

**EPIDEMIOLOGY OF ORAL CANCER IN INDIA
– A LIFE COURSE STUDY**

by

Sree Vidya Krishna Rao

Submitted for the degree of

Doctor of Philosophy

School of Dentistry



THE UNIVERSITY
of ADELAIDE

Supervised by

Professor Kaye Roberts-Thomson

Dr. Gloria Mejia

Professor Richard Logan

2014

'Knowing is not enough; we must apply.

Willing is not enough; we must do.'

-Johann Wolfgang von Goethe

Table of Contents

Table of Contents.....	ii
List of figures.....	ix
List of tables	x
Abstract.....	xii
Notes.....	xv
Declaration.....	xvii
Acknowledgements.....	xviii
1 Introduction	1
1.1 Rationale	3
1.2 Research hypotheses	5
1.3 Aims.....	5
1.4 Conceptual frameworks	5
1.5 Thesis structure	6
2 Review of the literature.....	8
2.1 Cancer of the oral cavity and oropharynx.....	8
2.2 Clinical aspects	8
2.2.1 Clinical features	9
2.2.2 Diagnosis	9
2.2.3 Histopathological types	9
2.2.4 Staging and prognosis.....	10
2.3 Screening for oral cancer	10
2.3.1 Principles of screening.....	11
2.3.2 Screening techniques and their predictive ability.....	13
2.3.3 What more could be done?	15
2.4 Global epidemiology of oral cancer.....	16
2.4.1 Incidence.....	16
2.4.2 Trend.....	17
2.4.3 Prevalence.....	17
2.4.4 Mortality and survival.....	18
2.4.5 Socioeconomic inequalities in oral cancer	18
2.4.5.1 Studies conducted using individual-based measures.....	19
2.4.5.2 Studies conducted using area-based measures	20

2.4.6	Global epidemiology of risk factors for oral cancer	21
2.4.6.1	Tobacco and alcohol	21
2.4.6.2	Diet	22
2.4.6.3	Poor oral health/hygiene	23
2.4.6.4	Body mass index	24
2.4.6.5	Human Papillomavirus	24
2.5	Epidemiology of oral cancer in Asia	26
2.5.1	Statement of Authorship	27
2.5.2	Introduction	28
2.5.3	Methods	28
2.5.4	Incidence	29
2.5.5	Trend	30
2.5.6	Age and gender	30
2.5.7	Sites in oral cavity	31
2.5.8	Recurrence	32
2.5.9	Mortality and survival	33
2.5.10	Socioeconomic conditions	37
2.5.11	Risk factors	38
2.5.11.1	Quid chewing	38
2.5.11.2	Tobacco use	39
2.5.11.3	Alcohol consumption	43
2.5.11.4	Diet	44
2.5.11.5	Viral infections	45
2.5.11.6	Oral hygiene	46
2.5.11.7	Family history of malignancy	46
2.5.11.8	Diabetes mellitus	46
2.5.11.9	Heavy metals	47
2.6	Conclusion	49
2.7	Causality in cancer epidemiology	50
2.7.1	Causality, causal inference and cancer	50
2.7.2	Study designs for causal inference	51
2.7.3	Conceptual model for causal analysis	52
2.8	The life course approach for oral cancer	52

2.8.1	Relevance of the life course approach	52
2.8.2	Life course models.....	53
2.8.2.1	The critical period model	53
2.8.2.2	The accumulation model	53
2.8.2.3	The pathway model	54
2.9	Social epidemiology of oral cancer.....	54
2.10	Socioeconomic measures in life course research	55
2.10.1	Individual-level measures.....	56
2.10.1.1	Occupation.....	56
2.10.1.2	Education.....	57
2.10.1.3	Income.....	57
2.10.2	Area-level measures.....	58
2.11	Conclusion.....	59
3	Methodology	60
3.1	Study design.....	60
3.2	Study setting.....	60
3.2.1	Cancer hospitals selected for the study.....	60
3.2.1.1	Location of the cancer hospitals.....	61
3.2.1.2	Description of the cancer hospitals	61
3.2.1.3	Source population for the cancer hospitals.....	62
3.3	Selection criteria	62
3.3.1	Selection criteria for cases	62
3.3.1.1	Case definition.....	62
3.3.1.2	Inclusion criteria.....	63
3.3.1.3	Exclusion criteria.....	63
3.3.2	Selection criteria for controls.....	64
3.3.2.1	Principles of control selection	64
3.3.2.2	Inclusion criteria.....	65
3.3.2.3	Exclusion criteria	65
3.4	Sampling	66
3.4.1	Cases.....	66
3.4.2	Controls.....	66
3.5	Estimated sample size	67

3.6	Study participant recruitment.....	68
3.6.1	Identification of oral cancer cases and controls.....	68
3.6.2	Contacting participants and obtaining informed consent.....	68
3.7	Data collection.....	69
3.7.1	Data collection instruments and methods.....	69
3.7.2	Preparing and pilot testing the questionnaire.....	69
3.7.3	Conducting direct interviews.....	70
3.7.4	Use of the life grid.....	70
3.7.5	Training the examiner.....	71
3.7.6	Oral examination.....	71
3.7.6.1	Infection control protocol.....	72
3.7.6.2	Oral examination equipment.....	72
3.7.6.3	Oral examination procedure.....	72
3.7.7	Record abstraction.....	74
3.7.8	Referral.....	75
3.7.9	Quality assurance of data collected.....	75
3.7.10	Confidentiality.....	75
3.10	Data entry and cleaning.....	75
3.8	Ethics clearance.....	76
4	The life course models: Exploring associations between socioeconomic conditions at three stages in life and oral cancer.....	77
4.1	Statement of Authorship.....	78
4.2	Abstract.....	80
4.3	Introduction.....	81
4.4	Methods.....	83
4.4.1	Study design.....	83
4.4.2	Selection of cases and controls.....	84
4.4.3	Data collection.....	84
4.4.4	Measures.....	85
4.4.5	Statistical analyses.....	85
4.5	Results.....	86
4.6	Discussion.....	90
4.6.1	Critical period.....	91
4.6.2	Social mobility.....	91

4.6.3	Accumulation.....	91
4.7	Conclusion	93
5	Estimating the effect of childhood socioeconomic disadvantage on oral cancer in India using marginal structural model.....	94
5.1	Statement of Authorship	95
5.2	Abstract.....	98
5.3	Introduction.....	99
5.4	Methods.....	101
5.4.1	Participant recruitment.....	101
5.4.2	Data collection	102
5.4.3	Measures	103
5.4.3.1	Outcome	103
5.4.3.2	Exposure.....	103
5.4.3.3	Confounders and mediators.....	103
5.4.4	Statistical analyses	104
5.4.5	Outcome models	105
5.4.6	Estimating stabilised weights	107
5.4.7	Sensitivity analysis	108
5.5	Results.....	108
5.6	Discussion	116
5.7	Conclusion	118
6	A screening model for oral cancer using risk scores: Development and validation	119
6.1	Statement of Authorship	120
6.2	Abstract.....	123
6.3	Introduction.....	124
6.4	Methods.....	125
6.4.1	Study design.....	125
6.4.2	Participant selection and recruitment.....	126
6.4.2.1	Cases.....	126
6.4.2.2	Controls	126
6.4.3	Potential predictors	127
6.5	Statistical analyses	128
6.5.1	Model development	128
6.5.2	Model validation.....	129

6.5.3	Analysis of prevalence.....	129
6.6	Results.....	129
6.7	Discussion.....	138
6.8	Conclusion.....	142
7	Discussion.....	143
7.1	Summary.....	143
7.2	Why life course epidemiology of oral cancer?.....	144
7.3	What are the study findings?.....	145
7.3.1	Which life course models can explain the association of oral cancer with socioeconomic conditions at various stages in life?.....	145
7.3.1.1	Critical period model.....	145
7.3.1.2	Social mobility models.....	146
7.3.1.3	Accumulation model.....	146
7.3.2	Does low SEC during childhood have a long-term effect on oral cancer in adulthood?.....	146
7.3.2.1	Mediation through behavioural factors.....	147
7.3.2.2	Genetic and epigenetic explanation for long-term effect of childhood low SEC on oral cancer.....	148
7.3.3	How well can the risk score model predict oral cancer?.....	149
7.4	How methodological issues have been addressed?.....	151
7.4.1	Using a case-control design for life course research on oral cancer.....	151
7.4.2	Causal and predictive modelling in life course research.....	153
7.4.3	Measures of association and effect measures used.....	155
7.5	What are the strengths and limitations of the study?.....	155
7.5.1	Strengths.....	156
7.5.2	Limitations.....	157
7.6	What are the study implications?.....	159
7.6.1	Implications for public health consideration.....	159
7.6.2	Implications for future research considerations.....	160
7.7	Conclusions.....	161
8	References.....	164
9	Appendices.....	186
1.	Ethics approval.....	186
2.	Information sheet and consent form.....	194

3. Questionnaire.....	201
4. Oral examination and record abstraction forms.....	232
5. Manual	236

List of figures

Figure 3.1 Places where the selected cancer hospitals are located in India.....	61
Figure 4.1 Conceptual model.....	83
Figure 5.1 Conceptual framework	101
Figure 6.1 ROC curve for the multivariable model	135
Figure 6.2 ROC curve for the multivariable model using the bootstrap sample	135
Figure 6.3 ROC curve for risk scores	136
Figure 6.4 ROC curve for risk scores using the bootstrap sample	136

List of tables

Table 2.1 Predictive ability of visual screening as reported in 36 studies.....	14
Table 2.2 Countries with highest and lowest age-standardised incidence rates of oral cancer in both sexes according to GLOBOCAN 2012 (Ferlay et al., 2013)	16
Table 2.3 A summary table showing data on various outcome estimates of oral cancer available from Asian countries from 2000 to 2012.....	36
Table 2.4 A summary table showing data on risk factors for oral cancer available from various Asian countries from 2000-2012*	48
Table 3.1 ICD-10 codes and corresponding oral and oropharyngeal sites	63
Table 3.2 Interpretation of OHI-S scores	74
Table 4.1 Distribution of cases and controls according to sociodemographic characteristics .	87
Table 4.2 Distribution of socioeconomic trajectories among oral cancer cases and controls .	88
Table 4.3 Unadjusted and adjusted odds ratio for socioeconomic conditions at various critical periods	89
Table 4.4 Unadjusted and adjusted odds ratio of oral cancer for accumulation of number of occasions being in the low socioeconomic condition	89
Table 4.5 Association of social mobility with oral cancer	90
Table 5.1 Distribution of life course socioeconomic measures, paternal and own habits among cases and controls, and crude risk ratio estimates with 95% confidence intervals	109
Table 5.2 Total effect, adjusted effect from conventional regression, and controlled direct effect from the marginal structural model for oral cancer	110
Table 5.3 Sensitivity analysis for unmeasured confounding considering three estimates of controlled direct effect of early-life socioeconomic conditions on oral cancer	112
Table 5.4 Frequency distribution of cases and controls according to joint exposure to childhood SEC (CHSEC) and smoking	113
Table 5.5 RERI between childhood SEC and smoking on the risk ratio scale.....	113

Table 5.6 Multiplicative interaction between childhood SEC and smoking on risk ratio scale	113
Table 5.7 Frequency distribution of cases and controls according to joint exposure to childhood SEC and chewing quid/tobacco (CHEWING)	114
Table 5.8 RERI between childhood SEC and chewing quid/tobacco on risk ratio scale	114
Table 5.9 Multiplicative interaction between childhood SEC and chewing on risk ratio scale	115
Table 5.10 Frequency distribution of cases and controls according to joint exposure to childhood SEC and alcohol	115
Table 5.11 RERI between childhood SEC and alcohol on risk ratio scale	115
Table 5.12 Multiplicative interaction between childhood SEC and alcohol on risk ratio scale	116
Table 6.1 Characteristics of cases and controls and unadjusted odds ratios	130
Table 6.2 Multivariable logistic regression with all predictors of oral cancer included in the full model	132
Table 6.3 Multivariable regression model for predictors of oral cancer and risk scores for the predictors	134
Table 6.4 Comparison of sensitivity and specificity of risk score cut-offs of 5, 6 and 7	137
Table 6.5 Comparison of predictive ability of multivariable model and risk scores with the cut-off as 6 in the study and bootstrap samples	137
Table 6.6 Frequency distribution of cases and controls by risk score cut-off at 6	138
Table 6.7 Expected PPV and NPV according to proportion of cases in the study population and prevalence at national level	138

Abstract

Background: Oral cancer is a malignant disease contributing to one third of the total cancer burden in India. There is a worldwide social disparity in oral cancer incidence and survival. Life course epidemiology has shown that early-life socioeconomic conditions (SEC) could influence adult health through various pathways. Thus, the socioeconomic disparities in the occurrence of oral cancer underscore the importance of understanding the ‘life course processes’ operating between SEC at different stages in life and oral cancer. In addition to understanding socioeconomic disparities, practical solutions are required to reduce the burden of the disease. Early diagnosis and prompt treatment could reduce morbidity and mortality. Though visual screening helps in early diagnosis, it requires training and calibration of the screeners. Developing a simple screening model that can be utilized by untrained health care workers will be helpful in triaging asymptomatic adults with oral cancer.

Therefore a study was designed with the following hypotheses.

Research hypotheses

1. Accumulation of socioeconomic disadvantage over the life course is associated with oral cancer in the Indian population.
2. Early-life socioeconomic disadvantage has a lasting effect on the oral cancer outcome in adulthood in the Indian population.
3. An oral cancer screening model developed for the Indian population, to screen high-risk people from rural/remote areas, has good predictive ability.

Methods: A multicentre hospital based case-control study was conducted between July 2011 and August 2012 in Karnataka, India. Cases were newly diagnosed oral and oropharyngeal cancer patients and controls were patient-visitors or patients seeking care for other reasons. Data were collected through direct interviews, oral examination and record abstraction. Cases were ascertained from hospital records. A semi-structured questionnaire was designed to collect life course information on SEC, family structure, housing conditions, parental habits of tobacco, quid and alcohol use, parental education, family history of malignancy, participants’ own diet, tobacco, quid, alcohol use and oral hygiene behaviour. A life-grid was used to improve recall accuracy. All consenting participants underwent an oral examination following an interview by a trained examiner. Oral soft and hard tissues were examined for the presence of any oral mucosal lesions, teeth present and oral hygiene status.

Data were analysed using SAS v 9.2. Conventional logistic regression models were used to determine the associations between life course SEC and oral cancer. Marginal structural model (MSM) was built to estimate the controlled direct effect of childhood SEC on oral cancer in adulthood. The validity of effect measures was checked with sensitivity analysis. A multivariable logistic regression model was used to develop a screening model for identifying individuals at high-risk for developing oral cancer. The development of the model involved deriving risk scores for the predictors of oral cancer. The predictive ability of the screening model was examined with *c* statistics, sensitivity, specificity and predictive values.

Results: A total of 180 incident cases and 272 controls participated in the study. Of them, 163 cases and 264 controls had complete information on SEC at all three stages. Nearly two-thirds (65%) of participants were stable in low SEC across all stages. Low SEC at all the three stages (childhood, early adulthood and later adulthood) was associated with oral cancer after adjusting for age and sex. The association was strongest for those who remained in the low SEC at all the three stages. Odds ratios (OR) for oral cancer in socially mobile groups were intermediate to that of the stable groups. The largest differences in OR for oral cancer were observed between the stable groups.

The total effect model showed that the risk was 63% [Risk ratio (RR) = 1.63 (95% CI = 1.38–1.92)] higher for those who lived in low SEC in childhood than for those in high SEC. From the MSM, the estimated risk for developing oral cancer for those in low SEC during early-life was 48% [RR = 1.48 (95% CI = 1.43–1.53)], 24% [RR = 1.24 (95% CI = 0.88–1.74)] and 94% [RR = 1.94 (95% CI = 1.66–2.27)] greater than those in the high SEC after controlling for smoking, chewing and alcohol respectively. However, the adjusted effect of low SEC on oral cancer was null when analysed using conventional regression.

A screening model was developed using statistical methods that involved smoking, chewing quid and/or tobacco, alcohol, family history of upper aero-digestive tract (UADT) cancer, diet and oral hygiene behaviour as predictors. Total risk score that was derived from odds ratio ranged from 0 to 28. Area under the curve of the Receiver Operating Characteristic (ROC) curve for risk scores was 0.866. The sensitivity (0.928) and negative predictive value (0.927) were higher while specificity (0.603) and positive predictive value (0.607) were lower for risk scores cut-off of 6.

Conclusions: Low SEC in childhood and early adulthood are important in determining oral cancer in later adulthood. Early-life socioeconomic disadvantage increases the risk for oral

cancer that is not mediated by later life risk factors when MSM was used. The developed screening model using risk scores had satisfactory predictive ability in the study population. However, validation of the model in other settings is necessary before it can be recommended to identify subgroups of the people to be referred for further clinical evaluation.

Notes

References

References in this thesis follow a generic style that provides author-date citations where the author(s) and date of publication is listed in the parentheses. In the text, to differentiate work by same authors in the same year, a letter after the year is included. In this Harvard author-date referencing system, where there are three or more authors, the first author is listed followed by “et al.” in the text. All authors are listed in the bibliography.

List of Abbreviations

ASIR	Age-standardised Incidence Rate
AUC	Area Under Curve
NSAOH	Australian National Survey on Adult Oral Health
ARCPOH	Australian Research Centre for Population Oral Health
BMI	Body Mass Index
CI-S	Calculus Index – Simplified
CVD	Cardiovascular Diseases
CCI	Commission on Chronic Illness
CDE	Controlled Direct Effect
DI-S	Debris Index – Simplified
DM	Diabetes Mellitus
DAG	Directed Acyclic Graphs
EBV	Epstein Barr Virus
FN	False Negative
FP	False Positive
HH	Head of Household
HCGSC	HealthCare Global Speciality Centre
HCG-BIO	HealthCare Global-Bangalore Institute of Oncology
HPV	Human Papillomavirus
HSV-1	Human Simplex Virus-1
ICD	International Classification of Diseases
IPW	Inverse Probability Weight
KMIO	Kidwai Memorial Institute of Oncology
MSM	Marginal Structural Model

MSW	Medical Social Worker
NIDCR	National Institute of Dental and Craniofacial Research
NIDR	National Institute of Dental Research
NSAOH	National Survey of Adult Oral Health
NPV	Negative Predictive Value
OR	Odds Ratio
OHI-S	Oral Hygiene Index-Simplified
OPMD	Oral Potentially Malignant Disorders
PNG	Papua New Guinea
PAH	Polycyclic Aromatic Hydrocarbons
PPV	Positive Predictive Value
RCT	Randomised Controlled Trial
RERI	Relative Excess Risk due to Interaction
RR	Risk Ratio
ROC	Receiver Operating Characteristic
SSSBCH	Shri Shirdi Sai Baba Cancer Hospital
SLT	Smokeless Tobacco
SEC	Socioeconomic Conditions
SEP	Socioeconomic Position
SCC	Squamous Cell Carcinoma
SW	Stabilised Inverse Probability Weight
SEER	Surveillance, Epidemiology, and End Results
US	United States
USPSTF	United States Preventive Services Task Force
UADT	Upper Aerodigestive Tract
WHO	World Health Organization

Declaration

I certify that this work contains no material which has been accepted for the award of any other degree or diploma in my name, in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. In addition, I certify that no part of this work will, in the future, be used in a submission in my name, for any other degree or diploma in any university or other tertiary institution without the prior approval of the University of Adelaide.

I give consent to this copy of my thesis when deposited in the University Library, being made available for loan and photocopying, subject to the provisions of the Copyright Act 1968.

The author acknowledges that copyright of published works contained within this thesis resides with the copyright holder(s) of those works.

I also give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the Library Search and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

Signed _____

Sree Vidya Krishna Rao

_____/_____/_____

Date

Acknowledgements

I am grateful to the Supreme power for giving me the energy and enthusiasm to undertake and complete this study.

I would like to express my deep sense of appreciation to Professor Kaye Roberts-Thomson for accepting me as her student, being an excellent supervisor and a source of strength in guiding me all along this PhD journey. I have no words to express the generosity and kindness she has shown me during my study.

I would like to express my heartfelt gratitude to Dr. Gloria Mejia for having been a tremendous mentor and a superb teacher. I am grateful to her for encouraging my research and helping me to think and work independently as a researcher and develop the scholarly qualities.

I wish to express my deep sense of admiration to Professor Richard Logan for his promptness and useful advice in discussing various aspects of oral medicine, diagnosis and pathology. His useful critiques as a supervisor have helped me immensely during this study.

I wish to thank Dr. Murthy N Mittinty, Bio-statistician at School of Population Health, the University of Adelaide, for making regular trips to ARCPOH, discuss the complexities of biostatistics which helped me in writing my research papers.

I wish to thank Professor Veena Kamath and Associate Professor Muralidhar Kulkarni at the Department of Preventive and Community Medicine, Kasturba Medical College, Manipal, Manipal University for their collaboration and research guidance during data collection.

I am grateful to Professor Donald Fernandes, Department of Radiotherapy, and Professor Satadru Ray, Department of Surgical Oncology, Shri Shirdi Sai Baba Cancer Hospital and Research Centre, Manipal, for their collaboration and guidance in the clinical setting.

A special thanks to Prof K. Ramnarayan, Vice Chancellor of Manipal University, for his help in establishing the collaboration with researchers at Manipal University.

I sincerely thank Dr. Amit Verma, Former Director, Triesta, for his help in data collection at HCG Hospitals in Bangalore. My thanks to Dr. Ramesh and Professor Vijai Kumar for their initial permission and assistance to collect data at the Kidwai Memorial Institute of Oncology, Bangalore.

I wish to thank Mr Serge Chrisopoulos for helping me setup a database. I wish to thank Dr. Yvonne Miels for her editorial assistance. I am thankful to Mrs Silvana Marveggio for her administrative support and I thank all my friends and colleagues who have contributed directly or indirectly towards this study.

Words are not enough in expressing my gratitude to all those study participants and their families who despite their difficult circumstances, grief and financial difficulties showed their strength, support and solidarity towards my research so that it could make a difference and help in preventing others from suffering this deadly disease. This study would not have been possible without their immense contribution. I shall continue to strive for their betterment.

I would like to extend my gratitude to all the nurses and medical social workers at the hospitals for their help in recruiting the study participants.

I wish to thank my family for their constant love, support and encouragement, without which I would not have been able to complete this study.