Bayesian Disease Mapping and the 'High-Risk' Oral Cancer Population in Hong Kong

John Adeoye, Siu-Wai Choi, Peter Thomson

Oral & Maxillofacial Surgery, Faculty of Dentistry, The University of Hong Kong, Hong Kong, SAR China

Correspondence: John Adeoye, Oral & Maxillofacial Surgery, Faculty of Dentistry, The University of Hong Kong, Hong Kong, SAR China.

E-mail: thomsonp@hku.hk

Abstract

Background: Preventive and early diagnostic methods such as health promotion and disease screening are increasingly advocated to improve detection and survival rates for oral cancer. These strategies are most effective when targeted at 'high-risk' individuals and populations. Bayesian disease-mapping modelling is a statistical method to quantify and explain spatial and temporal patterns for risk and covariate factor influence, thereby identifying 'high-risk' sub-regions or 'case clustering' for targeted intervention. Rarely applied to oral cancer epidemiology, this paper highlights the efficacy of disease mapping for the Hong Kong population.

Methods: Following ethical approval, anonymized, individual-level data for oral cancer diagnoses were obtained retrospectively from the Clinical Data Analysis and Reporting System (CDARS) of the Hong Kong Hospital Authority (HA) database for a 7-year period (January 2013 to December 2019). Data facilitated disease mapping and estimation of relative risks of oral cancer incidence and mortality.

Results: 3,341 new oral cancer cases and 1,506 oral cancer-related deaths were recorded during the 7-year study period. Five districts, located in Hong Kong Island and Kowloon, exhibited considerably higher relative incidence risks with 1 significant 'case cluster' hotspot. Six districts displayed higher mortality risks than expected from territory-wide values, with highest risk identified for two districts of Hong Kong Island.

Conclusion: Bayesian disease mapping is successful in identifying and characterising 'high risk' areas for oral cancer incidence and mortality within a community. This should facilitate targeted preventive and interventional strategies. Further work is encouraged to enhance global-level data and comprehensive mapping of oral cancer incidence, mortality and survival.

Keywords: Oral Cancer, Disease Mapping, Hong Kong Population Study

1. Introduction

The risk of disease varies substantially in both space and time, influenced by underlying genetic susceptibility and differences in environmental exposure and lifestyle behaviour[1]. Quantifying and explaining such variation within a defined population is termed disease mapping. Mapping facilitates counting of case numbers within non-overlapping 'areal units' over defined time periods, and attempts to explain and predict the patterns of disease outcome within a specific geographic location. By formulating and validating hypotheses of disease aetiology, mapping can enhance public health initiatives by identifying areas of elevated risk and then targeting appropriate health care intervention[2].

Cancer mapping has been trialled to determine spatial variations in disease clustering and to document risk factor distribution, case-related mortality and survival data; 'cancer atlases' have been published in Australia, Italy and the USA. Statistical models based on Bayesian conditional probability have proved most effective in generating robust 'smoothed' cancer outcomes combining information from data likelihood and prior distributions (an uncertainty measure) to generate reliable posterior probability estimates thus informing reliable map construction[2-9].

Oral cancer, primarily squamous cell carcinoma (SCC) arising from the mucosal lining of the lips, oral cavity and oropharynx, is the 14th highest malignancy worldwide in terms of both incidence and mortality[1,10,11]. With substantive geographic variation in disease distribution, proactive preventive measures and targeted screening of disease-prone individuals in 'high-risk' regions are pivotal methods to reduce the high, five-year mortality rates associated with late cancer detection[11,12]. Although lip cancer was studied in one of the first Bayesian disease mapping estimates, it is notable that there is little published research on oral cancer mapping.

In a series of previous investigations, we reported a significant increase in incidence in oral cancer within the population of Hong Kong; it is currently the tenth highest cancer in males[13-15]. Located on the South Eastern tip of China, Hong Kong is a semi-autonomous Special Administrative Region (SAR) with a total population of 7.4 million people and a land area of 1,106 square kilometres[16]. The territory comprises 3

principal geographic regions: the densely-populated urban areas of Hong Kong Island and Kowloon, and the territorially much larger and consequently less populated New Territories. These areas are divided into 4, 5, and 9 District Council Districts (DCD), respectively, totalling 18 constituencies for administrative and election purposes[16,17]. Although possessing one of the highest life expectancies worldwide, cancer has become the most frequent cause of death in Hong Kong with ten-year age-standardized incidence and mortality rates ranging between 213.9 - 232.2 per 100,000 and 84.4 - 98.2 per 100,000 persons, respectively[18-20].

Whilst screening programmes are available in Hong Kong for colorectal, ovarian, cervical, breast and nasopharyngeal cancers, none are provided for oral cancer and significant ignorance exists in the community regarding the disease. Development of a risk-targeted screening approach within the SAR could offer significant and pragmatic opportunities to reduce disease burden, and would be substantially facilitated by improved baseline data of oral cancer incidence and mortality[12,21].

The aim of this paper, therefore, was to construct Bayesian disease maps of oral cancer within Hong Kong and to attempt to illustrate definitively for the first time the local population at highest risk of disease.

2. Methods

2.1 Data Extraction

Following ethical approval, anonymised, individual-patient oral cancer data were obtained retrospectively from the Clinical Data Analysis and Reporting System (CDARS) of the Hong Kong Hospital Authority (HA) database, for the 7-year period January 2013 to December 2019. Oral cancer cases were identified from the CDARS database, using search terms aligned to the International Classification of Diseases, 10th revision (ICD-10) codes C00-C10 and C14. Exclusion criteria were cases with a provisional diagnosis only, or patients less than 20-years-old at the time of presentation. DCDs supplied the smallest non-overlapping geographic units for which

data could be obtained and satisfactory estimates generated. Total population estimates and counts of DCD residents aged 20 years and above (stratified by their age-groups at 5-year intervals) were obtained from the most recent population census conducted in 2016 by the Census and Statistics department of the Hong Kong SAR Government[17].

2.2 Spatial Modelling

Since the smallest areal units of Hong Kong could not provide outcome data, 95% credible intervals were used to represent the degree of uncertainty, achieved with hierarchical or full Bayes estimation[22]. The approach is based on the generalized linear mixed model, which fits outcome estimates according to the influence of pertinent covariates and a set of random effects that account for both within and between-area variations. This framework supports the principle of spatial dependence, which considers near objects more related than distant ones[2,4,23-25]. Model fitting criteria using original study data and the observation of realistic estimates favoured the use of the convolution model for data smoothing. The conventional standardized incidence and mortality ratio (θ_h), representing the risk of oral cancer occurrence and mortality respectively within non-overlapping geographic units as the response variable, was used. Details of these methods are in *Appendix 1*.

2.3 Computation and Statistical Software

Adjacency matrix was created using the R statistical software v 3.6.3 using the 'spdep' package v1.1-3[26]. The binary adjacency matrix for modelling neighbourhood dependence defined areas sharing common boundaries (Rook definition) using first order weights. Only a single DCD had no definite neighbours during initial matrix generation, after which neighbours were manually assigned prior to parameter estimation. Posterior sampling of parameter estimates was carried out using the Gibbs sampling Markov Chain Monte Carlo (MCMC) technique using WINBUGS14 and R v.3.6.3 via the CARBayes package v 5.2[27]. Convergence of this chain was assessed using the Monte Carlo error method, visual inspection of trace, density and auto-

correlation plot and Geweke diagnostic test on R via the CODA package v 0.19-3 (as shown in *Appendix 2*). Median smoothed estimates of incidence and mortality relative risks, their 95% credible interval obtained, and probability of state-average relative risk exceedance inferences were used to construct choropleth thematic maps using GeoDa v 1.14 and R via ggplot2 package[28]. Hong Kong DCD geometry data and accompanying shapefiles were obtained from http://opendata.esrichina.hk/. Evidence of spatial clustering and variation was assessed using the Global and Local indicator of spatial autocorrelation (LISA)/local Moran's I test on GeoDa with probability values calculated using Monte Carlo randomization at 99999 permutations *(Appendix 3)*.

3. Results

3.1 Oral Cancer Incidence

3,341 new oral cancer cases were identified during the 7-year period with crude incidence rates ranging between 7.22 and 8.92 per 100,000 persons; lowest and highest rates were seen in 2014 and 2019, respectively, with a substantive upward trend occurring between 2017 and 2019. Median standardized incidence ratios (SIR) varied between DCDs when compared to the average Hong Kong estimate (1.0), with minimum and maximum relative risks of 0.59 and 2.30, respectively; the spatial or clustering fraction of random effects represented in the model was 0.708. Based upon 5 categories of increasing risk (from very low to very high), SIR distribution within the 18 DCDs are listed in Table 1 whilst Figure 1 maps the posterior probability of exceeding average Hong Kong risk in each DCD. Whilst the risk of oral cancer development was compatible with the average in most DCDs, 3 districts in Kowloon (Kowloon City, Wong Tai Sin and Yau Tsim Mong) and 2 in Hong Kong Island (Southern, and Central and Western) displayed a considerably higher relative risk.

Global spatial autocorrelation analysis using the Moran's I test revealed significant positive autocorrelation with test statistic and p-value of 0.354 and 0.02 respectively. Further analysis to detect specific incidence clusters and their associations using the local Moran's I test revealed a significant 'hot-spot' cluster and low-high spatial outlier within the Hong Kong Island area only *(Appendix 3)*.

3.2 Oral Cancer Mortality

1,506 oral cancer-related deaths were recorded during the study period. Modelled median standardized mortality ratios (SMR) ranged from 0.70 to 1.54 across the DCDs. The proportion of variation explained by correlated and uncorrelated random effects in the model was 0.544 and 0.456 respectively. Based upon 5 categories of increasing risk (from very low to very high), SMR distribution within the 18 DCDs listed in Table 1 exhibit a decreasing mortality risk from Hong Kong Island to Kowloon to New Territories, with especially high mortality observed in Southern and Central and Western Districts. Figure 2 maps the posterior probability of exceeding average Hong Kong mortality risk in each DCD; with the addition of the Eastern District of Hong Kong Island a total of 6 DCDs were seen to exhibit higher mortality risks than territory-wide values.

Global Moran's I test showed significant positive spatial autocorrelation in the mortality relative risks with a test statistic value of 0.486 (p=0.004). Local spatial dependence analysis (LISA cluster and significance maps) identified clusters with high and low mortality risks in Hong Kong Island and the New Territories, respectively (*Appendix 3*).

4. Discussion

4.1 Oral Cancer in Hong Kong

Despite real advances in clinical management, oral cancer remains a lethal and deforming disease of rising incidence and global significance. This study has added further evidence of an increasing incidence within the Hong Kong population[15]. It is recognised that early diagnosis and timely intervention can significantly improve oral cancer outcomes and patient survival. From a public health perspective, screening of the 'high risk' population thus becomes an imperative[14]. The difficulty facing community–based studies is exactly how to identify this important sub-population. Although oral cancer population data from regional cancer registries are increasingly available in published text and tabular format, it remains difficult for investigators to

localize population clusters and track salient patterns of emerging disease and specific geographic variation.

In a previous publication, we documented the overall pattern of new oral cancer patient presentations in Hong Kong, based upon their initial registration within the HA CDARS system, and found clustering on Hong Kong Island and within Kowloon[15]. Whilst initial assessment considered the potential confounding influence of new case clustering around specialist teaching hospitals, disease mapping in this study confirms the increased risk of oral cancer within these specific local populations.

4.2 Bayesian Disease Mapping

This study utilised a Bayesian statistical approach, improving both data reliability and interpretation, to map oral cancer incidence and mortality risk within well-defined geographic regions of Hong Kong. Specific and consistent recognition of disease 'hotspots' within the geographic region is a demonstrable success. Five significant 'high-risk' patient clusters were confirmed, affecting Southern, and Central and Western Districts of Hong Kong Island, and Kowloon City, Wong Tai Sin and Yau Tsim Mong in Kowloon. With the additional recognition of Hong Kong Island's Eastern District, the resultant 6 DCDs also exhibited high cancer mortality, which is a potentially significant public health observation.

Many of these affected DCDs include highly populated urban areas with extensive networks of bars, restaurants and smoking areas concentrated in close proximity to public housing and residential areas. Whilst the aetiological factors responsible for oral carcinogenesis are well known, such detailed geographic and resultant demographic association will help guide subsequent studies in risk-modifying behaviour.

4.3 The 'High-Risk' Population

The Hong Kong population most at risk of oral cancer development also exhibited the highest mortality risk in this study. As there is reasonable access to public health care facilities within most DCDs, especially those concentrated on Hong Kong Island and within Kowloon, this is most likely the consequence of late presentation of aggressive, advanced stage disease and resulting high morbidity and poor survival[14,29]. Although overall mortality rates have been criticized as measures of disease detection, it is clear that earlier recognition of patients at risk of oral cancer is a priority for public health interventions. In the future, utilisation of time-bound mortality rates, relative survival and net survival estimates using period analyses may provide more accurate data and improve mapping to establish more reliable inference on early diagnosis rates[30,31].

4.4 Targeted Screening and Intervention

Accurate disease mapping informs a targeted approach to future screening, thus providing opportunities to improve cancer detection rates, facilitate preventive strategies and to deliver early interventional programmes to 'high-risk' communities. Longer term, such techniques may also inform causation hypotheses to explain observed disparities in disease incidence, convey pertinent health messages to the relevant public and improve monitoring of spatial parameter change over time. In the 21st Century, these approaches will prove invaluable in our attempts to lessen the global burden of oral cancer[1].

4.5 Study Limitations

This study was for a 7-year period only and the data analysed were not obtained from an official cancer registry, nor were additional data accessed from private medical institutes in Hong Kong out with the public healthcare system. However, as in our previous investigations, it was felt that the most detailed and consistent oral cancer information would be obtainable from HA databases populated by specialist surgical and oncology departments in a teaching hospital environment[15].

Whilst it is evident that Bayesian modelling and associated computation are complex processes, cartographic software, resource material and vignettes are all readily available for consultation without the specific requirement of extensive pre-existing statistical knowledge[23,24]. Perhaps the principal limitation in producing cartography outputs rests with the quality of secondary data utilized for parameter estimation. In the future, enhanced collaboration with spatial scientists, biomedical statisticians and cancer epidemiologists may all enhance the robustness of thematic disease maps in oral oncology[32,33].

Overall, however, it is hoped that the data and maps presented in this paper represent a significant step forward in formulating comprehensive oral cancer screening guidelines, not only for the Hong Kong population, but for a wider global application.

5. Conclusions

This paper has demonstrated that Bayesian disease mapping is effective in identifying sub-groups within the Hong Kong population at increased risk of oral cancer and, perhaps more significantly, highlighting mortality 'hot spots'. Whilst mapping may yet be unable to answer fundamental questions regarding aetiological influences, the authors hope that this article will encourage successive replication in different regions especially in those areas of high disease burden and result in on-going collaborations to improve understanding of oral cancer incidence, mortality and survival on a global level.

Funding

None declared.

Conflict of Interest

None declared.

Ethical Approval

Ethical approval to conduct the study and endorsement for the use of anonymized clinical data was granted by the Institutional Review Board of the University of Hong Kong/Hospital Authority Hong Kong West Cluster (Reference number UW-19-773).

Patient Consent

Not required.

References

1. Thomson PJ. Oral Cancer: From Prevention to Intervention. Newcastle upon Tyne: Cambridge Scholars Publishing; 2019.

2. Anderson C, Lee D, Dean N. Identifying clusters in Bayesian disease mapping. Biostatistics. 2014 15: 457-69.

3. Cramb SM, Mengersen KL, Baade PD. Developing the atlas of cancer in Queensland: methodological issues. Int J Health Geogr. 2011 10: 9.

4. Kang SY, Cramb SM, White NM, Ball SJ, Mengersen KL. Making the most of spatial information in health: a tutorial in Bayesian disease mapping for areal data. Geospat Health. 2016 11: 428.

5. Adeyemi R, Zewotir T, Ramroop S. A bayesian hierarchical analysis of geographical patterns for child mortality in Nigeria. Open Public Health J. 2019 12: 247-61.

6. Lee D. A comparison of conditional autoregressive models used in Bayesian disease mapping. Spat Spatiotemporal Epidemiol. 2011 2: 79-89.

7. Bilancia M, Fedespina A. Geographical clustering of lung cancer in the province of Lecce, Italy: 1992-2001. Int J Health Geogr. 2009 8: 40.

8. Duncan EW, Cramb SM, Aitken JF, Mengersen KL, Baade PD. Development of the Australian Cancer Atlas: spatial modelling, visualisation, and reporting of estimates. Int J Health Geogr. 2019 18: 21.

9. Johnson GD. Small area mapping of prostate cancer incidence in New York State (USA) using fully Bayesian hierarchical modelling. Int J Health Geogr. 2004 3: 29.

10. Conway DI, Purkayastha M, Chestnutt IG. The changing epidemiology of oral cancer: definitions, trends, and risk factors. Br Dent J. 2018 225: 867-73.

11. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2018 68: 394-424.

12. Sankaranarayanan R, Ramadas K, Thomas G, Muwonge R, Thara S, Mathew B, et al. Effect of screening on oral cancer mortality in Kerala, India: a cluster-randomised controlled trial. Lancet. 2005 365: 1927-1933.

13. Ushida K, McGrath CP, Lo EC, Zwahlen RA. Oral cavity cancer trends over the past 25 years in Hong Kong: a multidirectional statistical analysis. BMC Oral Health. 2015 15: 83.

14. Thomson P, Su R, Choi S. Oral cancer in Hong Kong: identifying and managing the 'high-risk' population. Fac Dent J. 2018 9: 116-21.

15. Choi SW, Thomson P. Increasing incidence of oral cancer in Hong Kong-Who, where...and why? J Oral Pathol Med. 2019 48: 483-90.

16. GOV.HK. Hong Kong - the facts. Hong Kong: Hong Kong Special Adminstrative Region Government; 2020 [Available from: https://www.gov.hk/en/about/abouthk/facts.htm.

17. Census and Statistics Department HG. Hong Kong 2016 population by census - District profiles 2020 [Available from: https://www.bycensus2016.gov.hk/en/bc-dp.html.

18. Census and Statistics Department HG. The mortality trend in Hong Kong from 1986 to 2018 Hong Kong: CSD HKSAR Govt; 2020 [Available from: https://www.censtatd.gov.hk/hkstat/sub/sp160.jsp?productCode=FA100094.

19. Hospital Authority HKCR. Cancer Statistics Query Systems - all ages 2020 [Available from: https://www3.ha.org.hk/cancereg/allages.asp.

20. Hospital Authority HKCR. Cancer Statistics Query Systems - all ages [Available from: https://www3.ha.org.hk/cancereg/pdf/overview/Summary%20of%20CanStat%202017.pdf.

21. Hong Kong Anti-Cancer Society. Cancer Screening Guidelines (Hong Kong) 2018 [Available from: https://www.hkacs.org.hk/en/screening_guideline.php.

22. Gomez-Rubio V, Lopez-Quilez A, editors. Empirical and full Bayes estimators for disease mapping. Proceedings of the International workshop on spatio-temporal modelling (METMA3) 2006 September 27 - 29; Pamplona, Spain. Spain: Universidad Publica de Navarra; 2006.

23. Waller LA, Carlin BP. Disease mapping. Chapman Hall CRC Handb Mod Stat Methods. 2010; 217-43.

24. Cramb S, Duncan E, Baade P, Mengersen K. Investigation of Bayesian Spatial Models. Brisbane, Australia.: Cancer Council Queensland and Queensland University of Technology (QUT); 2018.

25. Best N, Richardson S, Thomson A. A comparison of Bayesian spatial models for disease mapping. Stat Methods Med Res 2005 14: 35-39.

26. Bivand R, Altman M, Anselin L, Assuncao R, Berke O, Bernat A, et al. Spatial dependence: Weighting Schemes and Statistics 2014 [Available from: https://cran.r-project.org/web/packages/spdep/spdep.pdf.

27. Lee D. CARBayes: An R package for Bayesian spatial modeling with conditional autoregressive priors. J Stat Software 2013 55: 1-24.

28. Anselin L, Syabri I, Kho Y. GeoDa: An introduction to spatial data analysis Geogr Anal. 2006 38: 5-22.

29. Tsai WC, Kung PT, Wang YH, Huang KH, Liu SA. Influence of time interval from diagnosis to treatment on survival for oral cavity cancer: A nationwide cohort study. PLoS One. 2017 12:e0175148.

30. Perme M, Pavlic K. Nonparametric relative survival analysis with the R package relsurv J Stat Software. 2018 87:1-27.

31. Brenner H, Söderman B, Hakulinen T. Use of period analysis for providing more up-to-date estimates of long-term survival rates: empirical evaluation among 370,000 cancer patients in Finland. International Journal of Epidemiology. 2002 31: 456-62.

32. Fonseca EPD, Oliveira CDL, Chiaravalloti FN, Pereira AC, Vedovello SAS, Meneghim MC. Bayesian model and spatial analysis of oral and oropharynx cancer mortality in Minas Gerais, Brazil. Cien Saude Colet. 2018 23: 153-60.

33. Prado da Fonseca E, Cristina do Amaral R, Carlos Pereira A, Martins Rocha C, Tennant M. Geographical Variation in Oral and Oropharynx Cancer Mortality in Brazil: A Bayesian Approach. Int J Environ Res Public Health. 2018 15.

FIGURES

Figure 1

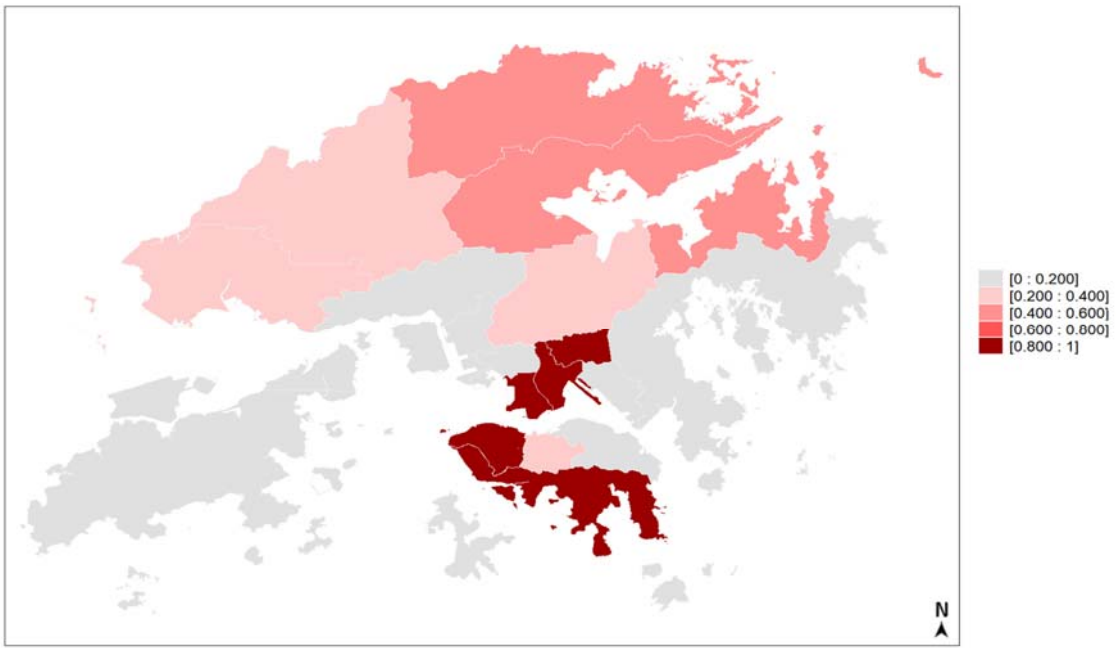


Figure 2

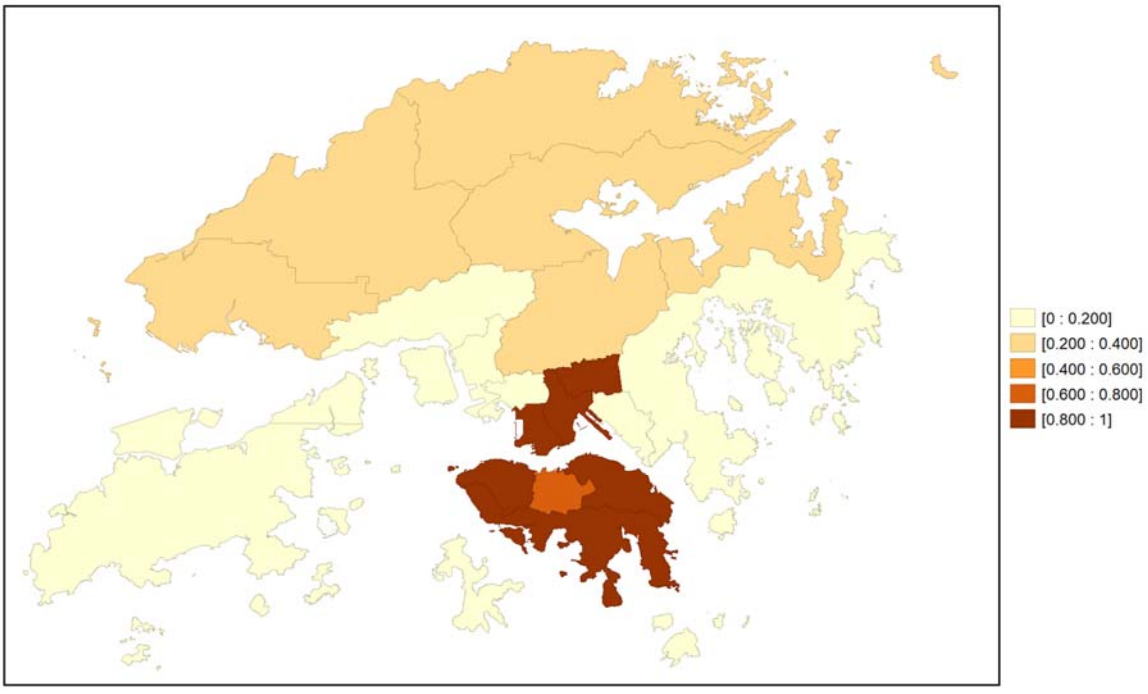


FIGURE LEGENDS

Figure 1: Bayesian mapping of oral cancer incidence within 18 DCDs in Hong Kong, highlighting posterior risk exceedance probability estimates [$Pr(\theta_h) > 1$].

Figure 2: Bayesian mapping of oral cancer mortality within 18 DCDs in Hong Kong, highlighting posterior risk exceedance probability estimates [$Pr(\theta_h) > 1$].

TABLE

Table 1: Median Standardized Incidence Ratios (SIR) and Median Standardized Mortality Rations (SMR) for 18 District Council Districts (DCD) in Hong Kong

Area	DCD	Estimated Source Population (≥20 years)	Incidence		Risk status[c]	Mortality		Risk status[c]
			Median SIR[a]	95% credible interval (LL, UL)[b]		Median SMR[d]	95% credible interval (LL, UL)[b]	
Hong Kong Island	Central and Western	208,923	2.302	2.036, 2.593	5	1.543	1.247, 1.888	5
	Eastern	469,548	0.950	0.836, 1.074	3	1.192	1.009, 1.396	4
	Southern	231,837	2.276	2.023, 2.549	5	1.526	1.249, 1.847	5
	Wan Chai	155,805	0.986	0.796, 1.205	3	1.229	0.944, 1.567	4
Kowloon	Kowloon City	351,765	1.171	1.027, 1.328	4	1.302	1.088, 1.549	4
	Kwun Tong	544,021	0.589	0.504, 0.675	1	0.696	0.569, 0.839	1
	Sham Shui Po	339,822	0.777	0.660, 0.906	2	0.790	0.628. 0.977	2
	Wong Tai Sin	358,785	1.171	1.027, 1.327	4	1.295	1.082, 1.542	4
	Yau Tsim Mong	290,381	1.164	1.005, 1.338	4	1.296	1.062, 1.568	4
New Territories	Islands	131,639	0.743	0.569, 0.950	2	0.755	0.528, 1.036	2
	Kwai Tsing	436,825	0.735	0.634, 0.847	2	0.741	0.600, 0.901	2
	North	259,972	0.982	0.831, 1.151	3	0.920	0.726, 1.151	3
	Sai Kung	387,644	0.608	0.511, 0.716	1	0.723	0.577, 0.890	2
	Sha Tin	552,602	0.974	0.869, 1.088	3	0.928	0.785, 1.087	3
	Tai Po	254,488	0.972	0.824, 1.139	3	0.917	0.727, 1.143	3
	Tseun Wan	268,981	0.748	0.623, 0.889	2	0.771	0.602, 0.966	2
	Tuen Mun	410,829	0.974	0.851, 1.108	3	0.944	0.777, 1.134	3
	Yuen Long	510,188	0.972	0.862, 1.092	3	0.950	0.800, 1.121	3

[a]SIR – Standardized Incidence Ratio; [b]LL – Lower limit, UL – Upper limit.

[c] Risk status classifiers were set according to their >10% and >30% difference from the overall Hong Kong SIR estimate and can be interpreted as 1 – very low, 2 – low, 3 – equivalent, 4 – high and 5 – very high.

[d]SMR – Standardized Mortality Ratio

APPENDIX 1

Sensitivity Analysis estimates for the Convolution and Leroux CAR models of the Hong Kong oral cancer incidence and mortality data

Estimates	**Convolution model**		**Leroux Model**	
	Incidence	Mortality	Incidence	Mortality
Median α (LL, UL)[a]	-0.01505 (-0.1229,0.0863)	-0.00676 (-0.1203,0.1032)	-0.0144 (-0.054,0.0241)	-0.00540 (-0.0614,0.0490)
pD[b]	17.069	15.196	19.471	16.254
$\overline{D(\theta)}$[c]	143.431	129.111	145.969	131.222
DIC[d]	160.499	144.307	165.440	147.476
WAIC[e]	155.860	142.50	165.373	147.606

[a]α - intercept; LL, UL – lower limit and upper limit of the 95% confidence interval

[b]pD – effective number of the parameters

[c]$\overline{D(\theta)}$ – posterior mean of the deviance

[d]DIC – Deviance information criteria (pD+ $\overline{D(\theta)}$)

[e]WAIC – Wantanabe-Akaike information criteria

Convolution CAR model

Taking the conventional standardized incidence and mortality ratio (θ_h) which represents the risk of oral cancer occurrence and mortality respectively within non-overlapping geographic units as the response variable:

$$\theta_h = \frac{y_h}{E_h} \qquad e_h = \frac{n_h [\sum_{h=1}^{N} y_h]}{[\sum_{h=1}^{N} n_h]}$$

{h = 1, 2, 3,…..,N}

Where y_h represents the count of observed cancer cases or cancer-related deaths within respective areas 'h' and E_h is the expected number of cases or deaths within matched geographical areas which is calculated from the population count at risk n_h.

According to the hierarchical structure for disease mapping models described by Best et al (26):

$y_h \sim$ Poisson $(e^{\hat{\theta}_h} E_h)$

where $\hat{\theta}_h$ is the log relative risk in area *'h'*.

$\hat{\theta}_h \sim \alpha + \beta X_{h\tau} + \eta$

where α is the intercept/overall risk effect, β is the coefficient of predictor variable X, and η represents the random effect modelled by convolution priors. Based on the previous socio-demographic description of individuals at risk of oral cancer in Hong Kong, the proportion of DCD residents above 60years was included to model the fixed effects. Both α and β were assigned vague normal distributions with mean (μ) 0 and variance (σ^2) of 10^6.

$\eta = \upsilon_h + v_h$

where υ_h and v_h represent the structured/spatial and unstructured random effects respectively. v_h is assigned a normal distribution while υ_h is modelled using an intrinsic CAR prior for spatial effects.

$v_h \sim N(0, \tau_v^2)$

where τ^2 is the inverse of the variance

$$\upsilon_h | \upsilon_{-h} \sim N\left(\frac{1}{\sum_k \omega_{hk}} \sum \upsilon_k \omega_{hk}, \frac{\sigma_u^2}{\sum_k \omega_{hk}}\right)$$

Where ω_{hk} represents a binary neighbourhood matrix to model the spatial closeness between the random effects.

$$\omega_{hk} \begin{cases} 1 & \text{value specified in the matrix if } h \text{ and } k \text{ DCDs are neighbours} \\ 0 & \text{value specified in the matrix if } h \text{ and } k \text{ DCDs are not neighbours} \end{cases}$$

Precision hyperparameters τ_v and τ_υ were also assigned uninformative gamma distributions. Different weakly informative priors were simulated to estimate these hyperparameters and sensitivity analyses was conducted to assess their individual effects on the smoothed risk estimates and the model's goodness-of-fit. Hyperprior distributions finally selected following sensitivity analyses are given below:

For Incidence model: $\tau_v \sim \Gamma(0.001, 0.001)$

$\tau_\upsilon \sim \Gamma(0.1, 0.1)$

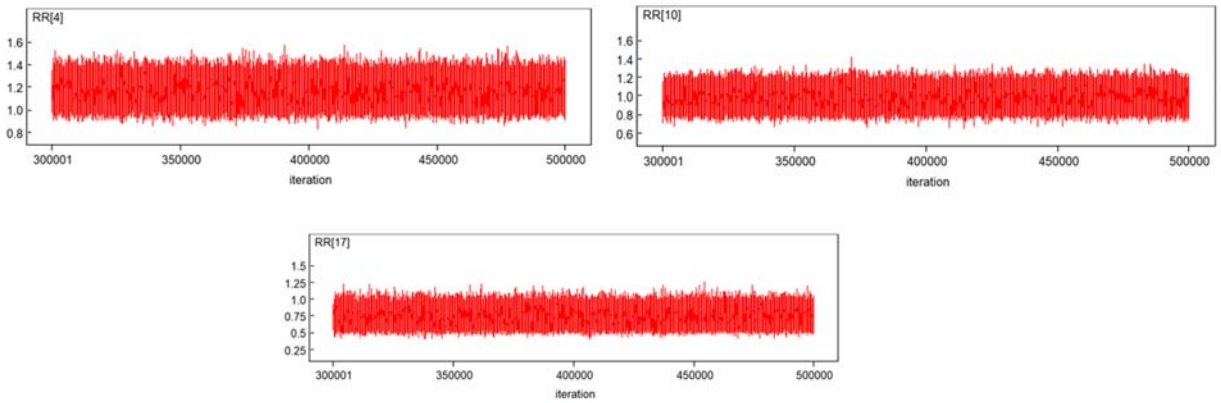
For Mortality model: $\tau_v \sim \Gamma(0.05, 0.005)$

$\tau_\upsilon \sim \Gamma(0.05, 0.005)$

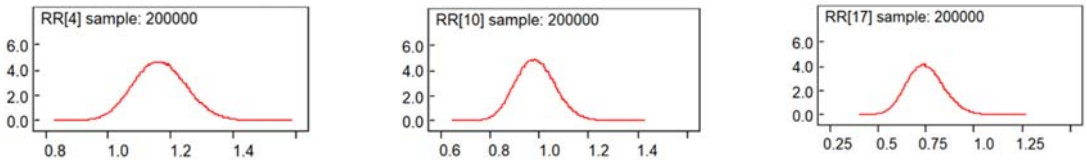
APPENDIX 2

Trace, density and autocorrelation plots indicating Markov chain convergence for the incidence relative risk estimates in three randomly selected DCDs.

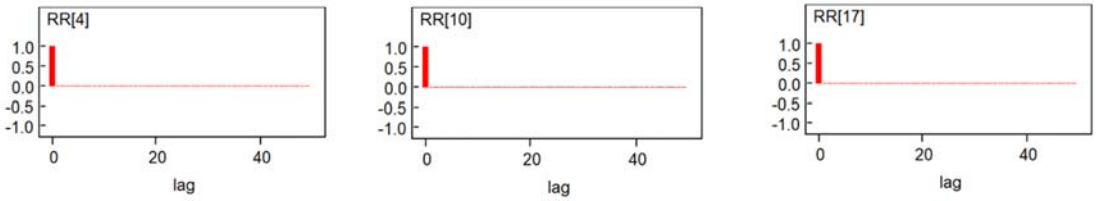
A. TRACE PLOTS



B. KERNEL DENSITY PLOTS



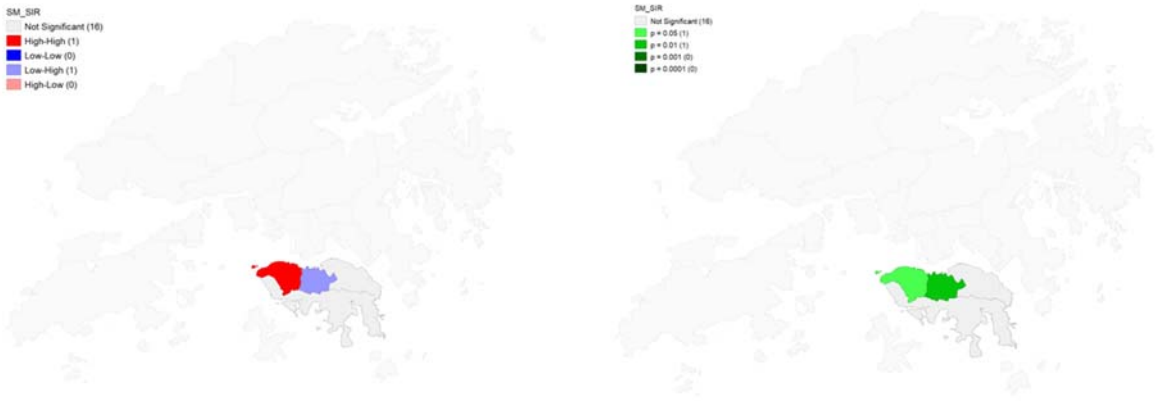
C. AUTOCORRELATION PLOTS



APPENDIX 3

Local indicator of spatial association (LISA) cluster maps and Significance maps for identification of incidence and mortality clusters in the DCDs in Hong Kong

A. INCIDENCE CLUSTER AND OUTLIER



B. MORTALITY CLUSTERS