Joint effects of intensity and duration of cigarette smoking on the risk of head and neck cancer
Di Credico, Gioia; Edefonti, Valeria; Polesel, Jerry; Pauli, Francesco; Torelli, Nicola; Serraino, Diego
Published in:
Oral Oncology
DOI:
10.1016/j.oraloncology.2019.05.006
Publication date:
2019
Document Version
Peer reviewed version
Link to publication in Discovery Research Portal

Citation for published version (APA):

General rights
Copyright and moral rights for the publications made accessible in Discovery Research Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from Discovery Research Portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain.
• You may freely distribute the URL identifying the publication in the public portal.

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Download date: 23. Sep. 2020
Joint effects of intensity and duration of cigarette smoking on the risk of head and neck cancer: a bivariate spline model approach


¹Department of Statistics, Padua University, Padua, Italy; ²Branch of Medical Statistics, Biometry

© 2019. This manuscript version is made available under the CC-BY-NC-ND 4.0 license http://creativecommons.org/licenses/by-nc-nd/4.0/
and Epidemiology "G. A. Maccacaro", Department of Clinical Sciences and Community Health, Università degli Studi di Milano, Milano, Italy; 3Cancer Epidemiology Unit, Centro di Riferimento Oncologico di Aviano (CRO), IRCCS, Aviano, Italy; 4Department of Economics, Business, Mathematics and Statistics, University of Trieste, Trieste, Italy; 5Department of Biomedical and Clinical Sciences, Università degli Studi di Milano, Milano, Italy; 6Univerité Rennes, INSERM, EHESP, Irset (Institut de Recherche en Santé, Environnement et Travail) - UMR_S 1085, Pointe-à-Pitre, France; 7Inserm, Center for Research in Epidemiology and Population Health (CESP), Cancer and Environment team, Université Paris-Sud, Université Paris-Saclay, Villejuif, France; 8Aichi Cancer Center Research Institute, Nagoya, Japan; 9International Agency for Research on Cancer, Lyon, France; 10Institute of Oncology Angel H. Roffo, University of Buenos Aires, Argentina; 11Institute of Oncology and Radiobiology, Havana, Cuba; 12Epidemiology - CIPE/ACCAMARGO, Sao Paulo, Brazil; 13Universidade Federal de Pelotas, Pelotas, Brazil; 14Hospital de Clinicas de Porto Alegre, Porto Alegre, Brazil; 15Escola Nacional de Saude Publica, Fundacao Oswaldo Cruz, Rio de Janeiro, Brazil; 16University of São Paulo School of Public Health, São Paulo, Brazil; 17Institute of Hygiene & Epidemiology, 1st Faculty of Medicine, Charles University, Prague, Czech Republic; 18Leibniz Institute for Prevention Research and Epidemiology, BIPS, Bremen, Germany; 19University of Bremen, Faculty of Mathematics and Computer Science, Bremen, Germany; 20Department of Hygiene, Epidemiology and Medical Statistics, School of Medicine, National and Kapodistrian University of Athens, Athens, Greece; 21University of Padua, Padova, Italy; 22Department of Medical Sciences, University of Turin, Turin,
Italy; 22Trinity College School of Dental Science, Dublin, Ireland; 24Cancer Registry of Norway, Oslo, Norway; 25Dental School, College of Medical, Veterinary and Life Sciences, University of Glasgow, Glasgow, UK; 26University of Aberdeen Dental School, Aberdeen, UK; 27University of Hong Kong, Hong Kong, China; 28Institut Catala Oncologia, Barcelona, Spain; 29University of São Paulo School of Medicine, São Paulo, Brazil; 30Penn State College of Medicine, Hershey, PA, USA; 31University of North Carolina School of Public Health, Chapel Hill, NC, USA; 32Division of Head and Neck Surgical Oncology in the Department of Otolaryngology/Head and Neck Surgery at Washington University School of Medicine, USA; 33National Cancer Institute, Bethesda, MD, USA; 34UT- M.D. Anderson Cancer Center, Houston, Texas, USA; 35Regional Authority of Public Health in Banska Bystrica, Slovakia; 36The M. Skasodowska-Curie Memorial Cancer Center and Institute of Oncology, Dept. of Cancer Epidemiology and Prevention, Warsaw, Poland; 37National Institute of Public Health, Bucharest, Romania; 38National Institute of Environmental Health to National Public Health Institute, Budapest, Hungary; 39Cancer Research Centre, Moscow, Russia; 40Institute of Occupational Medicine, Lodz, Poland; 41Roswell Park Cancer Institute, Buffalo, NY, USA; 42UCLA School of Public Health, Los Angeles, CA, USA; 43Departments of Epidemiology and Environmental Health Sciences, School of Public Health and Department of Urology, Medical School, University of Michigan, Ann Arbor, MI, USA; 44Institut Universitaire de Medecine Sociale et Preventive (IUMSP), Centre Hospitalier Universitaire Vaudois and University of Lausanne, Lausanne, Switzerland; 45College of Public Health, University of Iowa, Iowa City, IA, USA; 46Washington State University College of Pharmacy, Spokane, WA, USA; 47Department of
Oncology, IRCCS - Istituto di Ricerche Farmacologiche Mario Negri, Milan, Italy; 48Department of Otorhinolaryngology, School of Medicine and Surgery, University of Milano - Bicocca, Monza, Italy; 49Brown University, Providence, RI, USA; 50Boston University School of Public Health, Boston, MA, USA; 51University of Heidelberg, Germany; 52Fred Hutchinson Cancer Research Center, Seattle, WA, USA; 53Department of Epidemiology, Brown University School of Public Health, Providence, RI, USA; 54INSERM, Sorbonne Université, Institut Pierre Louis d’Epidémiologie et de Santé Publique IPLESP, Department of Social Epidemiology, Paris, France; 55Section of Hygiene - Institute of Public Health, Fondazione Policlinico Universitario A. Gemelli IRCCS, Rome, Italy; 56Università Cattolica del Sacro Cuore, Rome, Italy; 57Division of Epidemiology, New York University School Of Medicine, NY, New York, USA; 58Thoracic/Head and Neck Medical Oncology*, The University of Texas MD Anderson Cancer Center, TX, USA; 59New York Eye and Ear Infirmary, New York, NY, USA; 60Medical Informatics Center, Peking University, China; 61Division of Clinical Epidemiology and Aging Research, German Cancer Research Center (DKFZ), Heidelberg, Germany; 62Division of Preventive Oncology, German Cancer Research Center (DKFZ) and National Center for Tumor Diseases (NCT), Heidelberg, Germany; 63German Cancer Consortium (DKTK), German Cancer Research Center (DKFZ), Heidelberg, Germany; 64Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA; 65Department of Head and Neck Surgery, Division of Surgery, The University of Texas MD Anderson Cancer Center, Houston, TX, USA; 66Institute of Population Health Sciences, National Health Research Institutes, Taiwan; 67The Tisch Cancer Institute, Icahn School of Medicine at
Mount Sinai, New York, NY, USA; Division of Public Health, Department of Family & Preventive Medicine and Huntsman Cancer Institute, University of Utah School of Medicine, Salt Lake City, Utah, USA; Division of Public Health, Department of Family & Preventive Medicine, University of Utah School of Medicine, Salt Lake City, Utah, USA.

**Corresponding Author**: Valeria Edefonti, PhD, Branch of Medical Statistics, Biometry and Epidemiology "G. A. Maccacaro", Department of Clinical Sciences and Community Health, Università degli Studi di Milano, via Venezian 1, 20133 Milano, Italy; Tel: +39 02-50320873; fax: +39 02-50320866; email: valeria.edefonti@unimi.it.

**Word count**: 3197

**Conflicts of interest**: None declared.
Abstract – 246 words

Objectives: This study aimed at re-evaluating the strength and shape of the dose-response relationship between the combined (or joint) effect of intensity and duration of cigarette smoking and the risk of head and neck cancer (HNC). We explored this issue considering bivariate spline models, where smoking intensity and duration were treated as interacting continuous exposures.

Materials and Methods: We used individual-level pooled data from 33 case-control studies (18,260 HNC cases and 29,844 controls) participating in the International Head and Neck Cancer Epidemiology (INHANCE) consortium. In bivariate regression spline models, exposures to cigarette smoking intensity and duration (compared with never smokers) were modeled as a linear piecewise function within a logistic regression also including potential confounders. We jointly estimated the optimal knot locations and regression parameters within the Bayesian framework.

Results: For oral-cavity/pharyngeal (OCP) cancers, an odds ratio (OR) >5 was reached after 30 years in current smokers of ~20 or more cigarettes/day. Patterns of OCP cancer risk in current smokers differed across strata of alcohol intensity. For laryngeal cancer, ORs>20 were found for current smokers of ≥20 cigarettes/day for ≥30 years. In former smokers who quit ≥10 years ago, the ORs were approximately halved for OCP cancers, and ~1/3 for laryngeal cancer, as compared to the same levels of intensity and duration in current smokers.

Conclusion: Referring to bivariate spline models, this study better quantified the joint effect of intensity and duration of cigarette smoking on HNC risk, further stressing the need of smoking
cessation policies.

**Keywords:** bivariate spline models; cigarette smoking duration; cigarette smoking intensity; head and neck cancer; INHANCE; laryngeal cancer; oral cavity and pharyngeal cancers.

**Abbreviations:**

CI = credible interval

HNC = head and neck cancer

INHANCE = International Head and Neck Cancer Epidemiology

NUTS = No-U-Turn Sampler

OCP = oral cavity and pharynx

ORs = odds ratios

WAIC = Wanatabe-Akaike Information Criterion
Introduction

In recent years, tobacco smoking has confirmed its role as the worldwide leading cause of preventable diseases and death. It is responsible for at least 12% of all deaths and of 22% of all cancer deaths worldwide [1]. Tobacco has been strongly associated with head and neck cancer (HNC, i.e., cancers of the oral cavity, pharynx, and larynx), which is the sixth most common cancer type, accounting for nearly 900,000 new cases in 2018 [2].

A dose-response relationship with HNC risk has been reported for smoking intensity (cigarettes per day reported during the exposure period) and duration (years of exposure), as well as for a cumulative smoking exposure measured in pack-years [3,4]. Compared with intensity, smoking duration was shown to provide a greater risk of developing HNC [5,6], as well as, other tobacco-related cancers, such as lung [7,8] and bladder cancers [9,10].

So far, the association between tobacco smoking and HNC risk was evaluated by considering duration and intensity as either separate or interacting predictors. In the latter case, the combined exposure was modeled as either cross-product or pack-years [6,11,12]. Step-functions [4, page 369] were widely used to provide risk estimates for the cross-product of the categorized exposures. This approach assumed a constant rate within each combined category of exposure; however, this assumption may have been too rough and/or leading to some efficiency loss [13-15]. As an alternative, linear-exponential models estimated the excess odds ratio (OR) of HNC within models including pack-years and smoking intensity [6,9]. The pack-year approach was traditionally based on a strong assumption about the effects of smoking intensity and duration;
for instance, the effect of smoking $\frac{1}{2}$ pack per day (10 cigarettes) for 40 years was equivalent to smoking 2 packs (40 cigarettes) per day for 10 years. Indeed, both of them reflect a cumulative exposure of 20 pack-years [16]. However, the 2 exposures were measured on different metrics.

Bivariate spline models have the potential to provide a more realistic representation of the association between smoking intensity and duration and HNC risk [15]. First, they allow to model intensity and duration in their original scale of continuous exposures and, in this aspect, splines outperform step-function models. Second, the two exposures are allowed to determine disease risk in their separate or interacting role. Third, a non-linearity in the dose-response relationship may be modeled.

This study re-evaluated the joint effect of intensity and duration of tobacco smoking on HNC risk, by means of bivariate spline models, within the International Head and Neck Cancer Epidemiology (INHANCE) consortium [11].

**Materials and methods**

*Data source*

The INHANCE consortium was established in 2004 to elucidate the etiology of HNC through pooled analyses of individual-level data from several studies on a large scale [11]. Several aspects of tobacco smoking and HNC risk have been previously investigated within the consortium [5,6,11,17-19].

From the INHANCE consortium pooled dataset (version 1.5), we extracted all the available
case-control studies (35 studies) that collected information on cigarette smoking status, intensity, and duration at individual level (http://www.inhance.utah.edu, last accessed March 25th, 2019). Available data were harmonized at the study coordinating center [11]. In all the studies, information on smoking history was self-reported. While different studies had used different definitions of smoking status, the current study defined as never smokers those individuals who never smoked regularly, or smoked for a very short period, i.e., less than a 12 months [11]; likewise, former smokers were defined as those who had abstained from any type of smoking since at least 12 months before cancer diagnosis (cases) or interview (controls).

Details on individual studies, data harmonization, and pooling methods are summarized in eTable 1. Informed consent was obtained from all study subjects. The investigations were approved by the relevant Boards of Ethics, according to the regulation in force at data collection time.

Selection of subjects

The INHANCE protocol allowed inclusion of invasive cancer cases of the oral cavity, oropharynx, hypopharynx, oral cavity or pharynx not otherwise specified, larynx, or unspecified HNC. Cases with cancers of the salivary glands or of the nasal cavity/ear/paranasal sinuses were excluded [11]. The original study sample included 25,865 HNC cases and 37,248 controls, giving a total of 63,113 subjects.

We conducted subjects’ selection according to the following main steps (eAppendix, text and Figure 1), excluding: 1. cases with unspecified (95 subjects) or overlapping HNC (331 subjects);
2. subjects reporting smoking tobacco products other than cigarettes (i.e., cigar, pipe, and cigarillo), to avoid risk distortion due to their use [5] (6,255 subjects); 3. subjects with missing information on duration and/or intensity of cigarette smoking (1,897 subjects); 4. all subjects from studies that included only never (i.e., Japan 1988-2000) (822 subjects) or current (i.e., France 1987-1992) smokers (457 subjects), after the previous selection steps.

In order to prevent potential estimation distortion at the highest levels of the exposure distributions (due to small numbers of subjects or information bias in heavy tobacco consumers), further excluded were those subjects reporting the highest 5% of cigarette smoking intensity (>40 cigarettes per day) or duration (>51 years) (14% HNC cases and 6% controls excluded).

After all the described selection steps (eAppendix, text and Figure 1), the analysis included 33 studies [20-53] with 48,104 subjects (18,260 HNC cases and 29,844 controls) (Table 1). When a study reported to have conducted a case-control matching, separate sets of controls were matched for oral cavity and pharynx (OCP) cancers combined and for laryngeal cancer cases. In detail, the analysis included: 5,423 cancers of the oral cavity; 6,261 pharyngeal cancer cases (4,648 oropharyngeal and 1,613 hypopharyngeal cancers cases); 1,633 unspecified oral cavity/pharynx cancers (giving a total of 13,317 OCP cancer cases combined), and 4,943 laryngeal cancers.

**Statistical analysis**

The bivariate regression spline models [54] used to investigate the dose–response relationship between HNC and the joint exposure to smoking intensity and duration was
described extensively in eAppendix – Statistical Analysis and in Di Credico’s PhD Thesis [55]. We assumed a generalized semi-parametric logistic regression model where the two exposures were entered as a joined piecewise polynomial of a linear degree with constraints for continuity at each join point (called knot), together with potential confounders (i.e., age, sex, race, study, education, drinking status, drinking intensity, and drinking duration) [15]. Knots represented change points in the slope of the risk surface. The set of spline regression parameters described the shape of the risk surface. We further assumed that the knot locations for any of the two exposures were unknown parameters to be estimated, up to a maximum of 2 knots allowed for each exposure. At the maximum level of complexity, the risk surface was allowed to have 2 change-points in the slope for any exposure and 9 different areas with a possibly changing slope.

The optimal knot locations and regression parameters were jointly estimated within the Bayesian approach. Prior distributions expressed a priori knowledge on plausible values of knot locations and regression parameters [56,57]. Their posterior distribution combined prior information and available data through the Bayes theorem. A joint posterior distribution was simulated for each cancer site, smoking-status stratum (current/former smokers), and combination of number of knots using the NUTS (No-U-Turn Sampler) [58] algorithm. For each combination of site and stratum, the best model was selected among the convergent models with sensible knot locations (≤95th percentile of either exposures) as the one that minimized the Wanatabe-Akaike Information Criterion (WAIC) [59,60]. The optimal knot locations, the ORs, and their 95% credible intervals (CIs) were derived from the posterior distribution corresponding to
the best model. The ORs were presented through three-dimensional mesh plots that displayed the surface of risk for any combination of cigarette smoking intensity and duration, and through two-dimensional contour plots that showed iso-risk curves identifying the combinations of the two exposures with the same OR. Also explored were trade-offs between intensity and duration in two-dimensional contour plots by comparison of ORs for a fixed cumulative exposure (i.e., pack-years).

To evaluate potential modifying effects of the covariates, analyses were carried out in strata of alcohol drinking intensity (i.e., <1, 1-<5, ≥5 drinks per day) for current smokers, and years since quitting (i.e., <10, ≥10 years) for former smokers. Separate results for the risk of laryngeal cancer across alcohol drinking intensity strata were presented, due to absence of heterogeneity of the risk surfaces across strata for current smokers. As a comparison, risk estimates were further estimated according to the Bayesian step-function regression model.

When possible, all the models were fitted with the full set of potential confounders shown in Table 1; moreover, “Never smokers” were assumed as the reference category. Calculations were carried out using the open-source Stan program [61] within the open-source R program [62,63].

Results

Table 1 shows selected characteristics of cases and controls, according to the variables included in the models as potential confounders. Approximately 70% of the subjects were white. Studies from Europe contributed with approximately 44% of subjects; 31% of subjects were from
the United States, whereas the remaining ones were from Latin America (14%) and Asia (11%).

Six studies provided only cases of OCP cancer.

Table 2 shows the distribution of cigarette smoking habits for OCP cancer, laryngeal cancer, and controls. Never smokers were 21% of OCP and 7% of laryngeal cancers, versus 45% of controls. Current smokers were 66% of laryngeal cancer cases, 57% of OCP cancer cases, and 26% of controls. The prevalence of former smokers was similar in cases and controls, but the percentage of subjects who quit cigarette smoking ≥10 years ago was ~52% among HNC cases and 72% among controls.

Figure 2 shows the mesh and contour plots for cancers of the OCP and larynx among current cigarette smokers. Knot locations were indicated by thicker black lines; the intersections of the lines defined the areas where the risk surface had different slopes. For both cancer sites, the best model was characterized by one knot for duration (at 33 years for OCP cancer and 30 years for laryngeal cancer) and one knot for intensity (at 16 and 25 cigarettes/day, respectively). Smoking duration modified OCP cancer risk at any levels of intensity (Figure 2A). In contrast, for any duration up to 10 years, the ORs were always <2, regardless of the intensity of cigarette smoking. In addition, at a fixed value of 20 pack-years, an intensity of 40 cigarettes/day and a duration of 10 years led to an OR of ~2, whereas the OR was equal to ~4 with a duration of 40 years in smokers of 10 cigarettes/day.

For laryngeal cancer, ORs>20 were found for intensities of >20 cigarettes/day and durations of >28 years (Figure 2B). Moreover, ORs>10 of laryngeal cancer were reached by current
smokers of >20 cigarettes/day only when duration was >20 years; however, ORs>10 were not reached for any duration <15 years and any level of intensity. Finally, the OR was 6.2 for smokers of 40 cigarettes/day for 10 years, but it was higher (between 9 and 10) for 20 cigarettes/day smoked for 20 years, or for 10 cigarettes/day smoked for 40 years.

The shape of the risk surface and the values of the ORs were also very similar across major OCP cancer subsites (i.e., oral cavity and oropharyngeal cancers) (eFigure 2).

Figure 3 shows the joint effect of current smoking intensity and duration in strata of alcohol consumption. Among never drinkers (<1 drink/day), the shape of the risk surface was similar to the one presented for all alcohol intensities together (Figure 2A), but the ORs were generally lower; all the ORs were < 2 for durations ≤15 years and any intensity, whereas an OR>5 was observed only after ~25 years of duration or more (Figure 3A). However, the shape of the surface and/or the ORs of the joint effect of duration and intensity were different when light (Figure 3B) and heavy (Figure 3C) drinkers were considered.

As a comparison, the ORs and their corresponding 95% CIs were estimated for each cancer site in current smokers within the Bayesian logistic regression model that assumed the presence of step functions (Table 3, main ORs). These estimates were compared with the range of OR estimates derived from the spline models for each joint category of duration and intensity (Table 3, bracketed ORs). For both cancer sites, all Min-Max ranges included the OR estimates obtained from the Bayesian logistic regression, and this replication reassured that the spline model was valid. In addition, within the examined categories, the step-function intervals widely
overlapped with the Min-Max ranges of the ORs from the spline models, but failed to capture the ORs variability. For instance, for OCP cancer in current smokers of 26-40 cigarettes/day for 36-51 years, the categorical OR was 8.4 (95% CI: 8.0-8.9), whereas the OR varied from 7.1 to 10.6 under the spline model approach for the same combined category (Table 3). The same pattern emerged at any combination of duration and intensity for both cancer sites for former smokers who quit ≥10 years ago (eTable 2).

Among former smokers who quit ≥10 years ago, no ORs>4 were observed for OCP cancer (Figure 4A), and the ORs were approximately halved, as compared to the same levels of duration and intensity in current smokers (Figure 2A). A mean risk reduction of 2/3 was also found for laryngeal cancer in long-term former smokers, with ORs declines that varied from 1/3 to 4/5 depending on the different combinations of intensity and duration (Figure 4B). These estimates were derived from models that included one knot for intensity (27 cigarettes/day) for laryngeal cancer and no knots for OCP cancer. Results on the combined set of former smokers (regardless of time since quitting) were similarly presented in eFigure 3.

Discussion

Our large pooled analysis showed that cigarette smoking duration and intensity did not increase HNC risk to the same extent, but the effect was greater for longer durations. At the highest combined levels of cigarette intensity and duration, subjects may reach ORs>10 of OCP cancer and ORs>40 of laryngeal cancer, as compared to never smokers. Bivariate regression
spline models have been proven successful in exploring the separate and joint effects of intensity and duration of cigarette smoking. Therefore, when possible, models that allow this differential impact of intensity and duration on the risk should be considered [8,12,16].

Results from the present study were consistent with previous findings showing no threshold effect on HNC risk of shorter durations or lower intensities of tobacco smoking [13,64]. Indeed, the literature supported the presence of a dose-response relationship between cigarette smoking and HNC risk over the entire range of consumption [5,13]. Our contribution additionally suggested that the dose-response relationship with OCP and laryngeal cancer risk was still far from being linear, with a steeper increase for intermediate consumptions and a possible plateaux indicating a ‘saturation effect’ in smokers with >20 years of duration and >30 cigarettes/day.

Our results also strengthen the evidence [18,65] on the impact of smoking cessation on HNC cancer risk, with the OR in former smokers who quit ≥10 years ago being more than halved for OCP cancer and ~1/3 lower for laryngeal cancer, in comparison with current smokers. These results convey to the general population and to public health professionals the valuable message that it is never too late to give up cigarette smoking, as quitting provides far smaller risks of developing HNC cancer for OCP and, most of all, for laryngeal cancers. This is in accordance with another recently published meta-analysis of observational studies on laryngeal cancer based on 14,292 cases from 3 cohort and 15 case-control studies [65].

When the combined effect of smoking intensity and duration was modelled as a cross-product term, an interaction term between alcohol consumption and cigarette smoking was traditionally
added to the models [3,4]. In our study this effect was explored by stratified regression spline models estimated within categories of alcohol consumption; we compared the shape of the risk surface and the ORs values from the iso-risk curves across the three stratified analyses. However, our results also showed that alcohol acts as a substantial modifier of the association between OCP cancer risk and the joint effect of cigarette intensity and duration [4,17].

Major strengths of our study included a large sample size, international representativeness, and valid information on potential confounding factors. In addition, we applied a novel Bayesian approach to jointly estimate the optimal knot locations and the ORs of HNC for the joint effect of intensity and duration in a bivariate context. After examining various frequentist solutions [66,67], we opted for the Bayesian approach that allowed to put the knots where the data suggested, once we had taken the entire set of confounding variables into consideration. Within the Bayesian framework, we were also able to choose the optimal number of knots (up to 2) and knot locations by model comparison based on information criteria, instead of just comparing models with fixed number and location of the knots. Finally, we could incorporate in the model selection process evidence from the literature on the maximum number of changes in the risk pattern, which is likely to be 2. Indeed, cigarette smoking is supposed to have a protective or null effect on the risk at lower levels, a saturation effect at the highest intensity levels [68], as well as the expected increase previously described at the intermediate levels of consumption. The proposed Bayesian approach is applicable to other epidemiologic scenarios where continuous exposures in their potential interaction affect disease risk.
Among study limitations, the retrospective study design and the self-reported smoking history were the most relevant ones. Even if a large amount of literature has suggested acceptable correlations between self-reported smoking intensity and cotinine levels in blood or urine [69], inaccurate self-reporting may occur. Discrepancies between self-reported and objective information were more likely among long-term heavy smokers [70]; higher values of intensity and duration were therefore more prone to inaccurate reporting. Furthermore, smoking intensity may have varied over time and by age of exposure. However, its estimates were often based on the self-reported average number of cigarettes per day; these two aspects may have led to appreciable errors in measuring the true mean intensity of exposure over one’s lifetime. These issues may well be serious here, given the continuous nature of the exposures considered in spline models. To reduce information bias and residual confounding at the extreme values of the exposure distributions, we excluded subjects reporting higher (>95th percentiles) cigarette intensity and/or duration from the present analysis [13,15,71]. In addition, to avoid bias due to the use of other tobacco products, we excluded subjects reporting use of tobacco products other than cigarettes. We were also obliged to combine different subsites of OCP cancer, although differences in aetiology (i.e., the causal role of human papillomavirus in oropharyngeal cancer) have been demonstrated [34]. However, results were similar for major subsites, although based on less stable models. Finally, our Bayesian approach was computationally time consuming, asking for several hours of server computing for each model fitted.
Results from the present study confirmed previous evidence on the main contribution of cigarette smoking duration in influencing HNC risk and on the important reduction in HNC risk observed in former smokers who quit ≥10 or 15 years ago. Both results strongly support an even wider diffusion of public interventions to encourage smokers to quit and help them succeed.

Acknowledgements: The authors would like to thank Xavier Castellsagué who collected data in the IARC International Multicenter study and passed away in 2016. We thank Mrs Luigina Mei for editorial assistance.

Fundings: This work was supported by grants from the: National Institutes of Health (NIH) [no grant number provided for the INHANCE Pooled Data Project, grant numbers P01CA068384, K07CA104231 for the New York Multicenter study, grant numbers R01CA048996, R01DE012609 for the Seattle (1985-1995) study, grant number TW001500 for the Fogarty International Research Collaboration Award (FIRCA) supporting the Iowa study, grant number R01CA061188 for the North Carolina (1994-1997) study, grant numbers P01CA068384, K07CA104231, R01DE013158 for the Tampa study, grant numbers P50CA090388, R01DA011386, R03CA077954, T32CA009142, U01CA096134, R21ES011667 for the Los Angeles study, grant numbers R01ES011740, R01CA100264 for the Houston study, grant numbers R01CA078609, R01CA100679 for the Boston study, grant number R01CA051845 for the MSKCC study, grant number R01CA030022 for the Seattle-Leo study, grant number DE016631 for the Baltimore study]; National Cancer Institute (NCI) at the National Institutes of
Health (NIH) [grant number R03CA113157 for the INHANCE Pooled Data Project, no grant number provided for the Intramural Programs supporting the Puerto Rico and the US Multicenter studies, grant number R01CA90731-01 for the North Carolina (2002-2006) study]; National Institute of Dental and Craniofacial Research (NIDCR) at the National Institutes of Health (NIH) [grant number R03DE016611 for the INHANCE Pooled Data Project, grant numbers R01DE011979, R01DE013110 for the Iowa study, no grant number provided for the Intramural Program supporting the Puerto Rico study]; Italian Association for Research on Cancer (AIRC) [no grant number provided for the Milan (1984-1989) study, for the Aviano study, for the Italy Multicenter study and for the Rome study, grant number 10068 for the Milan study (2006-2009)]; Italian League against Cancer [no grant number provided for the Aviano and Italy Multicenter studies]; Italian Ministry of Research [no grant number provided for the Aviano and Italy Multicenter studies]; the Swiss Research against cancer/Oncosuisse [grant numbers KFS-700, OCS-1633 for the Swiss study]; World Cancer Research Fund [no grant number provided for the Central Europe study]; European Commission [grant number IC18-CT97-0222 (INCO-DC Program) for the Latin America study, grant number IC15-CT98-0332 (INCO-COPERNICUS Program) for the Central Europe study]; Veterans Affairs Merit Review Funds [no grant number provided for the Iowa study]; National Institute of Environmental Health Sciences (NIEHS) [grant number P30ES010126 for the North Carolina (1994-1997) study, grant number P30ES010126 for the North Carolina (2002-2006) study]; Alper Research Program for Environmental Genomics of the UCLA Jonsson Comprehensive Cancer Center [no grant number provided for the Los
Angeles study]; Fondo para la Investigacion Cientifica y Tecnologica Argentina (FONCYT) [no grant number provided for the Latin America study]; Institut Hospital del Mar d'Investigacions Mediquès (IMIM) [no grant number provided for the Latin America study]; Fundação de Amparo à Pesquisa no Estado de São Paulo (FAPESP) [grant number 01/01768-2 for the Latin America study, grant numbers GENCAPO 04/12054-9, 10/51168-0 for the Sao Paulo study]; Fondo de Investigaciones Sanitarias (FIS) of the Spanish Government [grant number FIS 97/0024, FIS 97/0662, BAE 01/5013 for the International Multicenter study]; International Union Against Cancer (UICC) [no grant number provided for the International Multicenter study]; Yamagiwa-Yoshida Memorial International Cancer Study Grant [no grant number provided for the International Multicenter study]; European Community (5th Framework Programme) [grant number QLK1-CT-2001-00182 for the Western Europe study]; Ministry of Science, Research and Arts Baden-Wurttemberg [no grant number provided for the Germany-Saarland study]; German Ministry of Education and Research [grant number 01GB9702/3 for the Germany-Heidelberg study]; Scientific Research grant from the Ministry of Education, Science, Sports, Culture and Technology of Japan [grant number 17015052 for the Japan (2001-2005) study]; Third-Term Comprehensive 10-Year Strategy for Cancer Control from the Ministry of Health, Labor and Welfare of Japan [grant number H20-002 for the Japan (2001-2005) study]; Johns Hopkins Richard Gelb Cancer Prevention Award [no grant number provided for the HOTSPOT study]; Italian Foundation for Cancer Research (FIRC) [no grant number provided for the Milan study (2006-2009)]; Italian Ministry of Education - PRIN 2009 Program [grant number X8YCBN for the
Milan study (2006-2009)]; VE was supported by Università degli Studi di Milano ‘Young Investigator Grant Program 2017’.
References


[24] Hashibe M, Boffetta P, Zaridze D. Evidence for an important role of alcohol- and aldehyde-


[54] Ruppert D, Wand M, Carroll R. Semiparametric Regression (Cambridge Series in Statistical

Ph.D. Thesis. Padua University: Italy; 2018

[56] Gelman A, Stern HS, Carlin JB, Dunson DB, Vehtari A, Rubin DB. Bayesian Data

[57] Gelman A, Jakulin A, Pittau MG, Su YS. A weakly informative default prior distribution for

[58] Hoffman MD, Gelman A. The No-U-turn sampler: adaptively setting path lengths in

[59] Watanabe S. Asymptotic equivalence of Bayes cross validation and widely applicable

[60] Gelman A, Hwang J, Vehtari, A. Understanding predictive information criteria for Bayesian


[63] R Development Core Team. R: A language and environment for statistical


Figure captions

Figure 1 - Flow chart of subjects' selection

Figure 2 - Odds ratios\textsuperscript{a,b} of oral and pharyngeal cancer and laryngeal cancer in current smokers, for the joint effect of intensity (cigarettes/day) and duration (years) of cigarette smoking estimated through bivariate spline models. INHANCE consortium

\textsuperscript{a}Fitted models included adjustment for age, sex, race, study, education, drinking status, drinking intensity, and drinking duration. The reference category was defined as “Never smokers”. \textsuperscript{b}On the grid, black thicker lines represent knot locations: 16 cigarettes/day and 33 years of duration for oral and pharyngeal cancer and 25 cigarettes/day and 30 years of duration for laryngeal cancer, respectively. Dark grey lines in contour plots indicate iso-risk curves at defined levels of risk.

Figure 3 - Odds ratios\textsuperscript{a,b} of oral and pharyngeal cancer in current smokers by alcohol drinking intensity, for the joint effect of intensity (cigarettes/day) and duration (years) of cigarette smoking, estimated through bivariate spline models. INHANCE consortium

\textsuperscript{a}Fitted models included adjustment for age, sex, race, study, education, and alcohol drinking status. The reference category was defined as “Never smokers”, in each strata of alcohol
drinking intensity. On the grid, black thicker lines represent knot locations: 32 years of duration for never drinkers, and 12 cigarettes/day and 25 years of duration for heavy drinkers, respectively. Dark grey lines in contour plots indicate iso-risk curves at defined levels of risk.
Figure 4 - Odds ratio\textsuperscript{a,b} of oral and pharyngeal cancer and laryngeal cancer in former smokers who quit ≥10 years ago, for the joint effect of intensity (cigarettes/day) and duration (years) of cigarette smoking estimated through bivariate spline models.

INHANCE consortium

\textsuperscript{a} Fitted models included adjustment for age, sex, race, study, education, drinking status, drinking intensity, and drinking duration. The reference category was defined as “Never smokers”. \textsuperscript{b} On the grid, the black thicker line represents the knot location: 27 cigarettes/day for laryngeal cancer. Dark grey lines in contour plots indicate iso-risk curves at defined levels of risk.