Cancer Epidemiology xxx (2015) xxx–xxx

Contents lists available at ScienceDirect

Cancer Epidemiology

The International Journal of Cancer Epidemiology, Detection, and Prevention

journal homepage: www.cancerepidemiology.net

Inferring lung cancer risk factor patterns through joint Bayesian spatio-temporal analysis

Susanna M. Cramb [a,b,*], Peter D. Baade [a,c,d], Nicole M. White [b,e], Louise M. Ryan [f], Kerrie L. Mengersen [b,e]

[a] Cancer Council Queensland, Brisbane, Australia
[b] Mathematical Sciences School, Queensland University of Technology, Brisbane, Australia
[c] School of Public Health and Social Work, Queensland University of Technology, Brisbane, Australia
[d] Griffith Health Institute, Griffith University, Gold Coast, Australia
[e] Cooperative Research Centre for Spatial Information, Australia
[f] School of Mathematical Sciences, University of Technology, Sydney, Australia

ARTICLE INFO

Article history:
Received 9 October 2014
Received in revised form 31 January 2015
Accepted 3 March 2015
Available online xxx

Keywords:
Lung cancer
Risk factor
Tobacco smoking
Bayesian methods
Spatio-temporal analysis
Shared component model

ABSTRACT

Background: Preventing risk factor exposure is vital to reduce the high burden from lung cancer. The leading risk factor for developing lung cancer is tobacco smoking. In Australia, despite apparent success in reducing smoking prevalence, there is limited information on small area patterns and small area temporal trends. We sought to estimate spatio-temporal patterns for lung cancer risk factors using routinely collected population-based cancer data.

Methods: The analysis used a Bayesian shared component spatio-temporal model, with male and female lung cancer included separately. The shared component reflected lung cancer risk factors, and was modelled over 477 statistical local areas (SLAs) and 15 years in Queensland, Australia. Analyses were also run adjusting for area-level socioeconomic disadvantage, Indigenous population composition, or remoteness.

Results: Strong spatial patterns were observed in the underlying risk factor estimates for both males (median Relative Risk (RR) across SLAs compared to the Queensland average ranged from 0.48 to 2.00) and females (median RR range across SLAs 0.53–1.80), with high risks observed in many remote areas. Strong temporal trends were also observed. Males showed a decrease in the underlying risk across time, while females showed an increase followed by a decrease in the final 2 years. These patterns were largely consistent across each SLA. The high underlying risk estimates observed among disadvantaged, remote and indigenous areas decreased after adjustment, particularly among females.

Conclusion: The modelled underlying risks appeared to reflect previous smoking prevalence, with a lag period of around 30 years, consistent with the time taken to develop lung cancer. The consistent temporal trends in lung cancer risk factors across small areas support the hypothesis that past interventions have been equally effective across the state. However, this also means that spatial inequalities have remained unaddressed, highlighting the potential for future interventions, particularly among remote areas.

© 2015 Elsevier Ltd. All rights reserved.

1. Introduction

Due to its high incidence and low survival, lung cancer is the leading cause of cancer-related death in Australia [1]. More males are affected by this disease than females [1]. Most lung cancers are caused by cigarette smoking, accounting for around 65% of lung cancers among females and 90% among males [2]. Other modifiable risk factors include exposure to air pollution, radon, asbestos and certain heavy metals [3].

In the absence of effective early diagnostic tools or treatments for advanced lung cancer [4], preventing the initiation of lung cancer by reducing exposure to risk factors is vital. In particular, there has been much progress in reducing the prevalence of tobacco smoking in many developed countries [5]. Between 1964 and 2010, the percentage of Australians who smoked cigarettes

* Corresponding author at: Cancer Council Queensland, PO Box 201, Spring Hill, QLD 4004, Australia. Tel.: +61 7 36345350; fax: +61 7 32598527.
E-mail address: susannacramb@cancerqld.org.au (S.M. Cramb).

http://dx.doi.org/10.1016/j.canep.2015.03.001
1877-7821/© 2015 Elsevier Ltd. All rights reserved.

Please cite this article in press as: Cramb SM, et al. Inferring lung cancer risk factor patterns through joint Bayesian spatio-temporal analysis. Cancer Epidemiology (2015), http://dx.doi.org/10.1016/j.canep.2015.03.001

decreased from 43 to 15% [6], although the prevalence of smoking among females increased until around the late 1970s, when it started to decline [7]. Yet this smoking prevalence varies geographically, with people living in rural and disadvantaged areas more likely to smoke [1]. However these geographical data are often compromised by small numbers and a reliance on self-reported surveys. This limits the ability to understand small area patterns of smoking prevalence, particularly over time.

Given the lack of data on most risk factors at the spatial level, recent work has sought to model selected cancers jointly to extract spatial or spatio-temporal estimates of the common underlying risk factor components. Where high quality, population-based cancer registry data are available, this can be used to obtain objective risk factor estimates. When a cancer has similar risk factors for both sexes, but a differential impact across space and/or time, there may be benefit in jointly modelling one cancer type and dividing into sex-specific components, e.g. male and female lung cancer. This joint modelling is often conducted using a shared component model.

The premise of the shared component model, as first proposed by Knorr-Held and Best [8], was to jointly model the relative risk by dividing into separate components, including one common to both diseases (e.g. representing the underlying risk factor exposure), as well as residual variation components (one for each disease). This enables information to be borrowed between diseases. In this model the shared component acts as a surrogate for spatially structured unobserved risk factors common to both diseases [8]. The model has been extended by incorporating covariates [9], adjusting the number of components [9], increasing the number of diseases [10], and including temporal trends [11,12]. The joint modelling of multiple cancers at the spatial or spatio-temporal level has commonly been applied within a Bayesian context [8,11].

When there is only one shared component in these models, this component provides an estimate of all the risk factors common to the included diseases. However, when a particular risk factor is prominent in developing disease, such as tobacco smoking with lung cancer, underlying risk estimates are likely to reflect the most prominent risk factor.

Our aims were to apply Bayesian spatio-temporal shared component models to routinely collected, population-based male and female lung cancer data to:

1. Infer the spatio-temporal patterns of underlying lung cancer risk factors in Queensland.
2. Determine how known influences (socioeconomic, remoteness and Indigenous status) impact on the modelled underlying risk factor patterns.
3. Identify geographical areas where the temporal underlying risk factor pattern differed from the pattern for total Queensland.

2. Methods

2.1. Data

Lung cancers diagnosed among Queensland residents between 1997 and 2011 were sourced from the Queensland Cancer Registry [13], a population-based cancer registry with high-quality data covering the entire state of Queensland. Australian legislation requires this Registry to be notified of every invasive cancer diagnosed in a Queensland resident, excluding only keratinocytic skin cancers. Ethical approval was obtained from Queensland Health (approval number: HREC/09/QHC/25).

Details about patients' usual residence at diagnosis was provided at the statistical local area (SLA) level. Geocoding was used to match the residence at diagnosis to the 2006 SLA definition, thus overcoming limitations of changing geographical boundaries over time. In 2006 there were 478 SLAs, with a median population of 5723.

Population estimates based on the 2006 SLA boundaries were obtained from the Australian Bureau of Statistics, for each SLA, year and 5-year age groups up to 85+ years. Due to zero population counts in one SLA for several years during the time period of interest, only 477 SLAs were used in our analyses (population range in 2006: 78–74,804).

Each SLA was assigned a value for area-level socioeconomic disadvantage (3 categories (advantaged: top 20%, middle SES: middle 60%, disadvantaged: lowest 20%), defined using the Index of Socioeconomic Advantage and Disadvantage (IRSAD) from the Australian Bureau of Statistics' Socioeconomic Indexes for Areas (SEIFA), remoteness (urban (major city), regional (inner/outer regional) and remote (remote/very remote) based on the Accessibility/Remoteness Index of Australia+), and Indigenous population (2 categories based on 2006 census data: $<10\%$ or $\geq 10\%$).

2.2. Model

Most shared component models use a standard Poisson likelihood, as is appropriate for rare and non-contagious diseases. However, when area-specific count data are particularly sparse, an alternative formulation allowing for excess zero counts may be preferred. Therefore, we extended previous approaches by incorporating and comparing alternate distributions for the counts within the shared component framework. Specifically, we compared four alternative variants of the Poisson count distribution [14,15]:

1. Poisson $O_{dij} \sim \text{Poisson}(\rho_{dij}E_{dij})$
2. Negative binomial $O_{dij} \sim \text{Poisson}(x_d\rho_{dij}E_{dij})$ where $x_d \sim \text{Gam-Gamma}(r_d, r_d)$
3. Zero-inflated Poisson (ZIP) $O_{dij} \sim \text{Poisson}((1-u_{dij})\rho_{dij}E_{dij})$ if $O_{dij} > 0$
4. Poisson hurdle $O_{dij} \sim \text{Poisson}(((1-u_{dij})/(1-\exp(-\rho_{dij}E_{dij}))\rho_{dij}E_{dij})$ if $O_{dij} > 0$

where O_{dij} are the observed lung cancer counts for each sex $d = 1,2$ (representing males and females, respectively), $i = 1, 2, \ldots, 477$ areas and $j = 1, 2, \ldots, 15$ years, ρ_{dij} is commonly referred to as the relative risk [16] and E_{dij} represents expected counts. To enable comparisons over time, the expected counts were calculated using the sex- and age-specific Queensland lung cancer incidence rates in 1997–1999. In the negative binomial model, here expressed as a Poisson-gamma mixture for comparability, r_d is the sex-specific overdispersion parameter, which forms the shape and scale parameters in the gamma distribution, while the u_{dij} is the probability of zero in the ZIP and hurdle models.

The Poisson hurdle model separates the zeros from anything above zero, modelling counts under a truncated Poisson distribution. The ZIP model can be considered a special type of hurdle model. Here the zero counts are separated into excess (those above what is expected under a Poisson distribution) and non-excess zeros (those expected under a Poisson distribution).

Using a modified version of the shared component model from Richardson et al. [11], the log relative risk for each of these models was then expressed as:

$$\log(\rho_{dij}) = \alpha_d + \mu_{dij}$$

The sex-specific intercept is given by α_d, while the space-time structure is modelled through μ_{dij}, which represents exposure to the risk factors for lung cancer, here referred to as the underlying risk factor component.

Please cite this article in press as: Cramb SM, et al. Inferring lung cancer risk factor patterns through joint Bayesian spatio-temporal analysis. Cancer Epidemiology (2015), http://dx.doi.org/10.1016/j.canep.2015.03.001

Download English Version:

https://daneshyari.com/en/article/10897321

Download Persian Version:

https://daneshyari.com/article/10897321

Daneshyari.com