Community Dental Health (2020) **37,** 96–101

doi:10.1922/CDH_SpecialIssueNicolau06

Quantitative Bias Analysis of misclassification in case-control studies: an example with Human Papillomavirus and Oropharyngeal Cancer

Doris Durán[1,4], Walid A. Al-Soneidar[2], Sreenath A. Madathil[3], Jay S. Kaufman[2] and Belinda Nicolau[3]

[1]*Escuela de Salud Pública, Facultad de Medicina, Universidad de Chile.* [2]*Department of Epidemiology, Biostatistics and Occupational Health, Faculty of Medicine, McGill University.* [3]*Faculty of Dentistry, McGill University.* [4]*Facultad de Medicina y Ciencia. Universidad San Sebastián.*

Objective: Laprise et al. (2019) observed a positive association between oral sex practices and oropharyngeal cancers (OPC) among HPV-negative individuals. Because oral HPV infections are likely to be transmitted through oral sex, these results are counterintuitive. We revisit Laprise et. al's analysis with the objective of estimating the impact of misclassification of HPV infection on the association between oral sex practices and OPC. ***Methods***: Data were drawn from the Head and Neck Cancer (HeNCe) Life study, a hospital-based case control study of head and neck cancer with frequency-matched controls by age and sex from 4 major referral hospitals in Montreal, Canada. We included only OPC cases (n = 188) and controls (n = 429) and used predictive value weighting, under differential and non-differential scenarios, to evaluate the misclassification. Subsequently, we used logistic regression and 95% confidence intervals to estimate the association between oral sex practice and OPC among HPV-negative individuals. ***Results***: Our results showed that the previously reported association between oral sex practices and OPC among HPV-negative individuals was attenuated or nullified both under differential and non-differential scenarios. ***Conclusion***: The association between oral sex practice and OPC could be explained by biases in the data (e.g., HPV mediator misclassification). Our results highlight the need for widespread adoption of Quantitative Bias Analysis in oral health research.

Keywords: *Head and neck cancer; HPV infection; Misclassification*

Introduction

Measurement error and misclassification of exposures are common problems in retrospective observational studies leading to biased results (Jurek *et al*., 2006). Several statistical techniques, known as Quantitative Bias Analyses (QBA), have been developed to estimate the magnitude and direction of this bias under different scenarios (Greenland, 1996, 2014; Lash *et al*., 2009; Lash *et al*., 2014; Lash and Schmidt, 2010; Orsini, 2007). In studies in which the exposure is measured using a diagnostic test, the sensitivity (Se) and specificity (Sp) of the test can significantly impact the level of misclassification and the resultant bias (Lash *et al*., 2009, 2014). In case-control studies, if the test performs equally well (same level of Se and Sp) in both cases and controls, bias in the association due to exposure misclassification will usually be towards the null; this is referred as "non-differential misclassification" (Hernán and Robins, 2020; Lash *et al*., 2009). However, if the diagnostic test performs differently (different levels of Se and Sp) in cases compared to controls ("differential misclassification"), it is often difficult to anticipate whether the bias over- or under-estimates the true association (Greenland, 1996; Lash *et al*., 2009). Furthermore, if a potential mediator in the exposure-outcome causal pathway is misclassified, it may lead to differential biases in stratum specific association estimates (Valeri and Vanderweele, 2014).

The QBA literature has increased dramatically in the last decade, providing numerous examples of how the estimates of effect measures change when accounting for various types of systematic errors (Alasbahi *et al*., 2014; Blakely *et al*., 2019; Rumball-Smith *et al*., 2013; Stott-Miller *et al*., 2010). Indeed, QBA has been suggested as an essential step not only in producing valid statistical inference from observational studies, but also to review the evidence aiming to inform public health decisions (Fox and Lash, 2017). Unfortunately, the oral health literature is lagging behind in the adoption of formal QBA. This manuscript portrays the application of QBA to correct for bias due to misclassification of a mediator (HPV) in the stratum specific association between the exposure (oral sex practices) and the outcome (oropharyngeal cancer (OPC)).

Oral sex, Human papillomavirus infection and risk of oropharyngeal cancers

The rapid increase in the incidence of OPC in North America has motivated researchers to report it as epidemic (Tota *et al*., 2019). There were over 90,000 new cases of OPC and 50,000 deaths in 2018, worldwide (The Global

Correspondence to: Belinda Nicolau, Faculty of Dentistry, McGill University, 2001 McGill College Avenue, Suite 527, Montreal, QC, Canada, H3A 1G1. Email. Belinda.f.nicolau@mcgill.ca Phone: xx1- 514-398-7203 ext 094655

Cancer Observatory, 2019). OPC are complex and expensive diseases to treat, lead to major impairment and deformity, and have major consequences for individuals' quality of life (Warnakulasuriya, 2010). The main risk factors are alcohol and tobacco consumption and their joint effect. In addition, oral HPV infections, now recognized as a major OPC risk factor, have modified the epidemiological profile of these neoplasms, in that they mainly affect socially advantaged young men (Ligier *et al*., 2011; Rettig and D'Souza, 2015).

HPVs are circular, double-stranded DNA viruses (de Villiers *et al*., 2004) that are likely to be sexually transmitted and have been associated with risky sexual behaviors (e.g., high number of sexual partners). Specifically, oral HPV infections are understood to be transmitted through oro-genital contact (e.g., practice of oral sex) (Schlecht *et al*., 2019). However, empirical evidence for the role of HPV mediating the association between oral sex and oropharyngeal cancer is still lacking. Laprise et al. (2019) investigated the association between oral sex practice and OPC, and considered HPV infection as a potential mediator to explain the relationship. Figure 1 mimics the directed acyclic graph (DAG) used by the authors.

The causal pathway described in the DAG is based on the evidence that the only biological mechanism by which oral sex practices are associated with OPC risk is through HPV infection. Under this assumption, the estimates for the association between oral sex practices and OPC risk, within a stratum defined by HPV status, is expected to be near null. However, the authors observed a positive association between oral sex practices and OPC risk among HPV-negative individuals. Because oral HPV infections are likely to be transmitted through oral sex, these results are counterintuitive. Laprise and colleagues attributed this finding to be partly due to potential misclassification of oral HPV status.

In their study, HPV genotyping was performed by generic PCR-ELISA using exfoliated cells from the oral cavity. Although HPV infections can be detected in exfoliated oral cells from oral rinse and swabs, this method does not indicate if HPV is present at the tumor site nor does it identify the specific HPV type colonizing the lesion. Therefore, there is a possibility of misclassification. Although there is no gold-standard technique for oral HPV detection, there have been attempts in the literature to summarize and describe HPV detection methods for the oral cavity (Gipson *et al*., 2018). We revisit the Laprise analysis with the objective of estimating the impact of misclassification of HPV infection on the association between oral sex practices and OPC as a methodological exercise.

Methods

The detailed study design and description of the data have been reported elsewhere (Laprise *et al*., 2019). Briefly, data were obtained from the Head and Neck Cancer (HeNCe) Life-study, a hospital-based case-control study conducted in 4 main referral hospitals in Montreal, Canada, between 2005 and 2013. All participants were born in Canada and lived within a 50 km area of the hospital from which they were recruited. For this analysis, we included all incident squamous cell carcinoma of the oropharynx (C01, C02.4, C05.01, C05.2, C09, C10, C12 and, C14 [base of the tongue, soft palate, tonsil, oropharynx, and uvula]). Controls, frequency matched to cases by age (5-year categories) and sex, were randomly recruited from several outpatient clinics at the same hospitals as the cases. Information on socio-demographic, occupational, behavioral, including oral sex practice (e.g., frequency of oral sex, age at the start of oral sex practices and number of oral sex partners) factors were collected using life-grid based in-person interviews.

Analysis

We emulated the analysis from Laprise et al. with a subsequent correction for misclassification in HPV status. The original analysis used unconditional logistic regression modelling for OPC, where reported years since first practice of oral sex was considered as the main exposure. Although other variables (e.g., frequency of oral sex) are available in this dataset, we chose years since first practice of oral sex. The reason for this choice was the strongest association of this variable with OPC among HPV negative reported by Laprise *et al* 2019. Adjustments included smoking, drinking, education and lifetime

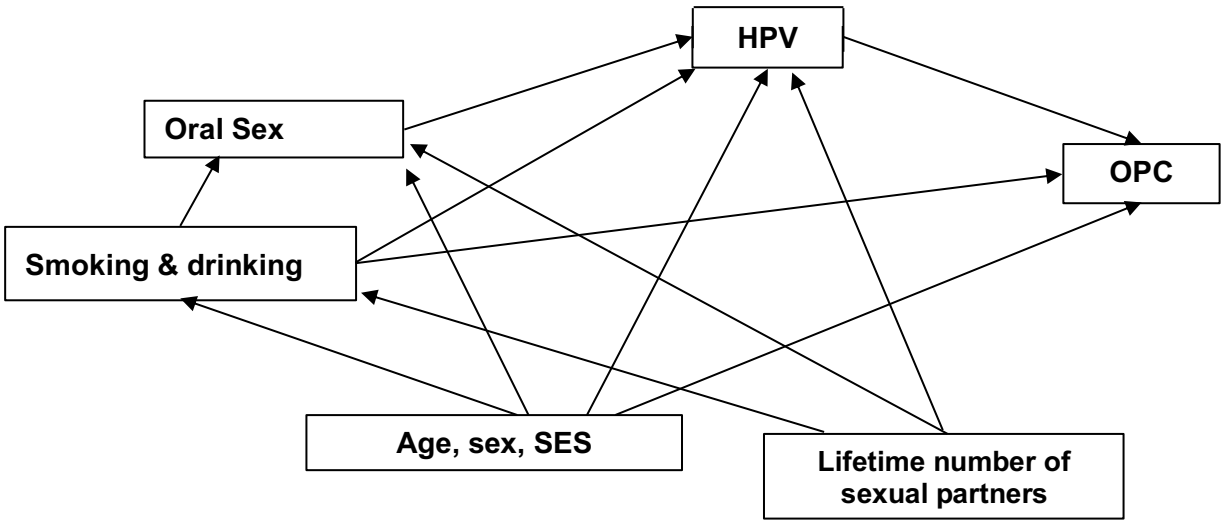


Figure 1. *Directed acyclic graph for the relation between Oral sex, HPV-infection and Oropharyngeal cancer.*

number of sexual partners as covariates and confounders. Then, two models were fitted with stratification on the observed HPV status (positive/negative) to evaluate the mediation effect of this variable in the oral sex practice /OPC relationship.

Quantitative Bias Analysis

For simplicity reasons, we consider only the misclassification in the HPV variable (mediator) in this manuscript. Figure 2 shows how HPV misclassification can be non-differential (A), or differential (B). The measured HPV status (HPV_{obs}) of the participant is assumed to be influenced by their true unknown HPV status (HPV_{true}) and other unknown variables, which may have influenced the measurement of HPV (U). Figures 2A and 2B depict the non-differential and differential misclassification scenarios, respectively.

We corrected for HPV misclassification using the predictive value weighting (PVW) approach described by Lyles and colleagues (2010). Briefly, the approach uses information on the observed value of the variable in the data and the positive and negative predictive values for the diagnostic test, which is obtained externally, to create predictive weights. Subsequently, the main analysis is performed in the re-weighted data, corrected for misclassification.

Because our goal was to quantify the amount of bias in the association between oral sex practice and OPC among individuals testing negative for HPV, we stratified the analysis using observed values of HPV. Confidence intervals for the estimates were obtained through bootstrapping (200 repetitions) and a sample size equal to the whole study.

We obtained the Se and Sp values from a systematic review and meta-analysis describing the accuracy of HPV detection methods for head and neck cancers including OPC (Gipson *et al*., 2018). Two types of misclassification scenarios were assessed.

Non-differential misclassification scenarios:

In this assumption, cases and controls were given the same Se and Sp. We explored two levels of misclassification: first, we used the point estimate of the meta-analysis (Se: 0.72 and Sp: 0.92); and second, we applied a more optimistic estimate, using the upper confidence limit (Se: 0.89 and Sp: 0.97) of the pooled estimate from meta-analysis.

Differential misclassification scenarios:

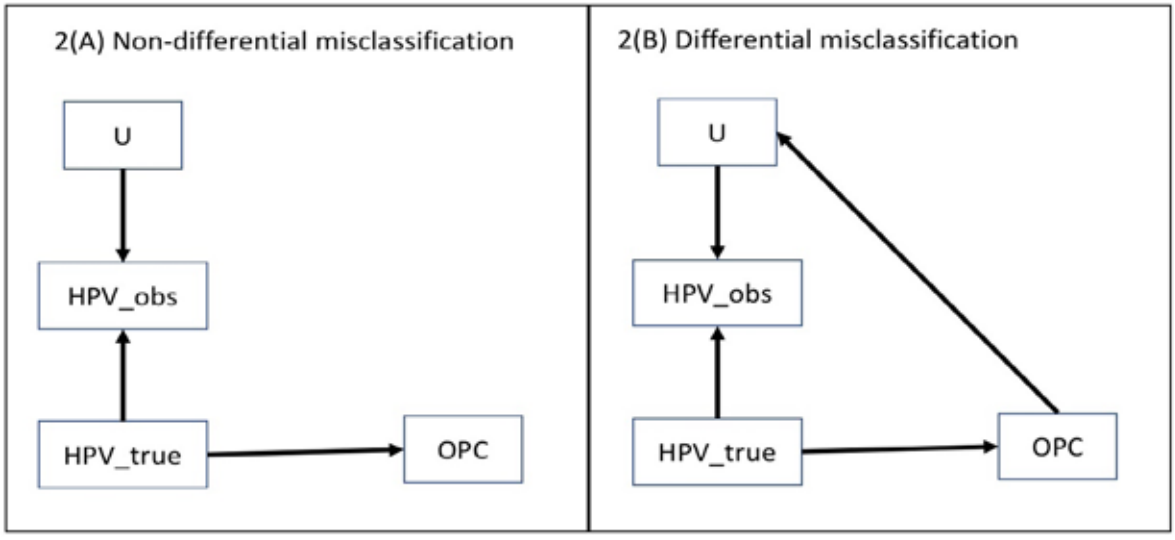
Two differential misclassification levels were explored: i) cases have higher Se compared to controls with same level of Sp; ii) cases have higher Se and Sp compared to controls. In scenarios with higher Se or higher Sp, we used corresponding upper confidence limit estimates from the meta-analysis and point estimates for all other parameters.

All analyses were conducted in Stata 15 (StataCorp., 2017), with the *pvw* package (Bartlett, 2019).

Results

The sample included 617 individuals, 188 cases and 429 controls. Both cases and controls were mostly men (72.7% for cases and 69.2% for controls). Oral sex practices were higher in cases: 49.5% had more than 8 sexual partners in their life; 68.6% had experienced their first oral sex experience more than 30 years ago (Table 1). HPV prevalence was more common among the cases (62.3%).

When emulating the previously reported analysis, the odds of developing OPC were 3.14 (CI 95% 1.80; 5.48) times higher among those reporting first oral sex experience more than 30 years ago, compared to those who reported experiencing oral sex for the first time less than 15 years ago. However, this association decreased to 1.20 (CI95% 0.37; 3.85) in HPV positive participants



HPV_{true}: True HPV status; OPC: True OPC status; HPV_{obs}: observed HPV; U: Unknown factors affecting misclassification

Figure 2. Independent misclassification of predictor (HPV)

Table 1. Sample characteristics

	Cases		*Controls*	
	N=188	*%*	*N=429*	*%*
Sex				
Male	137	72.7	297	69.2
Age	Mean (sd)		Mean (sd)	
	59.9 (9,47)		61.1 (10,9)	
Smoking				
Never	43	22.9	111	25.9
Yes	36	19.2	101	23.5
Past-smoker	109	58.0	217	50.6
Alcohol consumption				
Never	32	17.0	73	17.0
Yes	97	51.6	291	67.8
Past-consumption	59	31.4	65	15.2
Lifetime sexual partners				
Less than 2	35	18.6	126	29.4
Between 3 to 7	53	28.2	133	31.0
More than 8	93	49.5	155	36.1
NR	7	3.7	15	3.5
Time since first oral sex practice				
Less than 15 years	20	10.6	106	24.7
15 to 30 years	24	12.8	70	16.3
More than 30 years	129	68.6	226	52.7
NR	15	8.0	27	6.3
Frequency of oral sex practices				
Up to 16 years of age				
Occasionally	14	7.4	29	6.8
Frequently	4	2.1	4	0.9
Most of the time	7	3.7	5	1.2
Don't know-Prefer not to say	13	6.9	16	3.7
NR	150	79.8	375	87.4
From 17 to 30 years of age				
Occasionally	89	47.3	161	37,5
Frequently	42	22.3	76	17.7
Most of the time	17	9.0	44	10.3
Don't know-Prefer not to say	13	6.9	17	4.0
NR	27	14.4	131	31.5
After 30 years of age				
Occasionally	89	47.3	176	41.0
Frequently	46	24.5	85	19.8
Most of the time	23	12.2	55	12.8
Don't know-Prefer not to say	13	9.0	17	4.0
NR	17	6.9	96	22.4
HPV prevalence (Any type)				
Positive	119	62.3	61	14.2
Negative	69	36.7	368	85.8

suggesting a mediating role of HPV in this association. However, a strong association was observed in the HPV negative stratum (Table 2).

In every scenario of the QBA for the HPV-negative group (Table 3), the OR for oral sex practices decreased towards the null. Moreover, in the non-differential scenarios with the lowest sensitivity estimates, the direction of association changed. Furthermore, the association between oral sex practice and OPC in the HPV-negative stratum was attenuated (OR 1.90; 95% CI 0.71, 5.88).

Differential misclassification corrections yielded similar results, with the association decreasing toward the null in every case.

Discussion

In this manuscript, we demonstrate an application of QBA using information from a meta-analysis on a diagnostic test in case-control studies. Our findings highlight the significance of QBA to critically interpret research results under external evidence and optimistic assumptions. Furthermore, we contribute to the field by extending the analysis to the scenario of misclassification in a mediator.

Laprise et al. reported an unexpected association between oral sex practice and OPC among HPV-negative participants. The authors recognized the biological implausibility of this association and argued these findings could be explained by biases in data. In our analysis, we chose to explore one possible bias mechanism, which was the role of misclassification of the mediator (HPV status). Using Se and Sp values from a systematic review and metanalysis, we showed how the association between oral sex practice and OPC among the HPV-negative group could be attenuated after HPV classification correction. This attenuation happened under both non-differential and differential misclassification assumptions.

Although we have not considered misclassification in any other variables in the current analysis, this possibility cannot be excluded. When multiple variables are misclassified those errors could also be correlated resulting in further bias. One such misclassification of particular importance is misclassification in the outcome variable. Misclassification in both exposure and outcome can present in four different forms: dependent-differential, dependent-nondifferential, independent-differential, and independent-nondifferential (Hernán and Robins, 2020). In our analysis, because outcome ascertainment (OPC) is based on histopathological examination, there could not have been any dependence between HPV misclassification and OPC status.

We used a sample size equal to the whole study to obtain bootstrapped 95% confidence intervals. Although this results in an overestimation of precision, in QBA the quantity of interest is the change in point estimate and not change in confidence limits. Furthermore, we tested the extent of bias using different estimates for Se and Sp. Our results indicate that even correction for highly optimistic Se and Sp did not completely reduce the OR to null, thereby indicating residual bias due to an unmeasured/unaccounted confounder(s).

Misclassification of the mediator is not the only possible source of bias. Another possible explanation is

Table 2. Emulated original and stratified analysis for the relationship between oral sex practice and OPC

			Stratified by HPV positivity			
	Null		*HPV+*		*HPV-*	
	Model 1 [a]		*Model 2* [a]		*Model 3* [a]	
Variable	*OR*	*CI 95%*	*OR*	*CI 95%*	*OR*	*CI*
Time since first oral sex practice						
less than 15 years	ref		ref		ref	
15 to 30 years	1.63	(0.77; 3.43)	0.60	(0.13; 2.73)	1.28	(0.44; 3.71)
More than 30 years	3.14	(1.80; 5.48)	1.20	(0.37; 3.85)	2.35	(1.04; 5.30)

[a] All models adjusted by age, sex, smoking, alcohol and lifetime number of sexual partners

Table 3. Comparison of the logistic models with misclassification corrections

	Non-Differential ⊥				*Differential* ⊥			
	Se:0.72 Sp:0.92		*Se:0.89 Sp:0.97*		*SeCa:0.89 SeCo:0.72 SpCa0.92 SpCo: 0.92*		*SeCa: 0.89 SeCo: 0.72 SpCa: 0.97 SpCo:0.92*	
	Correction 1		*Correction 2*		*Correction 3*		*Correction 4*	
Variable	*OR*	*CI*[b]	*OR*	*CI*[b]	*OR*	*CI*[b]	*OR*	*CI*[b]
Time since first oral sex practice								
Less than 15 years	Ref		ref		ref		ref	
15 to 30 years	0.62	(0.13; 2.83)	1.09	(0.32; 3.78)	1.09	(0.32; 3.70)	1.08	(0.32; 3.66)
More than 30 years	0.88	(0.27; 2.91)	1.92	(0.70; 5.29)	1.90	(0.67; 5.37)	1.89	(0.68; 5.26)

[a] All models adjusted by age, sex, smoking, alcohol and lifetime number of sexual partners
[b] Confidence intervals obtained using bootstrap for the original sample size (n=617) if all participants were HPV-.
Se: Sensitivity; Sp: Specificity; Ca: Cases; Co: Controls
⊥ After correction, estimation was conducted only in HPV negative participants

measurement error in oral sex practices. The sexual behavior literature shows that self-reported sexual practices, even in the near past, can be subjected to self-disclosure and recall bias, which are extremely difficult to validate (Graham *et al*., 2010).

However, the HeNCe Life study used experienced interviewers and a life-grid technique to collect retrospective data which has been shown to reduce recall bias. A third possible source of bias is the introduction of collider bias in the case of differential misclassification, when stratified based on the observed value of HPV. However, our goal was to quantify the bias in a previous analysis which followed the same approach.

Our study illustrates how estimates of Se and Sp of a diagnostic test from a meta-analysis can be used to quantify bias due to misclassification. There is no gold standard for measurement of HPV in saliva samples, with the American Dental Association only recommending it for research purposes (Lingen *et al*., 2017). Hence, an estimate from meta-analysis is the next best potential source of evidence for Se and Sp to detect HPV infection. The meta-analysis by Gipson et al. showed high variability in primary studies, further reflecting the need for QBA in studies which use these techniques.

The above-mentioned factors make it impossible to completely disentangle the various sources of bias present. However, our analysis shows how HPV misclassification could be, among others, a potential explanation for biased results in the relationship between oral sex practice and OPC. These findings support the original hypothesis proposed by Laprise et al. 2019 on the mediating role of HPV in the relationship between oral sex and oropharyngeal cancer. In conclusion, our work highlights the need for widespread adoption of QBA in oral health research, critical use of oral health evidence is needed and a QBA is an useful tool to assess research results.

Acknowledgments

We acknowledge the contribution of HeNCe Life Study participants. We thank Claudie Laprise for providing syntax for the original analysis. We also would like to thank staff at the Department of Otolaryngology — Head and Neck Surgery as well as the control clinics at the Notre-Dame, Jewish General, Royal Victoria and Montreal General Hospitals for their support in the recruitment of study participants.

Financial Support

Emerging leaders of the Americas Scholarship Program (ELAP), Global Affairs Canada International Scholarships Program. Awarded to DD to conduct research at McGill University during 2019. BN holds a Canada Research Chair in life course oral epidemiology. The HeNCe Life study was funded by CIHR grant numbers: MOP-69062, MOP-201009, DCA-122700 and the Ministere du Developpement economique, de l'Innovation et de l'Exportation du Quebec.

References

Alasbahi, R. H., Melzig, M. F., Andreux, P. A, Houtkooper, R. H., Auwerx, J., Anne Stetler, R. and Zorzano, A. (2014): Clinical therapeutic strategies for early stage of diabetic kidney disease. *Nature Reviews. Endocrinology* **9**, 1–13.

Bartlett, J. (2019): PVW: Stata module to perform predictive value weighting for covariate misclassification in logistic regression. Retrieved from https://econpapers.repec.org/RePEc:boc:bocode:s457825

Blakely, T., Barendregt, J. J., Foster, R. H., Hill, S., Atkinson, J., Sarfati, D. and Edwards, R. (2013): The association of active smoking with multiple cancers : national census-cancer registry cohorts with quantitative bias analysis. *Cancer Causes & Control* **24**, 1243–1255.

de Villiers EM, Fauquet C, Broker TR, Bernard HU, zur Hausen H. (2014): Classification of papillomaviruses. *Virology* **324**,17-27.

Fox, M. P. and Lash, T. L. (2017): On the Need for Quantitative Bias Analysis in the Peer-Review Process. *American Journal of Epidemiology* **185**, 865–868.

Gipson, B. J., Robbins, H. A., Fakhry, C. and Souza, G. D. (2018): Sensitivity and specificity of oral HPV detection for HPV-positive head and neck cancer. *Oral Oncology* **77**, 52–56.

Graham, C. A., Catania, J. A., Brand, R., Duong, T., Jesse, A. and Canchola, J. A. (2010): Recalling sexual behavior : A methodological analysis of memory recall bias via interview using the diary as the gold standard. *Journal of Sex Research* **4499**, 325-332.

Greenland, S. (1996): Basic Methods for Sensitivity Analysis of Biases. *International Journal of Epidemiology* **25**, 1107–1116.

Greenland, S. (2014): Sensitivity Analysis and Bias Analysis. In *Handbook of Epidemiology* (pp. 685–706). Springer US.

Hernán, M. and Robins, J. (2020): Measurement Error. In Chapman & Hall CRC (Ed.), *Causal Inference: What If*. Boca Raton.

Jurek, A. M., Maldonado, G. and Church, T. R. (2006): Exposure-measurement error is frequently ignored when interpreting epidemiologic study results. *European Journal of Epidemiology* **21**, 871–876.

Laprise, C., Arekunnath, S. M., Schlecht, N. F., Allison, P., Coutlée, F., Hier, Rouseau, M.C., Franco, E.L. and M. Nicolau, B. (2019): Increased risk of oropharyngeal cancers mediated by oral human papillomavirus infection : Results from a Canadian study. *Head & Neck* **41**, 678–685.

Lash, T. L., Fox, M. P. and Fink, A. K. (2009): *Applying Quantitative Bias Analysis to Epidemiologic Data*. New York: Springer US.

Lash, T. L., Fox, M. P., Maclehose, R. F., Maldonado, G. and Mccandless, L. C. (2014): Good practices for quantitative bias analysis. *International Journal Of Epidemiology* **43**, 1969–1985.

Lash, T. L. and Schmidt, M. (2010): Methods to apply probabilistic bias analysis to summary estimates of association y. *Pharmacoepidemiology and Drug Safety* **19,** 638–644.

Ligier, K., Belot, A., Launoy, G., Velten, M., Bossard, N., Iwaz, J., Righini CA., Delafosse P. and Guizard, A.-V. (2011): Descriptive epidemiology of upper aerodigestive tract cancers in France: incidence over 1980-2005 and projection to 2010. *Oral Oncology* **47**, 302–307.

Lingen, M. W., Abt, E., Agrawal, N., Chaturvedi, A. K., Cohen, E., D'Souza, G., Gurenlian, J., Kalmar, J.R., Kerr, A.R., Lambert, P.M., Patton, L.L., Sollecito, T.P., Truelove, E., Tampi, M.P., Urquhart, O., Banfield, L. and Carrasco-Labra, A. (2017): Evidence-based clinical practice guideline for the evaluation of potentially malignant disorders in the oral cavity. A report of the American Dental Association. *Journal American Dental Association* **148**, 712-727.

Lyles, R. H. and Lin, J. (2010): Sensitivity analysis for misclassification in logistic regression via likelihood methods and predictive value weighting. *Statistics in Medicine*, **29**, 2297–2309.

Orsini, N. (2007): Probabilistic bias analysis of epidemiological results. https://www.stata.com/meeting/2sweden/orsini-etal_slide_7sep07.pdf

Rettig, E. M. and D'Souza, G. (2015): Epidemiology of Head and Neck Cancer. *Surgical Oncology Clinics of North America* **24**, 379–396.

Rumball-Smith, J., Blakely, T., Sarfati, D. and Hider, P. (2013): The Mismeasurement of Quality by Readmission Rate : How Blunt Is too Blunt an Instrument? *Medical Care* **51**, 418–424.

Schlecht, N. F., Masika, M., Diaz, A., Nucci-sack, A., Salandy, A., Pickering, S. and Strickler, H. D. (2019): Risk of Oral Human Papillomavirus Infection Among Sexually Active Female Adolescents Receiving the Quadrivalent Vaccine. *AMA Network Open* **2**, 1–10.

StataCorp. (2017): Stata Statistical Software: Release 15. College Station. TX: StataCorp LLC.

Stott-Miller, M., Heike, C. L., Kratz, M., and Starr, J. R. (2010): Increased risk of orofacial clefts associated with maternal obesity : case – control study and Monte Carlo-based bias analysis. *Paediatric and Perinatal Epidemiology* 502–512.

The Global Cancer Observatory. (2019): Cancer Factsheets: Oropharynx, 2018–2019.

Tota, J. E., Best, A. F., Zumsteg, Z. S., Gillison, M. L., Rosenberg, P. S. and Chaturvedi, A. K. (2019): Evolution of the Oropharynx Cancer Epidemic in the United States: Moderation of Increasing Incidence in Younger Individuals and Shift in the Burden to Older Individuals. *Journal of Clinical Oncology* **37**, 1538–1546.

Valeri, L. and Vanderweele, T. J. (2014): The estimation of direct and indirect causal effects in the presence of misclassified binary mediator. *Biostatistics* **15**, 498–512.

Warnakulasuriya, S. (2010): Living with oral cancer: Epidemiology with particular reference to prevalence and life-style changes that influence survival. *Oral Oncology* **46**, 407–410.