under the recessive model, based on three studies of NPC (two in Caucasians and one in Asians) [189]. Across four NPC case-control studies in Asian populations, the -94 ATTG insertion/deletion polymorphism in the nuclear factor-kappa-B subunit 1 gene (NFKB1), the meta-OR was 0.78 under the recessive model [190].

Conclusions

The landscape of NPC epidemiology has shifted in recent years, with more powerful and in-depth genetic studies, and new findings from several rigorously designed population-based studies in high-incidence regions. Where consumption of Chinese-style salted fish was once one of the most prominent known risk factors for NPC, our recent results suggest that it has a more modest influence on NPC risk than previously estimated, especially for intake during adulthood. The pivotal roles of EBV infection and HLA-A and HLA-B class I genotypes remain clear, and recent investigations into both viral and host genetic variation have identified key polymorphisms that may explain a substantial proportion of NPC in southern China. Viral and
host genetic factors provide the basis for promising NPC risk prediction models (e.g., [31, 191-193]), and EBV remains the foundation for developing new NPC screening strategies (e.g., [194]). Although causal, tobacco smoking appears to have a lesser impact on NPC development than on other head and neck cancers, especially squamous cell carcinomas. Otherwise, besides age, sex, and race/ethnicity, risk factors for NPC remain elusive (Table 2). Associations with indoor and outdoor air pollution merit further investigation due to the large potential population attributable risk from these widespread exposures, if causal. Occupational inhalation exposures, by contrast, are unlikely to account for a sizable fraction of NPC cases, even if causal. Collaborative pooled studies will be needed to achieve the statistical power required to reveal how viral, genetic, and environmental factors interact to cause NPC, and how intervention against modifiable characteristics can prevent it.
Acknowledgments

The authors thank Ms. Betty Dowd (Exponent) for her assistance with graphic design.

Funding support for this work was provided by the Sun Yat-sen University Cancer Center; and by the Foundation for Chemistry Research and Initiatives (FCRI), a non-profit organization established by the American Chemistry Council, as an honorarium to Exponent. The FCRI was not involved in the conception, writing, or submission of this article, and did not review any portion of the manuscript prior to journal submission or publication.

Ellen Chang is an employee of Exponent, an international science and engineering consulting company that provides consulting support to clients in industry, government, academia, and the private sector on some of the environmental exposures discussed in this review. The other authors declare no financial conflicts of interest.
References


Table 1. Results of genome-wide association studies of nasopharyngeal carcinoma

<table>
<thead>
<tr>
<th>Study</th>
<th>Cases (n)</th>
<th>Locus</th>
<th>Gene</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ng et al. (2009)</td>
<td>111 + 168</td>
<td>rs2212020, rs189897</td>
<td>Integrin-α 9 (ITGA9), 3p21</td>
</tr>
<tr>
<td>Tse et al. (2009)</td>
<td>277 + 339 + 296</td>
<td>rs2517713, rs2975042</td>
<td>HLA-A, 6p21.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>rs3129055, rs9258122</td>
<td>HLA-F, 6p21.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>rs29232</td>
<td>Gamma aminobutyric acid b receptor 1 (GABBR1), 6p21.3</td>
</tr>
<tr>
<td>Bei et al. (2010)</td>
<td>1,583 + 3,507 + 279</td>
<td>rs2860580, rs2894207</td>
<td>HLA-A, 6p21.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>rs28421666</td>
<td>HLA-B/C, 6p21.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>rs9510787</td>
<td>HLA-DQ/DR, 6p21.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>rs6774494</td>
<td>Tumor necrosis factor receptor superfamily member 19 (TNFRSF19), 13q12</td>
</tr>
<tr>
<td></td>
<td></td>
<td>rs1412829</td>
<td>Myelodysplasia 1 and ecotropic viral insertion site 1 (MDS1-EVI1), 3q26</td>
</tr>
<tr>
<td>Tang et al. (2012)</td>
<td>567 + 356 + 482</td>
<td>HLA-A 62Gln, HLA-B 16Leu, 116Leu</td>
<td>HLA-A<em>11:01, HLA-B</em>13:01, HLA-B<em>55:02, HLA-B</em>38:02</td>
</tr>
<tr>
<td>Chin et al. (2015)</td>
<td>184 + 260</td>
<td>rs11136697, HLA-A 99Cys</td>
<td>HLA-A<em>02:07, HLA-A</em>11:01</td>
</tr>
<tr>
<td>Cui et al. (2016a)</td>
<td>1,583 + 1,925 + 3,538</td>
<td>rs401681</td>
<td>Telomerase reverse transcriptase/cleft lip and palate transmembrane protein 1-like (TERT-CLPTM1L), 5p15</td>
</tr>
<tr>
<td></td>
<td></td>
<td>rs6498114</td>
<td>Class II major histocompatibility complex transactivator (CITA), 16p13</td>
</tr>
</tbody>
</table>

HLA: human leukocyte antigen
Table 2. Risk factors and preventive factors for nasopharyngeal carcinoma

<table>
<thead>
<tr>
<th>Risk factor</th>
<th>Direction of association*</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Well-confirmed risk factors</strong></td>
<td></td>
</tr>
<tr>
<td>Older age (up to ~60 years in high-incidence areas)</td>
<td>↑↑</td>
</tr>
<tr>
<td>Male sex</td>
<td>↑</td>
</tr>
<tr>
<td>Cantonese ethnic background</td>
<td>↑↑</td>
</tr>
<tr>
<td>Tobacco smoke</td>
<td>↑</td>
</tr>
<tr>
<td>Chinese-style salted fish (in early life)</td>
<td>↑</td>
</tr>
<tr>
<td>Epstein-Barr virus infection (positive IgA serology)</td>
<td>↑↑</td>
</tr>
<tr>
<td>First-degree family history of NPC</td>
<td>↑↑</td>
</tr>
<tr>
<td>Certain HLA-A and HLA-B alleles</td>
<td>↑/↓</td>
</tr>
<tr>
<td><strong>Possible risk factors, based on substantial data</strong></td>
<td></td>
</tr>
<tr>
<td>Indoor air pollution</td>
<td>↑</td>
</tr>
<tr>
<td>Other preserved foods</td>
<td>↑</td>
</tr>
<tr>
<td>Fresh fruits and vegetables</td>
<td>↓</td>
</tr>
<tr>
<td>Epstein-Barr virus sequence variation</td>
<td>↑↑</td>
</tr>
<tr>
<td>Chronic respiratory tract infection/inflammation</td>
<td>↑</td>
</tr>
<tr>
<td>HIV/AIDS</td>
<td>↑</td>
</tr>
<tr>
<td>Occupational wood dust</td>
<td>↑</td>
</tr>
<tr>
<td>Other types of occupational dust or smoke</td>
<td>↑</td>
</tr>
<tr>
<td>Certain HLA-D alleles</td>
<td>↑/↓</td>
</tr>
<tr>
<td><strong>Inconsistent findings</strong></td>
<td></td>
</tr>
<tr>
<td>Alcohol</td>
<td>↑/–</td>
</tr>
<tr>
<td>Tea</td>
<td>↓/–</td>
</tr>
<tr>
<td>Outdoor air pollution</td>
<td>↑/–</td>
</tr>
<tr>
<td>Occupational formaldehyde</td>
<td>↑/–</td>
</tr>
<tr>
<td>Traditional herbal medicines</td>
<td>↑/↓/–</td>
</tr>
</tbody>
</table>

HLA: human leukocyte antigen; IgA: immunoglobulin A; NPC: nasopharyngeal carcinoma

*Arrows indicate the approximate magnitude of the relationship, although the magnitude can vary substantially depending on the intensity, duration, and other characteristics of exposure.

↑: slight to moderate increase in risk
↑↑: moderate to large increase in risk
↓: slight to moderate decrease in risk
↓↓: moderate to large decrease in risk
↑/↓: slight to moderate increase or decrease in risk, depending on genotype or exposure type
–: no change in risk
Figure legends

Figure 1. Estimated age-standardized (to the world population standard) incidence rates of nasopharyngeal cancer among males by country, 2018. Data and map from the Global Cancer Observatory [3].

Figure 2. Estimated age-standardized (to the world population standard) incidence rates of nasopharyngeal cancer among females by country, 2018. Data and map from the Global Cancer Observatory [3].

Figure 3. Age-standardized (to the world population standard) incidence rates of nasopharyngeal cancer among men by city or county, China, 2008–2012. Data from Cancer Incidence in Five Continents, Volume XI [4].
Estimated age-standardized incidence rates (World) in 2018, nasopharynx, males, all ages

Data source: GLOBCAN 2018
Graph production: MSIC
World Health Organization

© International Agency for Research on Cancer 2018

All rights reserved. The designations employed and the presentation of the material in this publication do not imply the expression of any opinion whatsoever on the part of the World Health Organization / International Agency for Research on Cancer concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. Dotted and dashed lines on maps represent approximate borderlines for which there may not yet be full agreement.

ASR (World) per 100,000
- ≥ 1.6
- 0.84-1.6
- 0.53-0.84
- 0.36-0.53
- < 0.36
- Not applicable
- No data
Estimated age-standardized incidence rates (World) in 2018, nasopharynx, females, all ages

Data source: GLOBOCAN 2018
Graph production: IARC
[http://gco.iarc.fr/]
World Health Organization

© International Agency for Research on Cancer 2014

All rights reserved. The designations employed and the presentation of the material in this publication do not imply the expression of any opinion whatsoever on the part of the World Health Organization / International Agency for Research on Cancer concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. Dotted and dashed lines on maps represent approximate boundaries for which there may not yet be full agreement.
The evolving epidemiology of nasopharyngeal carcinoma

Ellen T. Chang, Weimin Ye, Yi-Xin Zeng, et al.

Cancer Epidemiol Biomarkers Prev  Published OnlineFirst April 13, 2021.