Colorectal cancer mortality in Mato Grosso, Brazil: Spatiotemporal trends

Christiane Maria Meurer Alves,¹ Vitor Gabriel Barra Souza,² Ronaldo Rocha Bastos²

¹Universidade Federal de Juiz de Fora, Programa de Pós Graduação em Saúde Coletiva; ²Universidade Federal de Juiz de Fora, Departamento de Estatística, São Pedro - Juiz de Fora, Minas Gerais, Brasil

Abstract

Mortality due to colorectal cancer is increasing in Brazil, but an organised approach to screening and prevention is lacking. Considering the importance of this disease, the present study examines recent mortality trends of colorectal cancer mortality in the meso- and microregions in the state of Mato Grosso with the objective of analysing spatiotemporal trends to help guide the allocation of health services related to this type of cancer. Mortality data from the Brazilian national public health system from 1996 to 2015 were analysed investigating spatiotemporal trends using Conditional Autoregressive (CAR) models, a class of Bayesian hierarchical models that rely on Markov Chain Monte Carlo (MCMC) simulations. Convergence issues arose with several types of CAR models, but notably not with the linear variant, which models the mortality within each spatial region as a linear function of time. Men and women of all ages displayed higher and increasing mortality rates in the Cuiabá and Rondonópolis microregions. Additional regions of increasing mortality were found for specific age and gender strata. It was concluded that spatiotemporal modelling is a useful tool for the characterisation of diseases, including cancer, which are influenced by several factors and need to be monitored over space and time. The combination of spatial and temporal analyses of mortality shown in this paper unveils important information regarding the small areas dynamics, which may guide discussions, regulation and application of decentralised public health policies.

Introduction

Colorectal cancer is hereditary in about 20% of the cases, most commonly related to familial adenomatous polyposis or hereditary nonpolyposis colorectal cancer (Lynch syndrome) (Lynch and Chapelle, 2003; Sinicrope, 2018). The remaining 80% occur sporadically and are presumed to be associated with alcohol consumption, red meat intake, smoking, obesity and diabetes. Physical activity has been shown to be a protective factor (Huxley et al., 2009; Johnson et al., 2013; World Cancer Research Fund/American Institute for Cancer Research, 2018).

Studies on the incidence and mortality of colorectal cancer have shown disparate rates across the world, with Brazil being one of the countries where these rates have been on the rise. The discrepant incidence and mortality patterns across different countries are influenced not only by geographic location, but also by socioeconomic changes, age, and gender. Fighting colorectal cancer requires knowledge of the strategic factors that may play a role in the distribution of the disease (Arnold et al., 2017; Douaiher et al., 2017). Mapping geographic patterns of disease incidence and mortality is important for understanding the dynamics of complex diseases, such as cancer, which are partly influenced by regional factors — be they related to healthcare, socioeconomic, or environmental characteristics — and cannot be sufficiently explained by individual factors alone. This is increasingly more relevant as data are collected over ever more finely partitioned areas, enabling more effective, directed health interventions to be implemented (Ugarte et al., 2012; Roquette et al., 2017). Studies of spatial patterns may also provide important information on the aetiology of diseases. Recent investigations of colorectal cancer in Brazil have found a regional component associated with mortality and incidence trends. It has been suggested that the overall increase in both indicators may be, in part, a consequence of improvements in the quality and reliability of the medical records, and partly a result of unfavourable lifestyles compounded with difficult access to diagnosis and treatment (Oliveira and Rêgo, 2016; Oliveira et al., 2018). Indeed, death registries in the country rose from an 80% coverage (1980-1991) to 95% in 2000-2010, at the same time that poorly defined causes decreased in occurrence (Lima and Queiroz, 2014). However, differences between spatially divergent regions remain little investigated. When studying diseases, spatially structured models that account for correlations between random effects and incorporate neighbourhood structures are able to more closely represent reality and hence capture the geographic dependency of variable of interest. Generalised Linear Mixed Models (GLMM) have been used to map diseases, more specifi...
cally Conditional Autoregressive (CAR) models (Ugarte et al., 2012; Jürgens et al., 2013; Kang et al., 2016). The importance of spatial analyses highlighted earlier has been attested by the CAR models applied to areas with a small number of cases per time unit (Etxeberria et al., 2014).

Situated in the Brazilian Midwest, Mato Grosso is the third largest state in the country, featuring a relatively low demographic density and serving as what might be considered the country’s breadbasket. Mato Grosso has a high overall Human Development Index (Governo de Mato Grosso, 2019), but marked differences between municipalities are known to exist (Cócaro et al., 2016). The present study explores the distribution of colorectal cancer throughout this highly heterogeneous state and tests the utility of CAR models in this and potentially similar endeavours.

Material and Methods

Mortality and population composition data for the state of Mato Grosso were obtained from the Mortality Information System (SIM) of the national public health system (Brasil, 2020). Mortality was recorded according to place of residence of the individuals whose data we investigated. For public health purposes, the state is divided into five mesoregions (North, Northeast, Southwest, Centre-South and Southeast) and 22 microregions (Aripuanã, Alta Floresta, Colider, Parecis, Arinos, Alto Teles Pires, Sinop, Paranaína, Norte Araguaia, Canarana, Médio Araguaia, Alto Guaporé, Tangaçu da Serra, Jauru, Alto Paraguai, Rosário Oeste, Alto Pantanal, Cuiabá, Primavera Leste, Rondonópolis, Tesouro, Alto Araguaia), as shown in Figure 1. Position and boundaries of these regions is shown in a map provided by the Brazilian Institute of Geography and Statistics (IBGE, 2019) (Figure 1).

Deaths between 1996 and 2015, falling into the C18 (colon cancer), C19 (rectosigmoid junction cancer) and C20 (rectal cancer) categories of the International Classification of Diseases, 10th edition (ICD-10), were selected. This period was selected because, prior to 1996, ICD-10 codes were not required to be reported, and past 2015 no data were available at the time of the study. The availability of ICD-10 codes is particularly important because anal cancer, which has different properties from colorectal cancer, is classified under a specific code (C21), which allows it to be separated from other types of cancer (WHO, 1995). Anal canal cancer is strongly associated with Human Papillomavirus (HPV) infection and appear usually as epidermoid carcinoma, with is histologically different from the other cancers mentioned.

The available data were initially grouped by mesoregion and categorised into three age groups (<45 years old, 45–64 years old, and ≥65 years old). Spatial and temporal analyses were then conducted. The data were afterwards partitioned by microregion and categorised into two age groups (<45 years old and ≥45 years old); similar analyses were then performed. The different age classification criteria between the two spatial partitioning schemes arose from the necessity to pool what would have been two tiny 45–64 and ≥65 age groups, with respect to the low mortality numbers of each kind in certain microregions, adding the slightly bigger ≥45-
year age group with more stable statistical properties. The cutoffs between the age classes were chosen to allow differentiation of patterns in younger and older individuals. In all cases, death at unknown age was excluded from the mortality calculations. Finally, the data were also grouped by gender as a third subdivision. Death counts were originally obtained as annual values and were kept as such for the analyses.

Stratified spatiotemporal analyses were performed in two steps: exploratory analysis followed by modelling based on the full set of observations for each age, gender, and combined age and gender stratum under each geographic discretionary scheme. First, the presence of spatial autocorrelation in the data was assessed through Moran’s I statistic. If the null hypothesis of no spatial autocorrelation could be rejected at a 5% significance level for at least one of the years under study, spatiotemporal modelling of the data was deemed to be in order. The aim of this initial procedure was exploratory in nature, serving as the basis for hypothesis generation, therefore no adjustment for multiple testing was carried out (Bender and Lange, 2001). Various CAR models, fitted via the CARBayesST package (Lee et al., 2018) for R (The R Project for Statistical Computing, 2017), were then investigated. These spatiotemporal models form a class of Bayesian hierarchical models based on Markov Chain Monte Carlo (MCMC) simulations. The general formula for the spatiotemporal model used was that of Lee et al. (2018):

\[
Y_{kt} | \mu_{kt} \sim f(Y_{kt} | \mu_{kt}, \nu^2) \quad \text{for } k=1,...,K, \ t=1,...,N; \\
\mu_{kt} = X_{kt} \beta + O_{kt} + \psi_{kt}; \\
\beta \sim N(\mu_\beta, \Sigma_\beta)
\]

where \( Y_{kt} \) is the matrix of response data (death counts), with one row per area unit \( k \) and one column per time period \( t \); \( \mu_{kt} \) the matrix of expected values \( \mu_{kt} = E(Y_{kt}) \); \( f(Y_{kt} | \mu_{kt}, \nu^2) \) a distribution of the binomial, Gaussian or Poisson families, with the parameter \( \nu^2 \) only applicable in the Gaussian case; \( g \) an invertible link function, respectively the logit, identity or natural logarithm functions for the above distribution families; \( X_{kt} \) the vector of \( F \) known covariates for each \( k \) and \( t \), including a leading unit entry for the intercept term; \( O_{kt} \) the vector of covariate regression parameters, including an intercept term; \( \psi_{kt} \) the matrix of latent components corresponding to spatio-temporally autocorrelated random effects; and \( \mu_\beta \) and \( \Sigma_\beta \) the mean and diagonal variance matrix, respectively, of the assumed multivariate Gaussian prior. Due to the very nature of (annual) death counts, the response variable was modelled by a Poisson distribution. (Lee et al., 2018).

The specification of the structured random spatial effect characterised an intrinsic CAR model, in which adjacent areas are assumed to have more similar mortality risk (Besag et al., 1991). Such spatial correlation is controlled by a non-negative, symmetric matrix \( W = W_{kj} \) representing the proximity of units. As is often adopted, the matrix \( W \) is binary where \( W=1 \) if there is a common boundary and \( W=0 \) if not (Lee et al., 2018). Any given region is defined as having no common boundaries with itself.

In order to predict the response variable, CAR models use a latent component for each spatial and temporal unit. Latent components can be in one of several competing forms, each with its theoretical advantages and potential indications. We investigated the linear (CARlinear), ANOVA-like (CARAanova), autoregressive (CARar) and discrete-level autoregressive, or localised (CARlocalised) variants, as implemented in the CARBayesST package, each representing a different spatiotemporal structure (Table 1).

There is no consensus as to the number of iterations in the MCMC simulation process (Hamra et al., 2013), but there exist numeric metrics such as the Geweke diagnostic test that can provide guidance on a suitable number (Lee et al., 2018). A total of 110,000 iterations was found to be satisfactory, with the first 10,000 being systematically discarded as the burn-in period of the

<table>
<thead>
<tr>
<th>Model</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear</td>
<td>( \psi_{kt} = \beta_1 + \varphi_k + (\alpha + \delta_k) t )</td>
</tr>
<tr>
<td></td>
<td>where ( \varphi_k ) modifies the intercept ( \beta_1 ) for each areal unit, ( \delta_k ) the temporal slope ( \alpha ) for each areal unit, and ( t ) is the linear temporal trend covariate:</td>
</tr>
<tr>
<td></td>
<td>( \beta = \frac{-1}{N} \sum_{t=1}^{N} t )</td>
</tr>
<tr>
<td>ANOVA-like</td>
<td>( \psi_{kt} = \varphi_k + \delta_k + \gamma_{kt} )</td>
</tr>
<tr>
<td></td>
<td>where ( \gamma ) and ( \delta ) are the common sets of spatial and temporal random effects, respectively, and ( Y ) a set of independent space–time interactions.</td>
</tr>
<tr>
<td>Autoregressive</td>
<td>( \psi_{kt} = \varphi_{kt} )</td>
</tr>
<tr>
<td></td>
<td>where ( \psi ) is the matrix of random effects, which evolve over time via a multivariate first-order autoregressive process.</td>
</tr>
<tr>
<td>Localised</td>
<td>( \psi_{kt} = \lambda(Z_{kt}) + \varphi_{kt} )</td>
</tr>
<tr>
<td></td>
<td>where ( Z_{kt} ) is the cluster to which a given space–time unit is assigned and ( \lambda() ) a piecewise constant clustering or intercept component, assuming a constant value for each cluster.</td>
</tr>
</tbody>
</table>
Markov chain. The convergence of the chain to a stationary distribution was assessed with tools such as Geweke’s diagnostic and plots of the model parameters over the iterations, as well as plots of the autocorrelation of each parameter at various lags (Hamra et al., 2013; Lee et al., 2018). Finally, when an equilibrium distribution of the parameters was deemed to have been reached, informative spatiotemporal maps were produced, showing changes in the spatial distribution of deaths due to colorectal cancer over time.

**Results**

From 1996 to 2015, there were 875 deaths due to colorectal cancer among men and 817 among women according to ICD-10 codes C18, C19, and C20 in the state of Mato Grosso, with respectively 49% and 47% of these cases occurring in individuals above 65 years of age.

The preliminary analyses with Moran’s I statistic revealed, under geographic partitioning by mesoregions, two (out of six) combinations of age group and gender, or strata, which lacked evidence against the null hypothesis of no spatial autocorrelation. While this would not strictly preclude further investigation with spatiotemporal modelling, the existence of only five mesoregions in Mato Grosso meant that this course of investigation could at best lead to only coarse results. Therefore, we decided to pursue only the partitioning by microregions for the remainder of the analyses.

In this case again, two (out of four) combinations of age group and gender did not show evidence against lack of spatial autocorrelation in at least one observed year. Still, especially given that it is known from the literature that strong disparities between municipalities exist in Mato Grosso (Cócaro et al., 2016), we could not rule out the existence of spatial autocorrelation in these two strata, and therefore proceeded to the fitting of CAR models.

The models with a linear latent component structure (CARlinear) converged and their results were accepted as meaningful, showing the best fit to our data and provided also more relevant interpretation of the micro-region patterns. Plots of autocorrelation of the parameters for all the models fitted with the full set of observations are displayed in Figure 2, which shows the superiority of the model with the linear latent component. The CARlinear structure treats the latent component as a linear function of time within each spatial unit (here, microregion), and is thus very suitable for identifying the regions with either increasing or decreasing mortality rates.

The spatiotemporal maps resulting from the models considering men (Figure 3) or women (Figure 4) show a strong growth in mortality centred at Cuiabá, which contains the capital city of the state, and spreading out to the surrounding microregions. This growth seemed particularly marked among men in the direction of Tesouro and Médio Araguaia.

In the <45-years age group, once again, in addition to the areas with the greatest mortality increase, namely Cuiabá and Rondonópolis, the microregions of Médio Araguaia and Sinop stand out in relation to their neighbouring regions (Figure 5).

The model fitted for men in the younger age group shows an increase in mortality over time in all microregions, more pronouncedly in the south-eastern, centre-southern and south-western mesoregions than in the north (Aripuanã, Alta Floresta, Arinos,
Figure 3. Linear CAR model for microregions with respect to men in the state of Mato Grosso, 1996–2015.

Figure 4. Linear CAR model for microregions with respect to women in the state of Mato Grosso, 1996–2015.
Alto Teles Pires e Colíder) and the Northeast (mostly Norte Araguaia). For men in the older age group, mortality rates are estimated to be on the rise mainly in the regions of Cuiabá, Rondonópolis, Médio Araguaia and Primavera Leste.

The model fitted for women aged less than 45 years shows a progressive mortality increase throughout the state in general, with an important emphasis on Cuiabá and apparently being less pronounced in Colíder and Arinos at the end of the period. For women aged over 45 years, the model shows increased mortality in the regions of Cuiabá, Rondonópolis, Parecis, Alto Teles Pires and Primavera Leste. The spatial mortality profile is very similar to that seen among men in the same age group, except for the addition of Parecis as an increased mortality area.

Discussion

The finding of increased mortality by colorectal cancer in the state of Mato Grosso matches reports by previous studies of increasing mortality in Brazil as a whole and also in the Brazilian Midwest more specifically (Douaiher et al., 2017; Oliveira et al., 2016; Oliveira e Rêgo, 2016; Oliveira et al., 2018). The state of Mato Grosso has 85 native Indian lands (IBGE, 2010). In the 2010 Census, 1.4% of the population of the state declared itself of Indian descent, the fifth largest participation among all twenty-seven Brazilian states. The city of Campinápolis, for instance, located in the Canarana microregion (the northeastern mesoregion), is among the ten Brazilian cities with the largest number of native Indian population living in rural areas (Fundação Nacional do Índio, 2018). One could argue that access to health services for prevention, diagnosis and treatment can be made difficult for the native Indian population in delimited areas, but we saw a predominance of mortality decrease in the North and Northeast among women and in the Northeast among men. Differences in risk factors may be influencing the profile we found.

More favoured regions, with a higher MHDI (Municipal Human Development Index), show higher colorectal cancer mortality, when considering cities with more than 100,000 inhabitants (Parreira et al., 2016). The state of Mato Grosso is considered to have a high MHDI (0.700-0.799), as stated earlier, but its northeastern mesoregion is a low-MHDI area (Atlas do Desenvolvimento Humano, 2013). This mesoregion displayed an increase in mortality among both women and men under 45 years of age, especially men.

As in other Brazilian Midwestern states, Mato Grosso received a lot of immigrants from other regions of the country, mainly with the expansion of agriculture during the 1970s and 1980s (Oliveira et al., 2016). Although the state has a population density of 3.36 inhabitants/km², its territorial distribution is quite variable and mainly concentrated in the microregion of Cuiabá, which includes the cities of Cuiabá (157,66 inhab/km²), Várzea Grande (240,98 inhab/km²), and Rondonópolis (47 inhab/km²).

The 2013 Pesquisa Nacional de Saúde (National Health Survey) evidenced that the diet of the Mato Grosso population is rich in fat meat or chicken, second in frequency of consumption among Brazilian states, whereas the consumption of fruit and vegetables is close to the national average. Alcohol consumption is also close to the national average (Claro et al., 2015; IBGE, 2013). The proportion of adults in the state classified as insufficiently active (52.9%) is slightly higher than the national average (46%) (IBGE, 2013). This information is an overall estimate for the state without considering how subpopulations are geographically spread in meso or microregions, but it is an indication of a population that has adopted risk behaviour in its lifestyle.

Another feature of the state is that it was ranked the greatest pesticide consumer in the country in 2012 (Pignati et al., 2014). A recent study correlated environmental factors with increased risk of colon cancer and identified an increase in mortality from this disease in Brazil concurrently with increased sales of pesticides in the country (Martin et al., 2018). It is interesting to note that the south-western mesoregion, which showed increasing mortality in the model for men under 45 (together with Cuiabá, Sinop and Rondonópolis), does not have any treatment services agreed-upon between the municipalities and the state, unlike other mesoregions. Another important point is the apparently slow increase in mortality for the same age and gender stratum in the two northern microregions of Aripuanã and Alta Floresta. These areas are locat-
ed far from available treatment areas (Ministério da Saúde, 2012), but are among the areas with the lowest pesticide use in relation to the rest of the state (Pignati et al., 2014). The fact that we considered deaths by place of residence implies that a decrease in mortality in areas far from treatment sites may indicate a real decrease, but also it could also be explained by underreporting of cases due to the difficult access to diagnosis and treatment.

Risk factors such as diet of red meat, fats and processed foods, obesity, smoking, and socioeconomic class were not included in the autoregressive model, nor were treatment-related factors. Still, this study showed potential microregion peculiarities in different gender and age strata, due to the small number of cases in certain regions. For men over 45, an apparent increase in mortality is seen in Alto Teles Pires, Jauru and Tangará da Serra. These are areas that offer diagnostic and/or treatment services, which may facilitate the recognition of the disease and the search for treatment.

The features of each microregion could be better evidenced with linear CAR models, further demonstrating that such models can be useful in spatiotemporal analyses. The inhabitants of Mato Grosso could benefit from interventions favouring the diagnosis and treatment of colorectal cancer in regions other than the South and Southeast, where treatment services are currently concentrated (Ministério da Saúde, 2012). The distance from treatment centres, which makes access to early diagnosis and treatment difficult, has been associated with regional differences in colorectal cancer mortality elsewhere in Brazil (Souza et al., 2014). The uneven distribution of specialised cancer services and prevention and diagnosis infrastructure has been linked to higher mortality by other authors (Oliveira et al., 2018).

Conclusions

Despite possible limitations related to the nature of the data and to a lack of more specific information on individual behaviours, the findings discussed in this text may provide guidance for the formulation, regulation and application of state-wide decentralised public health policies, whether in the form of population awareness campaigns about the disease or of better structuring and distribution of prevention, diagnosis and treatment services. Importantly, CAR models make it possible to expose priority areas for attention by policy makers allowing the design of special interventions.

References