CrossMark

CANCER

Combined effect of tobacco smoking and alcohol drinking in the risk of head and neck cancers: a re-analysis of case–control studies using bi-dimensional spline models

Luigino Dal Maso[1] · Nicola Torelli[2] · Elisa Biancotto[1] · Matteo Di Maso[1] ·
Andrea Gini[1] · Gianni Franchin[3] · Fabio Levi[4] · Carlo La Vecchia[5] ·
Diego Serraino[1] · Jerry Polesel[1]

Abstract The synergistic effect of tobacco smoking and alcohol consumption on the risk of head and neck cancers has been mainly investigated as a cross-product of categorical exposure, thus leading to loss of information. We propose a bi-dimensional logistic spline model to investigate the interacting dose–response relationship of two continuous exposures (i.e., ethanol intake and tobacco smoking) on the risk of head and neck cancers, representing results through three-dimensional graphs. This model was applied to a pool of hospital-based case–control studies on head and neck cancers conducted in Italy and in the Vaud Swiss Canton between 1982 and 2000, including 1569 cases and 3147 controls. Among never drinkers and for all levels of ethanol intake, the risk of head and neck cancers steeply increased with increasing smoking intensity, starting from 1 cigarette/day. The risk associated to ethanol intake increased with incrementing exposure among smokers, and a threshold effect at approximately 50 g/day emerged among never smokers. Compared to abstainers from both tobacco and alcohol consumption, the combined exposure to ethanol and/or cigarettes led to a steep increase of cancer risk up to a 35-fold higher risk (95 % confidence interval 27.30–43.61) among people consuming 84 g/day of ethanol and 10 cigarettes/day. The highest risk was observed at the highest levels of alcohol and tobacco consumption. Our findings confirmed a combined effect of tobacco smoking and alcohol drinking on head and neck cancers risk, providing evidence that bi-dimensional spline models could be a feasible and flexible method to explore the pattern of risks associated to two interacting continuous-exposure variables.

Keywords Spline models · Dose–response relationship · Head and neck cancers · Tobacco smoking · Alcohol intake

✉ Jerry Polesel
polesel@cro.it

1 Unit of Epidemiology and Biostatistics, CRO Aviano National Cancer Institute, IRCCS, via Franco Gallini, 2, 33081 Aviano, PN, Italy

2 Department of Economics, Business, Mathematics and Statistics, University of Trieste, Trieste, Italy

3 Division of Radiation Oncology, CRO Aviano National Cancer Institute, IRCCS, Aviano, Italy

4 Cancer Epidemiology Unit and Vaud Cancer Registry, Institute of Social and Preventive Medicine (IUMSP), Lausanne University Hospital (CHUV), Lausanne, Switzerland

5 Department of Clinical Sciences and Community Health, Università degli Studi di Milano, Milan, Italy

Introduction

Tobacco smoking and alcohol consumption are well established major risk factors for head and neck cancers (i.e. neoplasms of the oral cavity, pharynx, and larynx) [1, 2]. A synergistic effect of these two factors has been also consistently reported [2, 3].

Most of the current evidence on the effect of the interaction of tobacco smoking with alcohol drinking on the risk of head and neck cancers derives from case–control investigations [2, 3]. This interaction was mainly evaluated through step-functions, using cut-points to categorize the two exposures and providing risk estimates for the cross-product of the categorical exposures. Step-function analysis, however, assumes that the risk is constant within each category, which may lead to efficiency loss and interpretation difficulties [4–7].

The dose–response relations between alcohol consumption and tobacco smoking and head and neck cancers

risk have been already investigated by this research group through spline models in two previous studies [8, 9]. In this paper, we propose a bi-dimensional logistic spline model to investigate the interaction between tobacco smoking and alcohol consumption on the risk of head and neck cancers. This approach allows the estimation of a surface, which represents the risk associated to the concomitant exposure to ethanol and tobacco, and could be easily generalized to other continuous exposures.

Materials and methods

Data were obtained from a series of hospital-based case–control studies on head and neck cancers conducted in Italy and in the Swiss Canton of Vaud between, 1982 and 2000 [10–14]. These studies had similar designs and inclusion criteria. Briefly, cases were patients below the age of 80 years with incident, histologically confirmed cancer. Cases with a previous history of neoplastic diseases (metachronous) were excluded. The control groups included patients younger than 80 years who were admitted to the same network of hospitals where cases had been identified for a wide spectrum of acute, non-neoplastic conditions (e.g., traumas, other orthopaedic disorders, acute surgical conditions, disorders of nose, ear, eye, skin, or teeth). Patients admitted to hospital for any condition related to tobacco smoking or alcohol drinking, or with any disorder that might have induced long-term modification of diet were excluded from the control group. Trained nurses carried out all interviews in hospital settings, thus keeping the refusal rate of eligible cases and controls below 5 % in Italy and about 15 % in Switzerland. Controls were frequency-matched to cases according to area of residence, age, and year of interview. Patients could have served as controls for more than one study. The original control group was considered for site-specific analyses whereas, for pool estimates, the control groups of each study were collapsed and patients enrolled in more than one study were considered only once.

Similar questionnaires, including information on age, education, other socio-economic factors, a problem-oriented medical history, usual diet, alcohol drinking, and smoking habits were used in each study. Information on smoking included lifetime status (never, former, or current smoker), daily number of cigarettes/cigars and grams of tobacco pipe smoked, age at starting, and duration of the habit. Smokers were subjects who had smoked at least one cigarette/day for at least 1 year. In our computations, one gram of pipe-smoked tobacco corresponded to one cigarette, and one cigar to three cigarettes. Former smokers were defined as those who abstained from any type of

smoking since at least 12 months before cancer diagnosis or interview (for controls). In addition, the weekly numbers of drinks for several alcoholic beverages were elicited from the subjects. Taking into account the different ethanol concentration, one drink corresponded to approximately 125 ml of wine, 330 ml of beer and 30 ml of hard liquor (i.e., about 12 g of ethanol). Total alcohol intake, expressed in grams of ethanol per day (g/day), was computed as the sum of all reported alcoholic beverages. "Never drinkers" were patients who abstained from drinking lifelong; "Ex-drinkers" were individuals who had abstained from drinking for at least 12 months at the time of cancer diagnosis or interview (for controls). Satisfactory reproducibility of questions on self-reported smoking and drinking habits in our study populations has been previously reported [15, 16].

The main aim of this paper was to evaluate the interaction of alcohol consumption with tobacco smoking using the two exposures as continuous predictors, focusing on low- to medium exposures. Therefore, the analyses were restricted to people who had smoked ≤40 cigarettes/day and to those who had drunk ≤180 g/day of ethanol. This restriction also avoided odd behaviours of risk estimates due to sparse data. Thus, 1569 head and neck cancers cases (491 oral cavity cancers, 491 pharynx, and 587 of larynx) and 3147 controls were included in the present analysis (Table 1).

The dose–response relationship between the risk of cancer and the combined exposure to daily ethanol intake and cigarettes smoking was investigated through bi-dimensional regression spline models [17]. Briefly, the logit was estimated through a generalized semi-parametric model where the two exposures were included as a smoothly joined piecewise polynomial of defined degree, with constrains for continuity at each join point. Odds ratios (OR) and corresponding 95 % confidence intervals (CI) were also calculated according to the model presented in detail in "Appendix". An ad hoc SAS macro was developed.

The optimal number of segments was selected putting an increasing number of knots and selecting the model that minimizes the Akaike Information Criterion (AIC) [18]. This selection criterion compromises between a good fit and a simple model by taking into account the deviance and the number of parameters in the model. As a sensitivity analysis, risks estimated for models with higher AIC values were compared: the risk surface was quite stable in relation to modification of the number of knots and knots' location, reassuring on the robustness of the method.

In the present model, the number of estimated parameters increases exponentially with the degree of polynomial due to an increasing number of interactions, thus

Table 1 Distribution of 1569 cancer cases and 3147 controls according to gender, age, tobacco smoking and alcohol drinking

	Controls n (%)	All cancers n (%)	Cancers by anatomical sites		
			Oral cavity n (%)	Pharynx n (%)	Larynx n (%)
Sex					
Man	2287 (72.7)	1325 (84.5)	374 (76.2)	419 (85.3)	532 (90.6)
Woman	860 (27.3)	244 (15.6)	117 (23.8)	72 (14.7)	55 (9.4)
Age (years)					
<45	406 (12.9)	121 (7.7)	56 (11.4)	41 (8.4)	24 (4.1)
45–54	813 (25.8)	383 (24.4)	126 (25.7)	140 (28.5)	117 (19.9)
55–64	1031 (32.8)	614 (39.1)	180 (36.7)	191 (38.9)	243 (41.4)
≥65	897 (28.5)	451 (28.7)	129 (26.3)	119 (24.2)	203 (34.6)
Smoking habits					
Never	1280 (40.7)	123 (7.8)	69 (14.1)	29 (5.9)	25 (4.3)
Former	910 (28.9)	390 (24.9)	100 (20.4)	114 (23.2)	176 (30.0)
Current	957 (30.4)	1056 (67.3)	322 (65.6)	348 (70.9)	386 (65.8)
Tobacco smoking intensity (cigarettes/day)					
Never	1280 (40.7)	123 (7.8)	69 (14.1)	29 (5.9)	25 (4.3)
<15	773 (24.6)	374 (23.8)	113 (23.0)	136 (27.7)	125 (21.3)
15–24	765 (24.3)	742 (47.3)	221 (43.0)	228 (46.4)	303 (51.6)
25–40	329 (10.5)	330 (21.0)	98 (20.0)	98 (20.0)	134 (22.8)
Drinking habits					
Never	426 (13.5)	71 (4.5)	29 (5.9)	19 (3.9)	23 (3.9)
Former	128 (4.1)	151 (9.6)	44 (9.0)	66 (13.4)	41 (7.0)
Current	2593 (82.4)	1347 (85.9)	418 (85.1)	406 (82.7)	523 (89.1)
Alcohol drinking intensity (g/day)					
Never	426 (13.5)	71 (4.5)	29 (5.9)	19 (3.9)	23 (3.9)
<25	980 (31.1)	165 (10.5)	58 (11.8)	40 (8.2)	67 (11.4)
25–49	729 (23.2)	215 (13.7)	64 (13.0)	62 (12.6)	89 (15.2)
50–99	701 (22.3)	539 (34.4)	149 (30.4)	162 (33.0)	228 (38.8)
100–180	311 (9.9)	579 (36.9)	191 (38.9)	208 (42.4)	180 (30.7)

requiring an elevated sample size. Therefore, linear splines (i.e. polynomial degree = 1) were considered. Moreover, up to a total of 4 knots were considered for all head and neck cancers; in addition, different locations for the knots were evaluated. When analyses were conducted separately for each anatomical site, up to three knots were considered. The reference category was defined as "Never smokers and Never drinkers"; study area, sex, age at diagnosis, education, smoking and drinking habits (former vs. current) were included as covariates in the models' equations.

For the sake of comparison, risk estimates for all head and neck cancers were further calculated according to the step-function logistic model. Smoking intensity and ethanol intake were firstly categorized and a term for the cross-product of the categorical exposures was included in the regression equation. Considering the low number of abstainers from alcohol drinking in our population, the reference category for the step-function model was set to never smokers who drunk <25 g/day ethanol (i.e., 2 drinks/day). ORs were adjusted for the same covariates as the spline models.

Results

Cases and controls were predominantly men and reported similar distribution according to age (Table 1). Former smokers represented 25 % of all cancers and 29 % of controls whereas quitting from alcohol drink was less frequent in this study population (l0 % in cancer cases and 4 % in controls). As expected, the frequency of heavy smoker (i.e., 20–40 cigarettes/day) and heavy drinkers (i.e., ≥100 g/day of ethanol) was much higher among cancer cases than among controls.

Risk estimates for head and neck cancers for the best-fitting model were represented through a three-

Fig. 1 Mesh plot (**a**) and contour plot (**b**) of risk for head and neck cancers by ethanol consumption (g/day) and tobacco smoking (cigarettes/day) estimated through a bi-dimensional regression spline. The reference category was defined as "never drinkers and never smokers". Equation model includes terms for sex, age, centre, drinking and smoking habits (former vs. current). *Surface* is shown for best-fitting splines according to Akaike Information Criterion; *thicker grid lines* show knot locations. *Thicker lines* on graphic floor represent knots projection. *Black lines* in contour plots indicate iso-risk curves at defined levels of risk (odds ratios 2, 5, 10, 20, 50, 100)
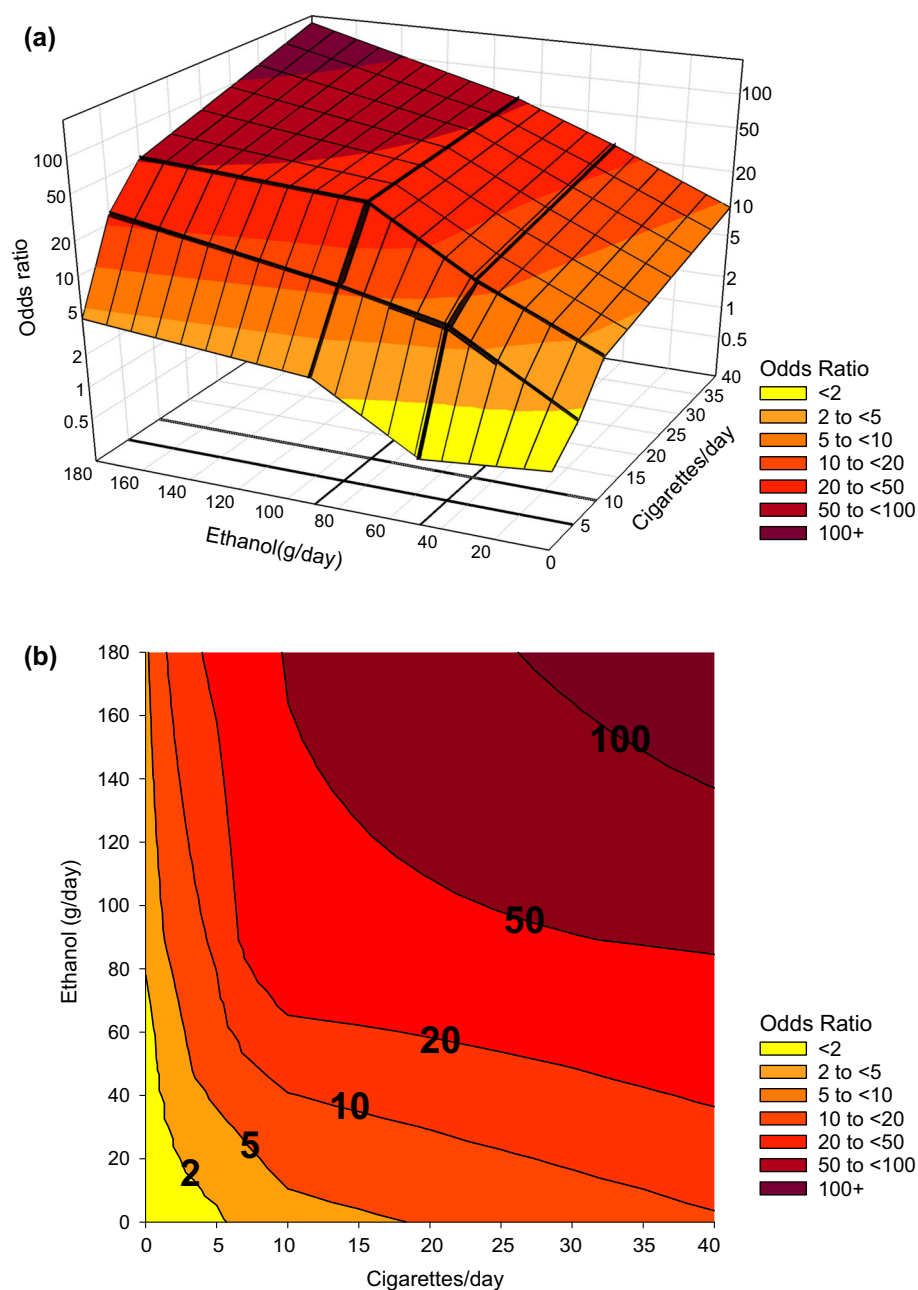
dimensional graph (mesh plot—Fig. 1a). Four knots, represented by thicker grid lines, were identified: two for ethanol intake (48 and 89 g/day) and two for tobacco smoking (5 and 10 cigarettes/day). Among never drinkers and for all levels of ethanol intake, the risk of head and neck cancer steeply increased with increasing smoking intensity, starting from 1 cigarettes/day. Conversely, the increased in risk associated to alcohol intake was less marked: the dose–response relationship between ethanol intake and cancer risk among never smokers showed a

threshold effect up to approximately 50 g/day, but this effect was no longer appreciable among smokers (Fig. 1a). The combined exposure to ethanol and/or cigarettes intake led to a steep increase of cancer risk up to a 35-fold higher risk (95 % CI 27.30–43.61) among those consuming 89 g/day of ethanol and 10 cigarettes/day, compared to abstainers from both tobacco and alcohol (Fig. 1a). The highest risk was observed at the highest levels of alcohol and tobacco consumption (OR 172.25; 95 % CI 78.21–379.40).

Risk estimates were also displayed through a contour plot (Fig. 1b), i.e. a two-dimensional graph showing lines that identify combinations of tobacco smoking and ethanol intakes for which the risk of head and neck cancers is constant. A steeper risk increase with tobacco smoking rather than with ethanol intake is outlined.

Risk estimates were reported separately for the three anatomical sites in Fig. 2. A similar pattern emerged for all the considered sites, but the effect of smoking seemed stronger for laryngeal cancer. For pharyngeal cancer, a threshold effect for ethanol intake up to 40 g/day emerged, similarly to that reported for all head and neck cancers; however, this threshold effect disappeared as the number of cigarettes increased.

Table 2 reports the ORs for the cross-products of ethanol intake and cigarettes/day smoked according to the step-function model with lowest and highest ORs estimated through spline models within the same cross-product category. The risk estimates of the two models were generally in agreement, except for the highest levels of the exposures, with ORs from the step-function model between the lowest and the highest ORs from the spline models. Nonetheless, ORs from the step-function model were averages of risks, which may vary considerably within the given cross-product. For instance, the OR for people smoking 1–14 cigarettes/day and drinking 100–180 g/day of ethanol was 21.61 (95 % CI 14.90–31.36) according to step-function model, whereas, according to spline model, the risk varied within that exposure categories from 3.89 to 61.41.

Discussion

Our results show that bi-dimensional regression spline models are a feasible and flexible method to estimate the interacting dose–response relationship between two continuous exposure variables, with higher power to detect the variation in risk estimates than traditional step-function models. Two- and three-dimensional graphical representations were provided (i.e., mesh plot and contour plot) to display the risk of head and neck cancers associated to the combined exposure to alcohol drinking and tobacco smoking.

Results from the present analysis were consistent with previous findings from the same data [8, 9]. In particular, the threshold effect of alcohol drinking observed in all head and neck cancers and in pharyngeal cancer for low consumption holds true only among never smokers. Conversely, among tobacco smokers, cancer risks increased across the whole range of ethanol intake starting from a very low number of cigarettes/day [9]. The risk found among heavy smokers and heavy drinkers was very high compared to never smokers/never drinkers, in agreement with what reported in a previous study from the same study areas [19]. Of course, the results for all head and neck cancers taken together should be considered with caution since they represent the risk pattern of a population with the same distribution for each cancer site.

Different dose–response relationship emerged for the three anatomical sites considered. The number of knots varied according to cancer site, as well as their location, ensuring the power of the methods to capture different relationship between the two exposures. In particular, the threshold effect for low ethanol consumption in never smokers emerged only for pharyngeal cancer and the effect of tobacco smoking was more relevant for laryngeal cancer, as previously reported by several studies [2, 9].

The standard approach to investigate interacting risk in case–control study is through logistic step-function. The two exposures were firstly categorized and risk estimates were calculated for the cross-product of the two categorical exposures, assuming that the risk was constant within each cross-product category. In this paper, we have compared risk estimates from step-function and from spline models, highlighting that the assumption of risk steadiness within the cross-product category was rarely respected. Bi-dimensional spline allowed to overcome this limitation, since the risk may vary within each cross-product category according to the polynomial degree. Moreover, the selection of the number of knots and their location may increase model flexibility. In addition, step-function did not allow the estimation of risk using never smokers who never drunk as reference category, as the sample size was inadequate for the model (i.e. 24 cases, 1.53 %). For this analysis, the inclusion in the reference category of never smokers who drunk <25 g/day of ethanol did not actually impact risk estimates, as they are not at increased risk of head and neck cancers. For the sake of caution and parsimony, only linear splines were considered in the present analysis. However, both the theoretical model and the developed SAS macro can be generalized to higher polynomial degree and to any other continuous exposure.

This method is affected by most of the limitations of univariate spline models. Firstly, the choice of some model setting parameters (e.g., maximum number of knots, highest polynomial degree, criteria for knots location) is arbitrary. Indeed, better estimates could be obtained forcing the allocation of knots, but this required an a priori knowledge of where the knots should be located. However, the selection of boundaries seems to be of less importance as compared to step-function analysis, since the flexibility
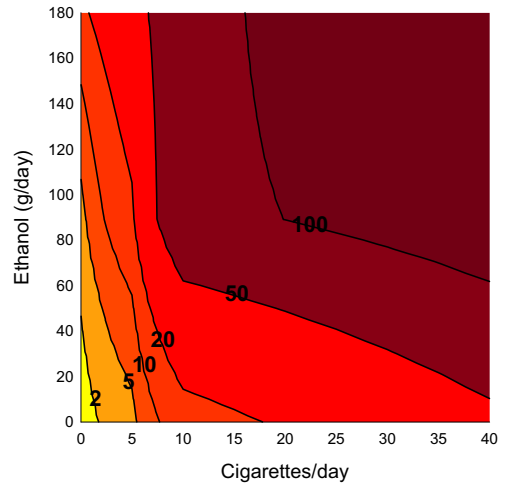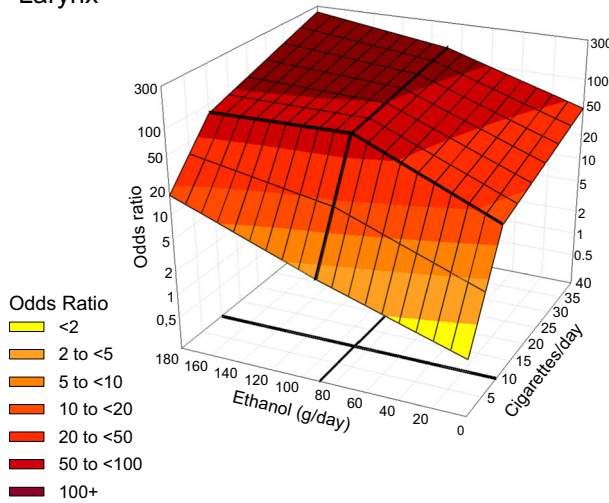
Oral cavity
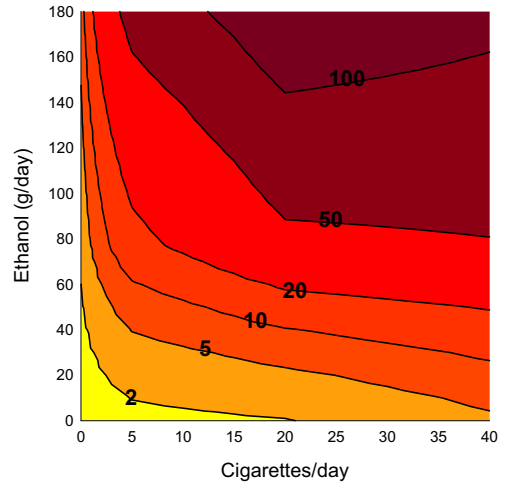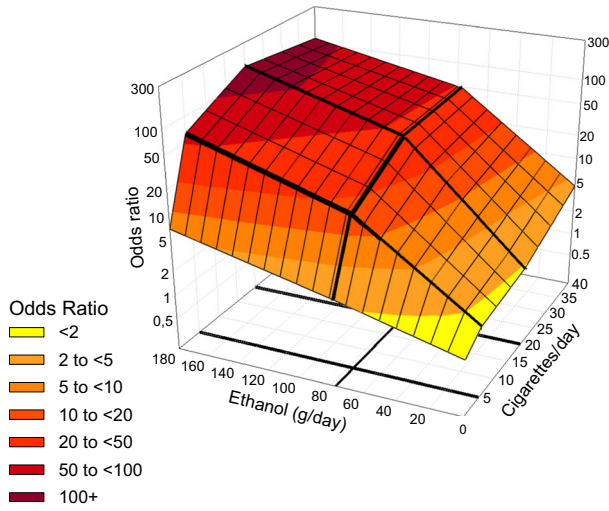
Pharynx
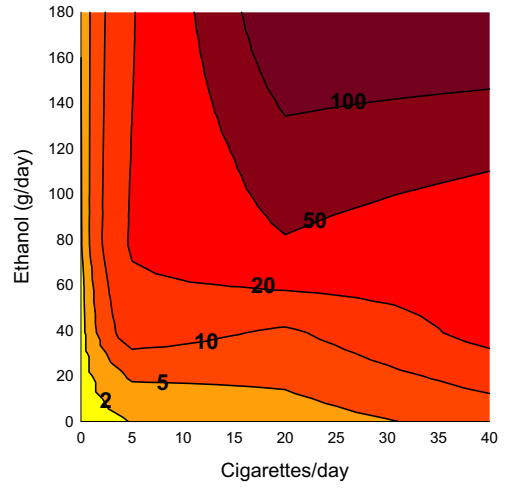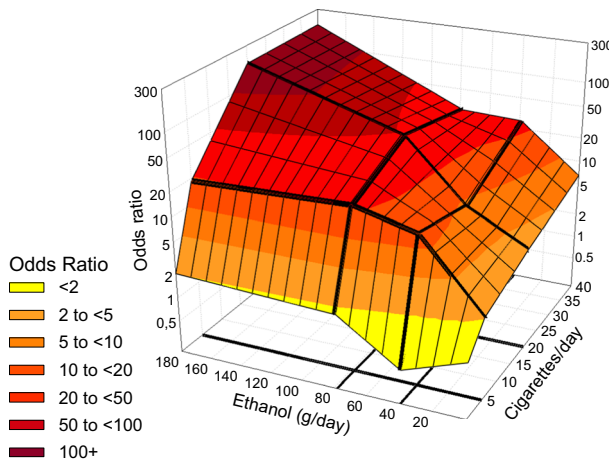
Larynx

◄**Fig. 2** Mesh plot and contour plot of risk for cancer of the oral cavity (n = 491), pharynx (n = 491), and larynx (n = 587), by ethanol consumption (g/day) and tobacco smoking (cigarettes/day). Risks were estimated through a bi-dimensional regression spline, taking "never drinkers and never smokers" as reference category. Equation model includes terms for sex, age, centre, drinking and smoking habits (former vs. current). *Surfaces* are shown for best-fitting splines according to Akaike Information Criterion; *thicker grid lines* show knot locations. *Thicker lines* on graphic floor represent knots projection. *Black lines* in contour plots indicate iso-risk curves at defined levels of risk (odds ratios 2, 5, 10, 20, 50, 100)

of the model could partly attenuate a bad selection of knots [6]. Secondly, the number of model parameters greatly increases with the power of the polynomial and with the number of knots, thus limiting the potentiality of the methods in absence of very large databases. The present pool analysis seemed to have an adequate sample size (at least for all head and neck cancers combined), which should be accounted for among the strengths of the study. Moreover, although confidence intervals were estimated, they could not be represented through the proposed plots. Finally, we did not formally investigate whether the two exposures interacted in an additive or multiplicative way.

Although this would have been possible through standard approaches (e.g., synergy index S), the great variability of risk estimates in the highest levels of exposures suggested caution.

Results of the present analysis may suffer from limitations common of a case–control design such as recall bias, selection bias, and use of hospital controls [20]. However, the reproducibility of alcohol intake and tobacco smoking was good in our study [15, 16]; in addition, the exclusion of very high exposures (i.e., >180 g/day of ethanol and >40 cigarettes/day) allowed us to rule out most of the possible misclassifications and the effect of confounding factors connected to severe alcoholism and/or heavy tobacco smoking [21, 22]. The almost complete participation of cases and controls and a direct interview through trained nurses contributed to the strength of the study.

Results from the present study strengthened the previous findings on the effect of tobacco smoking and alcohol drinking on the risk of head and neck cancers. This study also provided evidence that flexible models, such as bi-dimensional spline ones, may help to detect interactions between two continuous exposures without strong a priori hypothesis on the shape of such relationship.

Table 2 Odds ratio (OR) and 95 % confidence intervals (CI)[a] for head and neck cancers for tobacco smoking and alcohol drinking according to step-function regression model and bi-dimensional regression spline

Cigarettes/day	Ethanol (g/day)	Ca:Co	Step-function regression OR (95 % CI)	Bi-dimensional regression spline	
				Lowest OR (95 % CI)	Highest OR (95 % CI)
0	0–24	81:1035	1 (Reference)	0.89 (0.68–1.15)	1 (Reference)
0	25–49	17:375	0.67 (0.39–1.15)	0.78 (0.48–1.28)	0.88 (0.68–1.15)
0	50–99	14:271	0.98 (0.54–1.79)	0.83 (0.53–1.32)	2.91 (1.64–5.14)
0	100–180	10:81	2.27 (1.11–4.63)	2.92 (1.66–5.12)	4.20 (1.18–14.95)
1–14	0–24	59:372	2.85 (1.99–4.07)	1.11 (0.80–1.54)	7.58 (6.44–8.92)
1–14	25–49	64:282	5.33 (3.68–7.74)	1.17 (0.93–1.47)	13.43 (10.88–16.57)
1–14	50–99	126:336	10.71 (7.60–15.08)	1.29 (0.88–1.88)	43.52 (36.36–52.16)
1–14	100–180	123:155	21.61 (14.90–31.36)	3.89 (2.46–6.16)	61.41 (42.11–89.55)
15–24	0–24	68:307	4.68 (3.25–6.73)	4.53 (3.36–6.10)	9.93 (8.53–11.57)
15–24	25–49	90:244	9.63 (6.68–13.86)	7.96 (6.84–9.27)	17.28 (14.31–20.87)
15–24	50–99	282:348	22.32 (16.22–30.71)	14.17 (11.67–17.18)	49.72 (42.46–58.21)
15–24	100–180	294:207	41.11 (29.33–57.61)	44.30 (34.46–52.40)	91.31 (66.02–126.28)
25–40	0–24	28:109	6.21 (3.76–10.24)	6.02 (4.46–8.13)	15.33 (10.79–21.77)
25–40	25–49	44:111	10.88 (6.90–17.15)	10.43 (8.92–12.21)	25.87 (16.71–40.05)
25–40	50–99	117:141	25.83 (17.67–37.75)	18.16 (14.98–22.01)	61.52 (42.86–88.32)
25–40	100–180	136:80	50.12 (33.54–74.91)	50.78 (43.16–59.75)	172.25 (78.21–379.40)

[a] Estimates were adjusted sex, age, centre, smoking habits, and drinking habits

Appendix: univariate logistic spline model

Let $y = [y_1,\ldots, y_n]'$ be a $n \times 1$ vector of realizations of a random variable Y with binomial distribution so that $\pi = Pr\{Y = 1\}$ and $x = [x_1,\ldots, x_n]'$ be a $n \times 1$ vector of related realizations of a continuous predictor X of π. The additive logistic regression model:

$$\log \text{it}(\pi) = \log\left[\frac{\pi}{1-\pi}\right] = \alpha + f(X)$$

where f is an arbitrary spline function, i.e. a piecewise polynomial truncated function defined as:

$$f(X) = \beta_0 + \sum_{d=1}^{D} \beta_{1,d} X^d + \sum_{k=1}^{K} \beta_{2,k}(X - t_k)_+^D$$

where D is the degree of polynomial, t_k is the kth knot location, K is the number of knots, and

$$(X - t_k)_+^D = \begin{cases} (X - t_k)^D, & \text{if } X \geq t_k \\ 0, & \text{otherwise} \end{cases}$$

indicates a value of zero for negative values of the argument. The intercept α is absorbed into f and the continuity on knots is guaranteed by contrasts $(\cdot)_+$. The log-likelihood function of the additive logistic regression model as:

The pointwise $(1 - \alpha)$ confidence band is $\widehat{\beta} X \pm z_{(1-\alpha/2)}(XVX')^{1/2}$ where

$$V = \left(\sum_{i=1}^{n} \left[1 - \frac{e^{\breve{x}_i'\widehat{\beta}}}{1 + e^{\breve{x}_i'\widehat{\beta}}} \right] \breve{x}_i \breve{x}_i' \right)^{-1}$$

is a $(D + K + 1) \times (D + K + 1)$ covariance matrix.

Bi-dimensional logistic spline model

Let $x = [x_1,\ldots, x_n]'$ and $z = [z_1,\ldots, z_n]'$ be two $n \times 1$ vector of related realizations of two continuous predictors X and Z. The corresponding spline function:

$$
\begin{aligned}
f(X,Z) = {} & \beta_0 + \sum_{d_x=1}^{D_x} \beta_{1,d_x} X^{d_x} + \sum_{k_x=1}^{K_x} \beta_{2,k_x}(X - t_{k_x})_+^{D_x} \\
& + \sum_{d_z=1}^{D_z} \beta_{3,d_z} Z^{d_z} \\
& + \sum_{k_z=1}^{K_z} \beta_{4,k_z}(Z - t_{k_z})_+^{D_z} + \sum_{d_x=1}^{D_x}\sum_{d_z=1}^{D_z} \beta_{5,d_x,d_z} X^{d_x} Z^{d_z} \\
& + \sum_{d_z=1}^{D_z}\sum_{k_x=1}^{K_x} \beta_{6,d_z,k_x}(X - t_{k_x})_+^{D_x} Z^{d_z} \\
& + \sum_{d_x=1}^{D_x}\sum_{k_z=1}^{K_z} \beta_{7,d_x,k_z} X^{d_x}(z - t_{k_z})_+^{D_z} \\
& + \sum_{k_x=1}^{K_x}\sum_{k_z=1}^{K_z} \beta_{8,k_x,k_z}(X - t_{k_x})_+^{D_x}(Z - t_{k_z})_+^{D_z}
\end{aligned}
$$

where, D_x and D_z are the degrees of the polynomials, t_{k_x} and t_{k_z} are the k_xth and k_zth knot location, K_x and K_z are the number of knots in the distribution, respectively, of X and Z

$$l(\beta) = \sum_{i=1} y_i \ln\left(\frac{e^{x_i \beta}}{1 + e^{\breve{x}_i'\beta}}\right) + \sum_{i=1}(1 - y_i)\ln\left(\frac{1}{1 + e^{\breve{x}_i'\beta}}\right)$$

where $\beta = [\beta_0, \beta_{1,1}, \ldots, \beta_{1,D}, \beta_{2,1}, \ldots, \beta_{2,K}]'$ is the $(D + K + 1) \times 1$ vector of parameters and $\breve{x}_i = [1, x_i, x_i^2, \cdots, x_i^D, (x_i - t_1)_+^D, \cdots, (x_i - t_K)_+^D]'$ is the $(D + K + 1) \times 1$ vector of covariates for the ith observation.

The maximum likelihood estimator of parameters $\widehat{\beta}$ can be found solving:

$$\frac{\partial L(\beta)}{\partial \beta} = \sum_{i=1}^{n} \left[y_i - \frac{e^{\breve{x}_i'\beta}}{1 + e^{\breve{x}_i'\beta}} \right] \breve{x}_i = 0$$

continuous predictors,

$$(X - t_{k_x})_+^{D_x} = \begin{cases} (X - t_{k_x})^{D_x}, & \text{if } X \geq t_{k_x} \\ 0 & \text{otherwise} \end{cases}$$

and

$$(Z - t_{k_z})_+^{D_z} = \begin{cases} (Z - t_{k_z})^{D_z}, & \text{if } Z \geq t_{k_z} \\ 0 & \text{otherwise} \end{cases}$$

indicate the zero value for negative values of the argument for X and Z respectively.

The parameters and the corresponding confidence intervals will be estimated through the methodology described above.

References

1. IARC. IARC Monographs on the evaluation of carcinogenic risk of chemicals to humans. Alcohol drinking. IARC Sci Publ No. 44. Lyon: IARC. 1988.
2. IARC. IARC Monographs on the evaluation of carcinogenic risks to Humans. Vol. 83: Tobacco smoke and involuntary smoking. Lyon: IARC Press. 2004.
3. Hashibe M, Brennan P, Chuang SC, et al. Interaction between tobacco and alcohol use and the risk of head and neck cancer: pooled analysis in the International Head and Neck Cancer Epidemiology Consortium. Cancer Epidemiol Biomarker Prev. 2007;18:541–50.
4. Zhao LP, Kolonel LN. Efficiency loss from categorizing quantitative exposures into qualitative exposures in case-control studies. Am J Epidemiol. 1992;136:464–74.
5. Weinberg CR. How bad is categorization? Epidemiology. 1995;6:345–7.
6. Greenland S. Dose–response and trend analysis in epidemiology: alternatives to categorical analysis. Epidemiology. 1995;6:356–65.
7. Desquilbet L, Mariotti F. Dose–response analyses using restricted cubic spline functions in public health research. Stat Med. 2010; doi:10.1002/sim.3841.
8. Polesel J, Dal Maso L, Bagnardi V, et al. Estimating dose–response relationship between ethanol and risk of cancer using regression spline models. Int J Cancer. 2005;114:836–41.
9. Polesel J, Talamini R, La Vecchia C, et al. Tobacco smoking and the risk of upper aero-digestive tract cancers: a reanalysis of case-control study using spline models. Int J Cancer. 2008;122:2398–402.
10. Franceschi S, Talamini R, Barra S, et al. Smoking and drinking in relation to cancers of the oral cavity, pharynx, larynx, and esophagus in northern Italy. Cancer Res. 1990;50:6502–7.
11. Levi F, Pasche C, La Vecchia C, et al. Food groups and risk of oral and pharyngeal cancer. Int J Cancer. 1998;77:705–9.
12. Franceschi S, Levi F, Dal Maso L, et al. Cessation of alcohol drinking and risk of cancer of the oral cavity and pharynx. Int J Cancer. 2000;85:787–90.
13. Talamini R, Vaccarella S, Barbone F, et al. Oral hygiene, dentition, sexual habits and risk of oral cancer. Br J Cancer. 2000;83:1238–42.
14. Talamini R, Bosetti C, La Vecchia C, et al. Combined effect of tobacco and alcohol on laryngeal cancer risk: a case–control study. Cancer Causes Control. 2002;13:957–64.
15. D'Avanzo B, La Vecchia C, Katsouyanni K, et al. Reliability of information on cigarette smoking and beverage consumption provided by hospital controls. Epidemiology. 1996;7:312–5.
16. Ferraroni M, Decarli A, Franceschi S, et al. Validity and reproducibility of alcohol consumption in Italy. Int J Epidemiol. 1996;25:775–82.
17. Ruppert D, Wand MP, Carroll RJ. Semiparametric regression. Cambridge: Cambridge University Press; 2003.
18. Akaike H. Information theory and an extension of the maximum likelihood principle. In: Second international symposium on information theory. Budapest. 1973:267–281.
19. Barón AE, Franceschi S, Barra S, et al. A comparison of the join effects of alcohol and smoking on the risk of cancers across sites in the upper aerodigestive tract. Cancer Epidemiol Biomarker Prev. 1993;2:519–23.
20. Rothman KJ, Greenland S, Lash TL. Modern epidemiology. Philadelphia: Lippincott Williams & Wilkins; 2008.
21. Seitz HK, Stickel F, Homann N. Pathogenetic mechanisms of upper aerodigestive tract cancer in alcoholics. Int J Cancer. 2004;108:483–7.
22. Chiolero A, Wietlisbach V, Ruffieux C, et al. Clustering of risk behaviors with cigarette consumption: a population-based survey. Prev Med. 2006;42:348–53.