

UNIVERSIDADE FEDERAL DE SANTA MARIA
CENTRO DE TECNOLOGIA
PROGRAMA DE PÓS GRADUAÇÃO EM ENGENHARIA DE PRODUÇÃO

Alisson Castro Barreto

**MORTALIDADE POR CÂNCER DE PRÓSTATA E MAMA NO
BRASIL: UMA ANÁLISE SOB A ÓTICA DA ECONOMETRIA
ESPACIAL**

Santa Maria, RS
2021

Alisson Castro Barreto

**MORTALIDADE POR CÂNCER DE PRÓSTATA E MAMA NO BRASIL: UMA
ANÁLISE SOB A ÓTICA DA ECONOMETRIA ESPACIAL**

Dissertação apresentada ao Curso de Pós-Graduação em Engenharia de Produção, Área de Concentração Gerência da Produção, da Universidade Federal de Santa Maria (UFSM, RS), como requisito parcial para obtenção do grau de **Mestre em Engenharia de Produção**.

Orientador: Prof. Dr. Adriano Mendonça Souza

Santa Maria, RS
2021

Barreto, Alisson Castro
Mortalidade por câncer de próstata e mama no Brasil:
uma análise sob a ótica da econometria espacial / Alisson
Castro Barreto.- 2021.
80 p.; 30 cm

Orientador: Adriano Mendonça Souza
Dissertação (mestrado) - Universidade Federal de Santa
Maria, Centro de Tecnologia, Programa de Pós-Graduação em
Engenharia de Produção, RS, 2021

I. Econometria Espacial I. Souza, Adriano Mendonça
II. Título.

Sistema de geração automática de ficha catalográfica da UFSM. Dados fornecidos pelo autor(a). Sob supervisão da Direção da Divisão de Processos Técnicos da Biblioteca Central. Bibliotecária responsável Paula Schoenfeldt Patta CRB 10/1728.

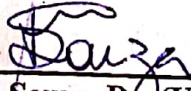
Declaro, ALISSON CASTRO BARRETO, para os devidos fins e sob as penas da lei, que a pesquisa constante neste trabalho de conclusão de curso (Dissertação) foi por mim elaborada e que as informações necessárias objeto de consulta em literatura e outras fontes estão devidamente referenciadas. Declaro, ainda, que este trabalho ou parte dele não foi apresentado anteriormente para obtenção de qualquer outro grau acadêmico, estando ciente de que a inveracidade da presente declaração poderá resultar na anulação da titulação pela Universidade, entre outras consequências legais.

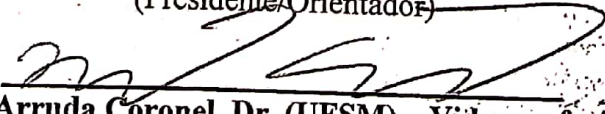
Alisson Castro Barreto

MORTALIDADE POR CÂNCER DE PRÓSTATA E MAMA NO BRASIL: UMA ANÁLISE SOB A ÓTICA DA ECONOMETRIA ESPACIAL

Dissertação apresentada ao Curso de Pós-Graduação em Engenharia de Produção, Área de Concentração em Gerência de Produção, da Universidade Federal de Santa Maria (UFSM, RS), como requisito parcial para obtenção do título de **Mestre em Engenharia de Produção**.

Aprovado em 18 de fevereiro de 2021:


Adriano Mendonça Souza, Dr. (UFSM) – Videoconferência
(Presidente/Orientador)


Daniel Arruda Coronel, Dr. (UFSM) – Videoconferência


Liane Werner, Dra. (UFRGS) – Videoconferência

**Santa Maria, RS
2021**

AGRADECIMENTOS

A concretização de mais uma etapa, neste início de jornada na Academia, não seria possível sem a contribuição e dedicação de diversos atores. Por isso, agradeço:

- a Deus, pelo dom da vida e pela sua infinita graça e misericórdia;*
- a meus pais Anibal Barreto e Suelmy Barreto, pela educação e incentivo;*
- a minha irmã Suani dos Anjos, pela amizade e apoio;*
- a minhas professoras da graduação, Eunice Paraguassu e Rita Rafael que foram minhas maiores apoiadoras e incentivadoras;*
- ao meu orientador Adriano Souza, pela amizade, acolhimento e principalmente pelos ensinamentos;*
- aos membros do Laboratório de Análise e Modelagem Estatística da UFSM, pelo companheirismo e aprendizado;*
- e ao Programa de Pós-Graduação em Engenharia de Produção da UFSM, pela oportunidade de desenvolvimento profissional e acadêmico.*

RESUMO

MORTALIDADE POR CÂNCER DE PRÓSTATA E MAMA NO BRASIL: UMA ANÁLISE SOB A ÓTICA DA ECONOMETRIA ESPACIAL

AUTOR: Alisson Castro Barreto
ORIENTADOR: Adriano Mendonça Souza

O câncer tornou-se um problema de saúde pública no Brasil, o que se evidencia pelo aumento de suas taxas de incidência e mortalidade por institutos de pesquisa especializados no assunto como o Instituto Nacional de Câncer e Agência Internacional de Pesquisa em Câncer. No Brasil, o câncer de Próstata (CaP) e Câncer de Mama (CaM) são atualmente os tipos de cânceres mais frequentes em homens e mulheres, respectivamente. Por isso, com o objetivo de determinar a distribuição espacial da mortalidade por CaP e CaM nas microrregiões brasileiras e verificar o seu inter-relacionamento com fatores sociodemográficos, econômicos, ambientais e de infraestrutura, fez-se uso das técnicas da Econometria Espacial: Análise Exploratória de Dados Espaciais – AEDE e Modelagem da dependência espacial. Foram escolhidas variáveis para representar os fatores, de acordo com literatura relevante e disponibilidade de dados, sendo estas: unidades de referência oncológica, número de médicos especialistas em oncologia, taxa de envelhecimento, taxa de urbanização, quantidade média de pesticidas, Produto Interno Bruto e número de mamógrafos. A base de dados utilizada é proveniente do Departamento de Informática do Sistema Único de Saúde do Brasil – DATASUS e do Instituto Brasileiro de Geografia e Estatística – IBGE. Os resultados indicaram autocorrelação espacial positiva para a taxa de mortalidade por câncer de próstata ($I = 0,4537$; $p \leq 0,001$) e câncer de mama ($I = 0,4842$; $p \leq 0,001$). A taxa de mortalidade por câncer de próstata apresentou correlação positiva com a taxa de envelhecimento e taxa de urbanização. A taxa de mortalidade por câncer de mama mostrou-se positivamente correlacionada com número de médicos especialistas, unidades de referência, taxa de urbanização e taxa de envelhecimento, mas negativamente correlacionada com número de mamógrafos. Houve indicativo de que a maior expectativa de vida e o estilo de vida, decorrente da urbanização possuem influência no aumento da mortalidade por câncer de próstata e mama no Brasil. Diante do exposto, considera-se que os objetivos gerais e específicos da pesquisa foram atingidos e a análise espacial da mortalidade por câncer possibilitou encontrar correlações estatísticas que merecem investigação epidemiológica para subsidiar ações no combate da mortalidade por câncer de próstata e mama em mulheres no Brasil.

Palavras-chave: Econometria Espacial. Mortalidade. Câncer. Autocorrelação Espacial.

ABSTRACT

PROSTATE AND BREAST CANCER MORTALITY IN BRAZIL: SPACIAL ECONOMETRY'S ANALYSIS

AUTHOR: Alisson Castro Barreto
ADVISOR: Adriano Mendonça Souza

Cancer has become a public health problem in Brazil, which is evident by the increase in its incidence and mortality rates by research institutes specializing in the subject such as the National Cancer Institute and International Agency for Research on Cancer. In Brazil, prostate cancer (PC) and breast cancer (BC) are currently the most common types of cancers in men and women, respectively. Therefore, in order to determine the spatial distribution of PC and BC mortality in Brazilian micro-regions and to verify their interrelationship with socio-demographic, economic, environmental and infrastructure factors, the techniques of Spatial Econometric were used: Exploratory Spatial Data Analysis – ESDA and Spatial dependence modelling. Variables were chosen to represent the factors, according to relevant literature and data availability, being this: oncology referral hospitals, number of oncology specialist doctors, aging rate, urbanization rate, average amount of pesticides used, Gross Domestic Product, and number of mammography devices. The database used comes from the Brazilian Health Informatics Department - DATASUS and the Brazilian Institute of Geography and Statistics - IBGE. Positive spatial autocorrelation was found for the mortality rate from prostate cancer ($I = 0.4537$; $p \leq 0.001$) and breast cancer ($I = 0.4842$; $p \leq 0.001$). The prostate cancer mortality rate has a positive correlation with the aging rate and urbanization rate. The breast cancer mortality rate was positively correlated with the number of specialist doctors, reference hospitals, urbanization rate, and aging rate, but it was negatively correlated with the number of mammography devices. It was an indicative that the higher life expectancy and Western lifestyle have an influence on the increase in mortality from prostate and breast cancer in Brazil. In view of the above, it is considered that the general and specific objectives of the research have been achieved and the spatial analysis of cancer mortality has made it possible to find statistical correlations that merit epidemiological research to subsidize actions in the fight against prostate and breast cancer mortality in women in Brazil.

Keywords: Spatial Econometrics. Mortality. Cancer. Spatial Autocorrelation.

LISTA DE ILUSTRAÇÕES

Figura 1 – Representação dos critérios de vizinhança rainha (Q), torre (R) e bispo (B).....	23
Figura 2 - Diagrama de dispersão de Moran Univariado.....	28
Figura 3 - Procedimento de especificação dos modelos espaciais.....	46
ARTIGO	
Figura 1 – Aninhamento dos modelos espaciais.....	54
Figura 2 - Mapa coroplético da taxa de mortalidade por câncer de próstata (A) e mama (B) em 2018 nas microrregiões do Brasil.....	57
Figura 3 - Mapa de clusters LISA da taxa de mortalidade por câncer de próstata (A) e mama (B) em 2018 nas microrregiões do Brasil.....	58

LISTA DE QUADROS

Quadro 1 – Síntese dos modelos espaciais para dados cross section e seus respectivos parâmetros.....	35
Quadro 2 – Variáveis utilizadas na pesquisa e suas fontes.....	43
ARTIGO	
Quadro 1 – Descrição das variáveis.....	50

LISTA DE TABELAS

ARTIGO

- Tabela 1 – Resultado para análise da autocorrelação espacial global das taxas de mortalidade por câncer de próstata e mama nas microrregiões brasileiras em 2018.....58
- Tabela 2 - Resultado da estimação dos parâmetros da regressão linear simples para a mortalidade por câncer de próstata e mama.....59
- Tabela 3 - Estimativa dos parâmetros e impactos espaciais para os modelos SDM e SAR.....69

LISTA DE ABREVIATURAS

AEDE	Análise Exploratória de Dados Espaciais
AIC	<i>Akaike Information Criterion</i>
BIC	<i>Bayesian Information Criterion</i>
CaM	Câncer de Mama
CaP	Câncer de Próstata
DATASUS	Departamento de Informática do Sistema Único de Saúde do Brasil
GIS	<i>Geographic Information Systems</i>
GSM	<i>General Spatial Model</i>
GSPRE	<i>Generalized Spatial Random-Effects model</i>
HPA	Hidrocarbonetos policíclicos aromáticos
IBGE	Instituto Brasileiro de Geografia e Estatística
INCA	Instituto Nacional do Câncer
IPEA	Instituto de Pesquisa Econômica Aplicada
LISA	<i>Local Indicators of Spatial Association</i>
MCRL	Modelo Clássico de Regressão Linear
MQO	Mínimos Quadrados Ordinários
OMS	Organização Mundial da Saúde
PIB	Produto Interno Bruto
PSA	<i>Prostate Specific Antigen</i>
SAC	<i>Spatial Autocorrelation Model</i>
SAR	<i>Spatial Auto Regressive Model</i>
SARMA	<i>Spatial Autoregressive Moving Average</i>
SDEM	<i>Spatial Durbin Error Model</i>
SDM	<i>Spatial Durbin Model</i>
SEM	<i>Spatial Error Model</i>
SLX	<i>Spatially Lagged X</i>
SLXMA	<i>Spatially Lagged X Moving Average</i>
SMA	<i>Spatial Moving Average</i>

SUMÁRIO

1	INTRODUÇÃO	12
1.1	PROBLEMA	14
1.2	OBJETIVOS	14
1.2.1	Objetivo Geral	14
1.2.2	Objetivos Específico	14
1.3	JUSTIFICATIVA	14
1.4	ESTRUTURA DO TRABALHO	16
2	REVISÃO DE LITERATURA	17
2.1	CÂNCER DE PRÓSTATA E MAMA	17
2.2	ECONOMETRIA ESPACIAL	21
2.2.1	Análise Exploratória de Dados Espaciais	25
2.2.2	Modelos Econométricos Espaciais	30
	2.2.2.1 Modelos Econométricos Espaciais para dados em corte transversal	30
	2.2.2.2 Modelos Econométricos Espaciais para dados em painel	37
3	MATERIAIS E MÉTODOS	42
3.1	BASE DE DADOS	42
3.2	ANÁLISE EXPLORATÓRIA DE DADOS ESPACIAIS	44
3.3	MODELAGEM ESPACIAL	45
4	ARTIGO: ESTUDO ECOLÓGICO ESPACIAL DA MORTALIDADE POR CÂNCER DE PRÓSTATA E MAMA NO BRASIL	47
5	CONSIDERAÇÕES FINAIS	72
	REFERÊNCIAS	74

1 INTRODUÇÃO

O destaque relativo ao câncer, suas taxas de incidência e de mortalidade por institutos de pesquisa especializados no assunto como, por exemplo, Instituto Nacional de Câncer e Agência Internacional de Pesquisa em Câncer, não se dá de forma aleatória. Segundo a Organização Mundial da Saúde (OMS) (2018), de algum modo, todas as famílias ao redor do mundo são afetadas pela doença e estima-se que até 2038 os números de novos casos de câncer devem aumentar em 70%.

A incidência de todos os tipos de câncer, quando combinados, é duas vezes maior em países desenvolvidos quando comparada com a de países em desenvolvimento. Mas essa proporção não se mantém em relação à taxa de mortalidade, que é apenas 8% a 15% maior. Essa discrepância reflete a diferença entre os perfis das doenças e o acesso ao tratamento (TORRE et al., 2015).

Países latino-americanos e o Caribe têm lutado para responder às crescentes taxas de incidência e mortalidade do câncer, mas, historicamente, enfrentam muitos problemas como: financiamento inadequado, má distribuição de recursos e serviços, falta de treinamento e distribuição de pessoal da saúde, falta de cuidados adequados com a população, com base em fatores socioeconômicos, étnicos e geográficos, entre outros. Esse cenário, se mantido, pode trazer danos generalizados e de ordem econômica (GOSS et al., 2013).

Em mulheres, o câncer é a principal causa de morte no mundo, tanto em países desenvolvidos quanto naqueles em desenvolvimento. Além disso, a incidência de câncer está aumentando em diversos países devido ao envelhecimento mundial da população (TORRE et al., 2017). O crescimento na incidência por câncer é esperado principalmente nos países em desenvolvimento, como o Brasil, onde a expectativa de vida está crescendo devido ao avanço nas políticas públicas relacionadas a saúde, como controle de doenças infecciosas e redução na mortalidade infantil (FRANCESCHI; WILD, 2013).

A incidência de câncer também cresce entre as mulheres nos países em desenvolvimento devido a alterações nos fatores de risco de câncer, à medida que os países experimentam transição econômica. Esses fatores de risco incluem tabagismo, excesso de peso corporal, inatividade física e alterações nos padrões reprodutivos, como maior idade no primeiro parto e menor número de partos. Por isso, os cânceres que antes eram comuns apenas nos países desenvolvidos estão se tornando mais prevalentes nos países em desenvolvimento (TORRE et al., 2017).

O câncer de mama (CaM) é o tipo de câncer mais frequente entre as mulheres tanto no Brasil como no mundo. Em 2018 houve 17.572 mortes por câncer de mama em mulheres no Brasil; e para 2020, estima-se 66.280 novos casos. Excetuando o câncer de pele não melanoma, o CaM é o mais frequente entre as mulheres nas regiões Sul, Sudeste, Centro-Oeste e Nordeste. Este tipo de câncer também acomete homens, porém é considerado raro e o total de casos é de apenas 1% (INSTITUTO NACIONAL DE CÂNCER - INCA, 2020).

Entre os homens, o segundo tipo de câncer mais frequente e o quinto tipo de câncer que mais mata no mundo é o câncer de próstata (CaP). Para 2018, foi estimado que apareceriam pelo menos 1.3 milhões de novos casos e 359 mil mortes por CaP entre todos os países (BRAY et al., 2018). No Brasil, entre os homens, excluindo o câncer de pele não-melanoma, o CaP é o tipo de câncer mais comum. Em 2018 houve 15.576 mortes pela doença e previa-se o registro de 65.840 novos casos em 2020 (INSTITUTO NACIONAL DE CÂNCER - INCA, 2020).

Dois fatores são responsáveis por tornar o câncer um problema de saúde pública no Brasil: o aumento da incidência e mortalidade da doença, proporcionalmente ao envelhecimento da população, ao desenvolvimento econômico e ao crescimento populacional, e o desafio para o sistema público de saúde brasileira no que diz respeito a acesso, diagnóstico e tratamento da população. O diagnóstico precoce e o tratamento adequado são uma das principais estratégias para o enfrentamento da doença (OLIVEIRA et al., 2011b).

Por isso, com o objetivo de contribuir para melhor compreensão da atual distribuição espacial da mortalidade por câncer de próstata e mama no Brasil, esta pesquisa faz uso de técnicas da econometria espacial. Entre as técnicas que serão utilizadas estão a Análise Exploratória de Dados Espaciais para identificar *outliers* espaciais e a formação de *clusters*, além da modelagem da dependência espacial.

A econometria espacial distingue-se da econometria tradicional pelos aspectos espaciais específicos dos dados, de modo que possibilita considerar a influência de uma região sobre outra. Tais aspectos podem ser denominados dependência espacial e heterogeneidade espacial. A dependência espacial costuma estar presente em dados *cross-section* e está ligada às áreas da ciência regional e geografia expressa na primeira lei da geografia desenvolvida por Tobler (1979): todas as coisas estão relacionadas entre si, mas coisas próximas estão mais relacionadas que coisas distantes. Já a heterogeneidade espacial reflete o fato de que os parâmetros variam com a localização e não são homogêneos em todo o conjunto de dados (ANSELIN, 1997).

1.1 PROBLEMA

O câncer se tornou um problema de ordem mundial e é a principal causa de morte em alguns países. O Brasil também segue essa tendência e caso não ocorra aprimoramento das políticas de prevenção e tratamento da doença, o câncer poderá se tornar a maior causa de morte brasileira, superando as doenças cardiovasculares.

A distribuição espacial das taxas de mortalidade por câncer de próstata e mama, nas microrregiões brasileiras, dá-se de forma aleatória no espaço ou é autocorrelacionada espacialmente? Fatores sociodemográficos, econômicos, ambientais e de infraestrutura influenciam na distribuição espacial?

1.2 OBJETIVOS

Esta seção refere-se aos objetivos que a pesquisa pretende alcançar e está subdividida em Objetivo Geral e Objetivos Específicos.

1.2.1 Objetivo Geral

Determinar a distribuição espacial da mortalidade por câncer de próstata e mama nas microrregiões brasileiras e verificar o seu inter-relacionamento com fatores sociodemográficos, econômicos, ambientais e de infraestrutura.

1.2.2 Objetivos Específico

Verificar se as microrregiões brasileiras apresentam autocorrelação espacial pela estatística I de Moran;

Determinar os fatores sociodemográficos, econômicos, ambientais e de infraestrutura de acordo com literatura relevante e disponibilidade de dados por microrregião;

Encontrar qual o melhor modelo espacial para representar a autocorrelação espacial das taxas de mortalidade por câncer de próstata e mama nas microrregiões brasileiras, baseado nas estatísticas de ajustes.

1.3 JUSTIFICATIVA

Realizou-se pesquisa na base de dados *Web of Science, Scopus, Lilacs, Medline e Pubmed*, relacionando econometria espacial e mortalidade por câncer de próstata e mama. Foram encontrados apenas 10 artigos, entre os anos de 1996 e 2020, e nenhum dos artigos encontrados analisou a mortalidade por câncer de próstata e mama nas microrregiões do Brasil, indicando uma lacuna de pesquisa.

A análise espaço-temporal sobre casos de incidência, mortalidade e sobrevida de neoplasias pode ser uma importante ferramenta de apoio à tomada de decisão para poderes públicos e órgãos ou institutos de saúde interessados. Segundo Domingues (2017), métodos estatísticos para estudar variáveis temporais distribuídas espacialmente iniciaram-se na geociência há cerca de 50 anos, mas já possuem aplicações em áreas como epidemiologia, climatologia, economia, entre outras.

A crescente atenção para especificar, estimar e testar a presença de interação espacial pode ser atribuída a dois fatores. O primeiro é a aferição de como a magnitude de uma variável de interesse em um determinado local é determinada pelos valores da mesma variável em outros locais do sistema. E o segundo é a necessidade de lidar com dados espaciais, em virtude da popularização de sistemas de informação geográfica (GIS) (ANSELIN, 2003a).

Por meio da análise da distribuição espacial das observações é possível identificar se existe variação nos dados, conglomerados que possam vir a representar áreas de riscos (*hot spot*), analisar a tendência das taxas de interesse como, por exemplo, maiores e menores percentuais de óbitos, verificar se a mortalidade em uma microrregião afeta ou é afetada pelas áreas que a circundam e realizar previsão espaço-temporal.

A análise dos padrões espaciais do câncer no Brasil ainda é pouco explorada e se configura como importante ferramenta para o planejamento e implementação de novos programas de controle de neoplasias. Esta pesquisa se torna inédita por verificar a distribuição espacial da mortalidade dos dois tipos de cânceres mais frequentes para homens e mulheres, no Brasil, investigar se há uma relação da mortalidade dessas doenças nas microrregiões brasileiras, levando em conta sua localização e investigar qual seria a fonte dessa autocorrelação espacial. A investigação da fonte da autocorrelação espacial inclui encontrar correlação da mortalidade da doença com outras variáveis e isso pode ajudar a identificar hipóteses que mereçam futura investigação epidemiológica para contribuir com políticas públicas relacionadas à saúde (HERRMANN et al., 2015; MADHU et al., 2016). Justifica-se, desse modo, o desenvolvimento da pesquisa.

1.4 ESTRUTURA DO TRABALHO

Visando atender aos objetivos propostos, esta pesquisa encontra-se estruturada em 5 capítulos, que serão especificados a seguir.

O primeiro capítulo refere-se à Introdução da pesquisa. Neste é apresentado a introdução e o delineamento do tema, bem como da técnica a ser utilizada, encontrando-se subdividida em: Problema, Objetivos, Justificativa e Estrutura do Trabalho.

O segundo capítulo, a Revisão de Literatura, está subdividida em Câncer de próstata e mama e Econometria Espacial. Esta última encontra-se dividida em: Análise Exploratória de Dados Espaciais e Modelos Econométricos Espaciais.

Já o terceiro capítulo, Materiais e Métodos, refere-se a abordagem metodológica da pesquisa e subdivide-se em: Análise Exploratória de Dados Espaciais, Modelagem Espacial e Base de Dados.

No quarto capítulo, os resultados são apresentados no formato de um artigo relacionado ao tema.

No quinto capítulo, encontram-se as considerações finais da pesquisa.

Após estas, são especificados as Referências.

2 REVISÃO DE LITERATURA

Para dar suporte teórico e apresentar estudos relevantes à pesquisa, a Revisão de Literatura encontra-se estruturada em 2 subseções: Câncer de próstata e mama e Econometria Espacial.

2.1 CÂNCER DE PRÓSTATA E MAMA

O câncer é uma doença que tem origem na multiplicação desordenada de células causada por mutações genéticas que codificam as proteínas reguladoras do ciclo celular. Essas alterações provocam diferentes características como a capacidade de multiplicação celular mesmo sem estímulos, a metástase que configura a migração da enfermidade para outras partes do corpo e também a não submissão à morte celular programada, denominada apoptose (BERNARDES et al., 2019).

No Brasil, os indicadores para o triênio 2020-2022 estimam a ocorrência de 625 mil novos casos de câncer para cada ano, sendo destes 177 mil casos de câncer não melanoma. O câncer não melanoma é o mais frequente tanto no Brasil quanto no mundo. No Brasil ele representa cerca de 30% de todos os tumores malignos registrados, mas possui baixa taxa de letalidade (INSTITUTO NACIONAL DE CÂNCER - INCA, 2020).

Com exceção do câncer não melanoma em homens, os cânceres presumidamente mais frequentes para o triênio 2020-2022 serão próstata (29,2%), cólon e reto (9,1%), pulmão (7,9%), estômago (5,9%) e cavidade oral (5,0%). Para mulheres, câncer de mama (29,7%), cólon e reto (9,2%), colo do útero (7,5%), pulmão (5,6%) e tireoide (5,4%) seriam os de maior incidência (INSTITUTO NACIONAL DE CÂNCER - INCA, 2020)

O Brasil apresenta estimativas com perfil semelhante ao de países desenvolvidos, porém ainda possui altas taxas de cânceres associados a infecções, que refletem características de países em desenvolvimento. Isso é reflexo das desigualdades regionais brasileiras, que englobam diferenças no acesso à saúde, expectativas de vida e condições socioeconômicas (SANTOS, 2018).

O câncer de próstata é a neoplasia mais frequente diagnosticada em homens em mais da metade dos países e é a principal causa de morte por câncer em 46 países. As incidências por câncer de próstata nas últimas décadas têm aumentado, mas tais registros podem ter influência na maior utilização de testes de detecção da doença, como o teste PSA - *Prostate Specific Antigen*, realizado através de coleta de sangue ou através de procedimentos mais invasivos com

a remoção de tecidos para análise. Países como Reino Unido, Japão, Brasil, Costa Rica e Tailândia, que demoraram a utilizar amplamente o teste PSA, apresentam aumento nas taxas por CaP (BRAY et al., 2018).

O câncer de próstata apresenta como fatores de risco a idade, que possui impacto tanto na incidência quanto na mortalidade; pai ou irmão com a doença antes dos 60 anos de idade; excesso de gordura corporal e exposição a aminas aromáticas, ao arsênio, derivados de petróleo, hidrocarbonetos policíclicos aromáticos (HPA), fuligem e dioxinas (INSTITUTO NACIONAL DE CÂNCER - INCA, 2019a).

Nas mulheres, o câncer de mama se tornou um desafio para o enfrentamento de doenças crônicas não transmissíveis no Brasil e para a tendência do envelhecimento da população brasileira. Excetuando o tipo de câncer não melanoma, é o câncer de maior incidência e mortalidade no Brasil. Em 2018 ocorreram 17.572 óbitos por câncer de mama em mulheres. As maiores taxas foram registradas nas regiões Sul e Sudeste do Brasil (INSTITUTO NACIONAL DE CÂNCER - INCA, 2019b).

A idade é o principal fator de risco para o desencadeamento do câncer de mama, outros fatores que impactam na morbidade da doença são fatores ambientais e comportamentais, como sobrepeso, sedentarismo, consumo de bebidas alcóolicas; fatores da história reprodutiva e hormonal, como menopausa, não ter tido filhos, uso de contraceptivos hormonais; e fatores genéticos e hereditários, como histórico familiar e alterações genéticas (INSTITUTO NACIONAL DE CÂNCER - INCA, 2019c).

Algumas pesquisas já utilizaram as técnicas de econometria espacial para investigar a mortalidade e/ou incidência dos cânceres de próstata e mama, verificar a existência de dependência espacial e seu inter-relacionamento com outras variáveis. As 10 pesquisas serão citadas abaixo.

Guerra et al. (2020) utilizaram a análise exploratória de dados espaciais para verificar a distribuição espacial da incidência e mortalidade por câncer de mama nos estados do Brasil e ajustaram um modelo de painel espacial para analisar o inter-relacionamento com o índice sócio demográfico. Dentre os resultados encontrados, observou-se que regiões mais desenvolvidas possuem maiores taxas de incidência e mortalidade por câncer de mama.

A pesquisa de Torres-Roman et al. (2019) buscou avaliar as taxas de mortalidade por câncer de próstata de acordo com a região geográfica no Peru, entre os anos de 2005 a 2014. A autocorrelação espacial foi determinada pela estatística I de Moran. Os autores chegaram à conclusão de que a taxa de mortalidade por câncer de próstata continua a crescer no Peru.

Em outra pesquisa, Amin, Fritsch e Retzlöff (2019) estudaram agrupamentos espaciais formados pela mortalidade e incidência de câncer de mama feminino nos Estados Unidos. Foram utilizados *cluster* espacial e análise de regressão linear com variáveis como idade, fumaça, poluição do ar, inatividade física, vida urbana, nível de educação e raça. Os achados foram que existem vários grupos significativos nos EUA, tanto para a incidência de câncer de mama quanto para a mortalidade.

Góes et al. (2018) desenvolveram estudo ecológico da mortalidade por câncer de próstata no estado de Sergipe, no Brasil, utilizando a análise exploratória de dados espaciais. Foi encontrada uma distribuição heterogênea da mortalidade por câncer de próstata, identificando-se áreas de risco e foi observado aumento das taxas de mortalidade por câncer de próstata no estado.

Núñez et al. (2016) avaliaram a associação entre os níveis de cromo e arsênico no solo e a mortalidade por câncer na Espanha; para tanto, utilizaram modelo espacial bayesiano. Os resultados sugeriram que a exposição decorrente de baixos níveis de arsênico e cromo no solo superficial pode ser um fator de risco potencial para o desenvolvimento de câncer. Em outra pesquisa realizada nos Estados Unidos, Mueller et al. (2015) utilizaram uma abordagem de Sistemas de Informações Geográficas (GIS) para realizar análises espaciais e regressão linear para testar duas variáveis de exposição à mineração de carvão e suas associações com a incidência e mortalidade por câncer em Illinois. A autocorrelação espacial global indicou um agrupamento espacial significativo de incidência, mortalidade e produção agregada de carvão. As exposições à mineração de carvão foram associadas ao risco elevado de vários cânceres, incluindo próstata e mama.

Os autores Vinnakota e Lam (2006) mostraram o uso de uma abordagem de extração de regras para investigar associações entre variáveis socioeconômicas selecionadas e as quatro principais causas de mortalidade por câncer nos Estados Unidos: câncer colorretal, pulmão, mama e próstata. As associações foram mapeadas para visualizar a distribuição espacial e verificou-se que as áreas de serviços de saúde com baixa escolaridade, alto desemprego e baixos salários estão associadas a altas taxas de mortalidade por câncer.

Em pesquisa realizada no Brasil, Barbosa et al. (2016) objetivaram analisar a distribuição espacial da mortalidade por câncer e sua correlação com os fatores socioeconômicos nos municípios brasileiros no período de 2010 a 2012. Para a autocorrelação e formação de *clusters* foi utilizada a estatística I de Moran Global e local (LISA - *Local Indicators of Spatial Association*). Avaliou-se a correlação espacial entre as taxas de mortalidade e outras variáveis por meio da análise bivariada LISA. Os autores concluíram que

existe desigualdade na distribuição das taxas de mortalidade e que elas estão mais correlacionadas às áreas mais desenvolvidas e urbanizadas.

Em outra pesquisa realizada no Brasil, Silva et al. (2015) buscaram explorar correlações entre variáveis relacionadas à produção agrícola, uso de serviços de saúde, consumo de alimentos, características sociodemográficas e taxas de mortalidade por câncer de próstata nos estados brasileiros. A produção de soja e milho possuem correlação positiva com a mortalidade por câncer de próstata. Por meio da regressão linear espacial múltipla, algumas variáveis mostraram associação com as taxas de mortalidade por câncer de próstata: toneladas de soja produzida, proporção da população com 80 anos ou mais e consumo de bebidas. Foi identificada ainda correlação positiva entre toneladas de soja plantada e mortalidade por câncer de próstata, indicando associação entre exposição a pesticidas e câncer de próstata. O modelo escolhido para representar a autocorrelação espacial não foi informado.

Rocha-Brischiliari et al. (2018) analisaram espacialmente a taxa de mortalidade por câncer de mama e seu relacionamento com disparidades socioeconômicas e acesso a tratamentos especializados em 399 cidades do Estado do Paraná, no Brasil. Foram utilizadas para a autocorrelação espacial a estatística I de Moran global e local, além do LISA *map* para os *clusters*. Um modelo SAR (*Spatial Auto Regressive Model*) foi conduzido para estimar os coeficientes de regressão das variáveis independentes (analfabetismo, renda e índice de acessibilidade ao diagnóstico de CaP) em relação à variável dependente. Os resultados mostraram que uma menor taxa de analfabetismo não está relacionada com uma menor taxa de mortalidade específica por câncer de mama, e quanto maior a pontuação de acessibilidade aos serviços de câncer, maior a mortalidade específica por câncer de mama.

Observa-se que os estudos indicaram associação entre mortalidade de câncer de próstata e/ou de mama com variáveis como cromo e arsênico no solo, exposição a mineração de carvão, baixa escolaridade, alto desemprego, baixos salários, taxa de urbanização, exposição a pesticidas, acessibilidade a serviços de saúde e índice sócio demográfico. Estas variáveis podem, então, ser testadas quanto à existência de associação com as taxas de mortalidade por câncer de próstata e mama nas microrregiões brasileiras, dependendo, entretanto, da disponibilidade dos dados por microrregião nas principais bases oficiais do Brasil.

A próxima subseção, Econometria Espacial, abordará as especificidades relacionadas à ferramenta estatística.

2.2 ECONOMETRIA ESPACIAL

Econometria espacial pode ser definida como o subconjunto da econometria que lida com aspectos espaciais presentes nas observações transversais e espaço-temporais. Variáveis relacionadas a localização, distância e arranjo e topologia são tratadas explicitamente na especificação do modelo, estimativa, verificação e previsão de diagnóstico. Trata-se de um campo cujas técnicas analíticas são projetadas para incorporar dependência entre observações que estão em estreita proximidade geográfica (ANSELIN; SYABRI; KHO, 2006; LESAGE; P., 2008).

A incorporação de efeitos espaciais aos modelos de regressão diferencia a econometria espacial da econometria convencional. Pressupostos do Modelo clássico de regressão linear e algumas das hipóteses de Gauss-Markov, como média condicional zero, homocedasticidade e independência dos erros, não se sustentam na presença de dependência espacial e da heterogeneidade espacial (ALMEIDA, 2012).

Em termos gerais, dependência espacial, ou autocorrelação espacial, pode ser entendida como a existência de uma relação funcional entre o que acontece em um ponto do espaço e o que acontece em outro lugar. Duas condições podem levar a isso: erros de medição para observações espaciais e a existência de uma variedade de fenômenos de interação espacial (ANSELIN, 1988).

A heterogeneidade espacial pode ser entendida como uma variabilidade do comportamento econômico no espaço. Pode ocorrer, portanto, formação de padrões espaciais ou agrupamentos ao longo do conjunto de dados. Desse modo, os parâmetros variam e ocasionalmente mudam a forma estrutural do modelo, podendo inclusive gerar heterocedasticidade com possíveis erros de especificação (VIEIRA, 2009).

Grande parte dos problemas gerados pela heterogeneidade espacial pode ser resolvida com o uso da econometria tradicional, porém o conhecimento teórico sobre a estrutura espacial dos dados costuma apresentar resultados mais eficientes, além de que, há casos em que não é fácil diferenciar heterogeneidade espacial de autocorrelação espacial, ou mesmo quando as duas ocorrem ao mesmo tempo (ANSELIN, 1988).

Os efeitos espaciais são a principal razão da existência de um campo específico da econometria denominado econometria espacial. A saber: dependência espacial e heterogeneidade espacial, dois aspectos que merecem particular atenção do ponto de vista metodológico (ANSELIN, 1988).

Quatro dimensões importantes podem ser identificadas que definem o escopo da moderna metodologia econométrica espacial: especificação do modelo; estimativa; teste de especificação; e previsão espacial (ANSELIN, 2010).

Para Câmara et al. (2004), a análise espacial é um conjunto de procedimentos que objetiva escolher o melhor modelo inferencial que incorpore os efeitos espaciais presentes no fenômeno. Comumente, antes de realizar a modelagem, é feita a análise exploratória de dados espaciais (AEDE), associada à representação dos dados de forma visual por meio de mapas e gráficos, além da identificação de padrões de dependência no estudo.

A noção de dependência espacial implica a necessidade de determinar outras unidades no sistema espacial, que possuem influência sobre a unidade em consideração. Formalmente, isso é expresso nas noções topológicas de vizinhança e vizinho mais próximo (ANSELIN, 1998).

Um aspecto crucial na definição da autocorrelação espacial é determinar o que deve ser considerado locacionalmente próximo, ou seja, observações que circundam outras e que podem exercer determinada influência sobre aquelas. Infelizmente, a determinação da vizinhança não acontece sem um determinado grau de arbitrariedade (FISCHER; WANG, 2011). Busca-se encontrar um valor que represente o grau de conexão entre as regiões de interesse.

Para tal, tem-se a matriz de pesos que possui a função, na econometria espacial, de representar a estrutura de dependência entre as observações analisadas (PIRAS; LOZANO-GRACIA, 2012). Seus valores são definidos conforme as hipóteses do grau de dependência (CORRADO; FINGLETON, 2012). Por isso, um ponto crucial na determinação da matriz de pesos é que ela deve incorporar fielmente as interações entre as observações estudadas (GOLGHER, 2015).

Os vários tipos de matrizes de ponderação espacial podem ser subdivididos em duas categorias: matrizes de proximidade geográfica e matrizes de proximidade socioeconômica. Nesta pesquisa, serão abordados os conceitos das matrizes de contiguidade e distância geográfica, como matrizes de proximidade geográfica. Para as matrizes de proximidade socioeconômica, serão descritas as matrizes baseadas nos conceitos de similaridade, dissimilaridade e fluxos.

A ideia de vizinhança baseada em contiguidade pode ser verificada quando duas regiões geográficas compartilham de uma fronteira física comum. Matematicamente, a participação de observações na vizinhança pode ser expressa por meio de uma contiguidade espacial $n \times n$ ou matriz de pesos W_{ij} (ALMEIDA, 2012). A matriz de pesos W_{ij} pode ser expressa pela equação (1):

$$W_{ij} = \begin{bmatrix} W_{11} & \cdots & W_{1n} \\ \vdots & \ddots & \vdots \\ W_{n1} & \cdots & W_{nn} \end{bmatrix} \quad (1)$$

Onde:

n representa o número de observações;

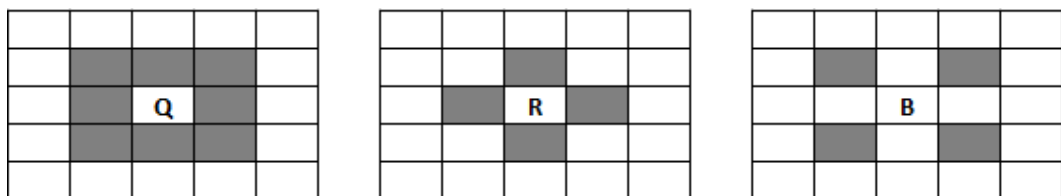
Os elementos da diagonal principal da matriz possuem, por convenção, o valor de zero;

Os demais elementos assumem valores de 1, quando as localizações i e j são consideradas vizinhas; e 0 quando não o são.

A contiguidade binária entre unidades espaciais significa que, se duas unidades espaciais possuírem uma borda comum de comprimento diferente de zero, elas serão consideradas contíguas e um valor de 1 será atribuído. Essa definição de contiguidade, obviamente, pressupõe a existência de um mapa, a partir do qual os limites podem ser discernidos. Para um arranjo irregular de unidades de área espacial, isso é razoavelmente direto. No entanto, quando as unidades espaciais se referem a uma grade regular ou a uma coleção de pontos dispostos irregularmente, a determinação da contiguidade não é única (ANSELIM, 1998).

Para tal, deve-se definir como será considerado o critério de vizinhança na matriz de contiguidade. A contiguidade é definida como rainha (*Queen - Q*), se cada fronteira e vértice configurarem como vizinhos. Caso somente sejam consideradas as fronteiras para o critério de vizinhança, a contiguidade é definida como torre (*Rook- R*); e se somente forem considerados os vértices, bispo (*Bishop - B*) (ALMEIDA, 2012). As representações dos critérios de vizinhança podem ser visualizadas na Figura 1.

Figura 1 – Representação dos critérios de vizinhança rainha (Q), torre (R) e bispo (B)



Fonte: Elaborado pelo autor

Em relação à matriz de distância geográfica, o princípio é que regiões próximas geograficamente possuem maior interação espacial. Uma matriz desse tipo bastante abordada

em literaturas é a matriz dos k vizinhos mais próximos, $w_{ij}(k)$. Trata-se de uma matriz binária cujo critério de proximidade é baseado na distância geográfica (ALMEIDA, 2012), podendo ser expressa pela equação (2):

$$w_{ij}(k) = \begin{cases} 1 & \text{se } d_{ij} \leq d_i(k) \\ 0 & \text{se } d_{ij} > d_i(k) \end{cases} \quad (2)$$

Onde:

$d_i(k)$ representa a distância limite para que a região i tenha k vizinhos.

Nem sempre a interação espacial é representada por ligações geográficas, ela pode ser por forças socioeconômicas. As matrizes socioeconômicas por similaridades baseiam-se na ideia de que quem é mais similar influencia mais do que os diferentes e as distâncias são construídas com base na diferença de similaridade. No que tange às matrizes socioeconômicas baseadas em dissimilaridade, o princípio norteador é o de que regiões dissimilares são mais conectadas entre si. Outro tipo de matriz socioeconômica são as matrizes por fluxos; assim, quanto maior o fluxo estabelecido, como, por exemplo, as relações comerciais, maior a interação espacial (ALMEIDA, 2012).

Como a determinação da matriz é feita de forma arbitrária, caso ela não venha a retratar o processo real de geração de dados, o modelo espacial poderá gerar conclusões empíricas equivocadas (HARRIS; MOFFAT; KRAVTSOVA, 2011). A arbitrariedade na determinação da matriz de pesos pode ser entendida com uma limitação na utilização de modelos espaciais (ARBIA; FINGLETON, 2008).

Para determinar a matriz de pesos, pode-se recorrer a outros estudos semelhantes que retratam a interação entre as observações analisadas; porém, quando isso não é possível, uma opção é comparar os modelos espaciais com as matrizes de pesos, utilizando para tal o critério da máxima verossimilhança. Ou, se mesmo assim não houver indicativo de qual utilizar, deve-se adotar uma matriz simples, como a de contiguidade (STAKHOVYCH; BIJMOLT, 2009).

Com o objetivo de contornar a arbitrariedade na escolha da matriz de pesos, dada a sua importância na análise estatística espacial, esta pesquisa adotará a metodologia proposta por Baumont (2004), que consiste nas seguintes etapas:

- a) estima-se o Modelo Clássico de Regressão Linear (MCRL) por Mínimos Quadrados Ordinários (MQO);
- b) deve-se testar a autocorrelação espacial, por meio da estatística I de Moran, dos resíduos do melhor modelo ajustado com os diferentes tipos de matrizes de ponderação;

- c) Seleciona-se a matriz que gerou o maior valor para o teste I de Moran, mas que seja também estatisticamente significativo.

Após a definição da matriz de pesos, é necessário conhecer os dados com que se está trabalhando; isso é possível por meio da Análise Exploratória de Dados Espaciais.

2.2.1 Análise Exploratória de Dados Espaciais

A Análise Exploratória de Dados Espaciais (AEDE) pode ser definida como a coleção de técnicas para descrever e visualizar distribuições espaciais, identificar locais atípicos (*outliers* espaciais), descobrir padrões de associação espacial (*clusters*) e sugerir diferentes regimes espaciais e outras formas de instabilidade espacial ou não estacionariedade espacial (ANSELIN, 1997).

O objetivo da AEDE é possibilitar ao pesquisador conhecer melhor os dados com que está trabalhando e permitir captar as interrelações dos dados espaciais. Destaca-se que a AEDE é mais adequada para investigar variáveis espacialmente densas ou intensivas; isto é, variáveis divididas por indicadores de intensidade como variáveis *per capita* ou por área (ALMEIDA, 2012).

A primeira etapa da AEDE é verificar se os dados em estudo seguem um padrão espacial ou estão distribuídos aleatoriamente, ou seja, pretende-se verificar se existe autocorrelação espacial. O princípio básico da análise espacial de dados é a proposição de que os valores das variáveis que estão próximos são mais similares ou estão mais relacionados que valores em localizações distantes. Tobler (1979) corrobora essa ideia quando da definição da Primeira Lei da Geografia: tudo está relacionado com todas as coisas, porém coisas que estão perto são mais relacionadas que coisas distantes.

Se as observações que estão mais próximas também possuem similaridade no valor de suas variáveis, então diz-se que o padrão apresentado é de autocorrelação espacial positiva. Ao oposto denomina-se autocorrelação espacial negativa, quando observações espaciais próximas não possuem os valores de suas observações semelhantes, contrariando a primeira lei da geografia. A autocorrelação espacial nula ocorre quando os valores das variáveis são independentes da localização (FISCHER; WANG, 2011).

A autocorrelação pode ser mensurada por meio de um coeficiente global de autocorrelação espacial. Nesta pesquisa, será abordado a Estatística I de Moran. O coeficiente I de Moran é a estatística mais difundida para mensurar a autocorrelação espacial e se mostra mais sensível que outros coeficientes. É recomendável ainda que a análise de autocorrelação

não seja realizada com menos de 30 localidades (CARRIJO; DA SILVA, 2017; EPPERSON; LI, 1996; FORTIN; DRAPEAU; LEGENDRE, 1989). O coeficiente de autocorrelação espacial I de Moran é obtido por meio de uma medida de autocovariância mediante um produto cruzado (MORAN, 1948), conforme equação (3):

$$I = \frac{n}{\sum_i^n \sum_j^n w_{ij}} \times \frac{\sum_i^n \sum_j^n w_{ij} (y_i - \bar{y})(y_j - \bar{y})}{\sum_i^n (y_i - \bar{y})^2} \quad (3)$$

Onde:

n representa o número de regiões;

i e j são as regiões em análise;

$(y_i - \bar{y})(y_j - \bar{y})$ representa os valores da variável de interesse padronizada;

w_{ij} representa a matriz de ponderação espacial.

O resultado para a estatística I de Moran possui a seguinte interpretação:

$I = 1$ significa extrema autocorrelação espacial positiva (Lei de Tobler);

$I = -1$ significa extrema autocorrelação espacial negativa (oposição a lei de Tobler);

$I = 0$ significa ausência de autocorrelação: quando atributos são independentes da localização.

Para testar a significância estatística dos resultados encontrados para o coeficiente de autocorrelação espacial, existem duas formas. A primeira é assumir que a variável padronizada segue a distribuição normal (ALMEIDA, 2012). A fórmula pode ser expressa pela equação (4):

$$Z(I) = \frac{[I - E(I)]}{DP(I)} \quad (4)$$

Onde:

$E(I)$ representa o valor esperado de I que pode ser calculado por $E(I) = -[1/(n-1)]$;

$DP(I)$ significa o desvio padrão teórico de I .

A segunda forma é assumir aleatoriedade no mecanismo gerador dos dados espaciais. O teste consiste em permutar os dados espaciais para as diversas regiões e calcular o coeficiente para cada permutação, originando uma distribuição de referência empírica. É então comparada à estatística calculada do teste com a distribuição de referência empírica e verificado se está de

acordo com a região crítica de rejeição especificada pelo pesquisador. A hipótese nula do teste é de independência espacial, enquanto que a hipótese alternativa é de autocorrelação espacial (ALMEIDA, 2012; ANSELIN; SYABRI; KHO, 2006).

Outra forma de visualização da autocorrelação espacial é por meio do diagrama de dispersão de Moran. O diagrama é o gráfico da nuvem de pontos entre a variável de interesse y e sua defasagem espacial (Wy), ambas padronizadas (z e Wz), com indicação da reta de regressão, estimada por Mínimos Quadrados Ordinários (MQO) (ALMEIDA, 2012). A equação da reta de regressão pode ser representada conforme equação (5):

$$Wz = \alpha + \beta z + \varepsilon \quad (5)$$

Onde:

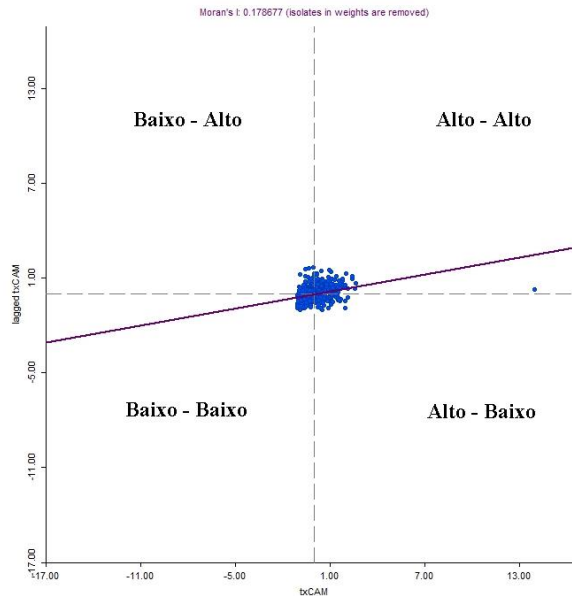
α é a constante da regressão;

βz é o coeficiente angular e equivale ao coeficiente I de Moran e;

ε é o erro aleatório.

O diagrama também fornece o tipo de associação linear, por meio dos seus quadrantes, a saber: Alto-Alto, Baixo-Baixo, Alto-Baixo e Baixo-Alto. Essa associação significa a associação entre os valores da variável de interesse e os valores das regiões que circundam, por exemplo: o primeiro quadrante representa a associação linear Alto-Alto, que significa que as regiões possuem valores acima da média e estão cercadas de regiões que também possuem valores altos (ANSELIN, 1996). O diagrama pode ser visualizado conforme Figura 2.

Figura 2 - Diagrama de dispersão de Moran Univariado



Fonte: Elaborado pelo autor

Pode-se, ainda, verificar a autocorrelação espacial global no contexto bivariado. O objetivo é verificar se valores de uma variável (Z_k) observada em determinado local mostram associação sistemática (mais do que provável sob aleatoriedade espacial) com outra variável (Z_l) observada nos locais vizinhos (ANSELIN; SYABRI; SMIRNOV, 2002). A estatística I de Moran no contexto bivariado pode ser expressa pela equação 6.

$$I_{kl} = \frac{Z'_k W Z_l}{Z'_k Z_l} \quad (6)$$

Em que:

n representa o número de observações;

W a matriz de pesos espaciais padronizadas;

$Z'_k Z_l$ representam as variáveis padronizadas.

Pode-se também gerar o diagrama de Dispersão de Moran bivariado, que expressa o gráfico da dispersão da nuvem de pontos entre os valores de uma variável padronizada em uma região com o valor de outra variável padronizada em regiões que a circundam e uma reta de regressão ajustada. Segue-se a mesma interpretação realizada para o diagrama de dispersão de Moran univariado. O coeficiente angular da reta pode ser interpretado como o coeficiente I de Moran e a associação linear dos quadrantes demonstra a associação entre os valores da variável

de interesse e os valores de outra variável em regiões vizinhas, podendo ser: Alto-Alto, Baixo-Baixo, Alto-Baixo e Baixo-Alto (ALMEIDA, 2012; ANSELIN; SYABRI; KHO, 2006).

O coeficiente de autocorrelação espacial global fornece apenas um valor para toda a área analisada. Com o objetivo de obter informações locais para as regiões identificando diferentes regimes de associação espacial, *clusters* e *outliers* espaciais locais, existe a estatística de autocorrelação espacial local (ALMEIDA, 2012; ANSELIN, 1995).

Para Anselin (1995), um indicador de autocorrelação espacial local (LISA) é qualquer estatística que satisfaça os seguintes critérios: ser capaz de indicar uma associação espacial significativa para cada observação e o somatório de todos os LISA's serem proporcionais ao indicador espacial global de autocorrelação relativo. Nesta pesquisa, será adotado o coeficiente *I* de Moran local.

O coeficiente de autocorrelação local I_i de Moran para a variável padronizada z_i observada na região i pode ser expressa matematicamente pela equação (7).

$$I_i = z_i \sum_{j=1}^j w_{ij} z_j \quad (7)$$

Para cada observação é calculado um coeficiente de autocorrelação espacial local I_i ; logo, se os resultados forem gerados em forma de tabelas, eles podem não facilitar a leitura e interpretação dos resultados. Nesse caso, costuma-se grafá-los em forma de mapa de *clusters* ou mapa de significância, denominado *LISA map*. O mapa sintetiza as informações do diagrama de dispersão de Moran com a significância da associação entre as observações e resulta em *clusters* destacados conforme o nível significância da associação. O *LISA map* e o diagrama de dispersão de Moran bivariado também são utilizados para identificar *outliers* espaciais. (ALMEIDA, 2012).

A autocorrelação espacial local também pode ser expressa em um contexto bivariado. Adaptando a equação do coeficiente *I* de Moran local, proposta por Anselin et al. (2002), obtém-se (8):

$$I_i^{z_1 z_2} = z_{1i} \sum_{j=1}^j W_{z_{2i}} \quad (8)$$

Onde:

z_{1i} e z_{2j} representam as variáveis de interesse padronizadas;
 $W_{z_{2i}}$ representa a defasagem espacial da variável z_{2i} .

2.2.2 Modelos Econométricos Espaciais

Após confirmada a existência de autocorrelação espacial, deve-se incorporar ao modelo de regressão linear convencional (equação 9) termos capazes de capturar esses efeitos, como: Wy , $W\xi$ ou $W\varepsilon$ e WX , que representam a defasagem espacial da variável dependente, do termo do erro e das variáveis independentes, respectivamente.

$$y = X\beta + \varepsilon, \quad \varepsilon \sim N(0, \sigma^2 I_n) \quad (9)$$

Onde:

X matriz de variáveis explicativas exógenas;
 β representam os coeficientes de regressão;
 ε representam os erros estocásticos.

Pode-se, ainda, adotar modelos econométricos espaciais para dados em corte transversal (*cross-section*) ou para dados em painel, conforme as seções seguintes.

2.2.2.1 Modelos Econométricos Espaciais para dados em corte transversal

Nesta seção serão abordados os principais modelos econométricos espaciais capazes de capturar os efeitos da dependência espacial global e/ou local e adequados para dados em corte transversal (*cross section*).

Os modelos de autocorrelação espacial de alcance global relacionam todos os locais do sistema entre si. Para um processo de erro, um choque no erro em qualquer local será transmitido para todos os outros locais de acordo com o multiplicador espacial expresso na sua equação (ANSELIN, 2003b). Nesta pesquisa, serão abordados, então, como modelos de alcance global, os modelos Modelo de *Lag* Espacial (SAR), Modelo de Erro Autorregressivo Espacial (SEM) e Modelo de Defasagem Espacial com Erro Autorregressivo Espacial (SAC).

O Modelo de *Lag* Espacial (SAR) foi proposto por Whittle (1954) e pode ser representado algebricamente pela equação (10):

$$y = \rho W y + X \beta + \varepsilon, \quad \varepsilon \sim N(0, \sigma^2 I_n) \quad (10)$$

Onde:

- ρ representa o coeficiente autorregressivo espacial e $|\rho| < 1$;
- $W y$ representa a defasagem espacial da variável dependente;
- X matriz de variáveis explicativas exógenas;
- β representa o coeficiente de regressão;
- ε é o erro estocástico.

Esta especificação implica que cada valor de Y é função de um grupo de variáveis explicativas X e uma média ponderada dos valores da variável dependente ($W y$) para as observações vizinhas. O parâmetro ρ indica o grau de correlação espacial entre as observações e, caso seja positivo, há autocorrelação espacial positiva; se ρ for negativo, a autocorrelação espacial é negativa. Se o parâmetro ρ não for significativo, pode-se considerar que o coeficiente é nulo, não existindo indicação de autocorrelação espacial (KLIER; MCMILLEN, 2008).

O Modelo de Erro Autorregressivo Espacial ou SEM foi proposto por Ord (1975) e possui a seguinte representação matemática indicada pelas equações (11) e (12):

$$y = X \beta + \xi \quad (11)$$

$$\xi = \lambda W \xi + \varepsilon \quad (12)$$

Onde:

λ é o parâmetro do erro autorregressivo espacial.

Nesse modelo, a dependência espacial se dá apenas pelo erro e não pela variável dependente. No modelo SEM observa-se que o erro estruturado espacialmente (ξ) é uma média dos erros nas regiões vizinhas ($\lambda W \xi$) mais um erro aleatório não explicado pelo modelo (ε). E λ representa o parâmetro do erro autorregressivo que é ponderado por uma matriz de pesos, indicando a defasagem do erro ($W \xi$). Quando não há correlação espacial entre os erros, $\lambda = 0$.

A influência sobre y não se dá apenas pelos choques do termo de erro da região; ela também é influenciada por choques provenientes de outras regiões. O efeito global de um choque ocorre porque é transmitido também para locais que são vizinhos de vizinhos por meio das potências de W . Um choque em uma região afeta as regiões vizinhas, e os vizinhos dos

vizinhos, e, eventualmente, retorna. Em outras palavras, o efeito total de um choque não é simplesmente o choque em si, mas o choque inicial mais a retroalimentação dos outros locais (FINGLETON, 2008).

O Modelo de Defasagem Espacial com Erro Autorregressivo Espacial ou SAC representa o modelo tradicional de regressão por MQO, acrescido de uma defasagem espacial e um termo de erro autocorrelacionado espacialmente. Nesse modelo, a dependência espacial ocorre tanto na variável dependente quanto nos erros (KELEJIAN; PRUCHA, 1998). O modelo pode ser representado pelas equações (13) e (14):

$$y = \rho W_1 y + X\beta + \xi, \quad |\rho| < 1 \quad (13)$$

$$\xi = \lambda W_2 \xi + \varepsilon, \quad |\lambda| < 1 \quad (14)$$

Onde:

W_1 e W_2 podem ser matrizes de ponderação espacial diferentes.

Nos modelos globais, o foco se encontra na interação entre todas as observações ou locais observados simultaneamente; em contraste, os modelos locais são apropriados quando os efeitos não modelados estiverem restritos às regiões vizinhas de primeira ou segunda ordem (ANSELIN, 2003b). Serão, então, abordados os seguintes modelos de alcance local: Modelo de Erro de Média Móvel Espacial (SMA), Modelo Regressivo Cruzado Espacial (SLX) e Modelo Regressivo Cruzado Espacial com Erro de Média Móvel (SLXMA).

O Modelo de Erro de Média Móvel Espacial ou SMA foi inicialmente proposto por Haining (1978) e pode ser expresso pelas equações (15) e (16):

$$y = X\beta + \xi \quad (15)$$

$$\xi = \gamma W\varepsilon + \varepsilon, \quad |\gamma| < 1 \quad (16)$$

Onde:

γ é o coeficiente de média móvel espacial.

Por se tratar de um modelo de alcance local, os efeitos espaciais ficam limitados para os vizinhos de primeira e segunda ordem. Observa-se que o erro espacial segue uma dinâmica de média móvel de primeira ordem, onde o erro é composto pelos efeitos na região de origem e pela média dos efeitos nas regiões adjacentes.

O Modelo Regressivo Cruzado Espacial ou SLX é o modelo de regressão linear espacial mais simples capaz de captar o transbordamento espacial local (*spillover effect*) por meio do parâmetro WX . Algebricamente o modelo pode ser expresso pela equação (17):

$$y = X\beta + WX\tau + \varepsilon \quad (17)$$

Onde:

τ é um vetor de coeficientes espaciais.

No que diz respeito à econometria espacial, existem pelo menos dois assuntos de interesse: a defasagem espacial e os efeitos de transbordamento espacial. A defasagem espacial mede o impacto da variável dependente y , das variáveis explicativas X ou do termo de erro ε observado em outras unidades de corte transversal. Um efeito de transbordamento espacial é definido como o impacto marginal que a alteração em uma variável explicativa gera sobre os valores das variáveis dependentes em outra região (ELHORSTA; VEGAA, 2017).

O Modelo Regressivo Cruzado Espacial com Erro de Média Móvel ou SLXMA trata-se de um modelo capaz de captar os efeitos do transbordamento local através de parâmetros nas variáveis explicativas e no termo do erro (ALMEIDA, 2012). Algebricamente pode ser expresso pelas equações (18) e (19):

$$y = X\beta + WX\tau + \xi \quad (18)$$

$$\xi = \gamma W\varepsilon + \varepsilon \quad (19)$$

A seguir, serão apresentados modelos que possuem alcance tanto global quanto local; a saber: Modelo de Durbin Espacial (SDM), Modelo de Defasagem Espacial com Erro de Média Móvel Espacial (SARMA), Modelo de Durbin Espacial do Erro (SDEM) e Modelo de Dependência Espacial Geral (GSM).

O Modelo de Durbin Espacial ou SDM possui alcance global por meio do parâmetro da variável dependente defasada e alcance local que se dá pela defasagem das variáveis independentes. O modelo pode ser expresso pela equação (20):

$$y = \rho Wy + X\beta + WX\tau + \varepsilon, \quad \varepsilon \sim N(0, \sigma^2 I_n) \quad (20)$$

O Modelo de Defasagem Espacial com Erro de Média Móvel Espacial ou SARMA considera que a dependência ocorre tanto pela defasagem da variável dependente quanto pelo termo do erro. Os efeitos não modelados dos erros possuem alcance local, enquanto os efeitos modelados das variáveis explicativas possuem alcance global. A representação algébrica do modelo encontra-se nas equações (21) e (22):

$$y = \rho W_1 y + X\beta + \xi, \quad |\rho| < 1 \quad (21)$$

$$\xi = \gamma W_2 \varepsilon + \varepsilon, \quad |\gamma| < 1 \quad (22)$$

O Modelo de Durbin Espacial do Erro ou SDEM incorpora componentes espaciais de alcance local por meio das variáveis explicativas exógenas e componente espacial de transbordamento global que afeta o termo de erro. O modelo pode ser representado pelas equações (23) e (24):

$$y = X\beta + WX\tau + \xi \quad (23)$$

$$\xi = \lambda W\xi + \varepsilon, \quad |\lambda| < 1 \quad (24)$$

Por fim, na literatura, existe um modelo que engloba todos os demais termos dos outros modelos: a interação endógena da variável dependente, a interação exógena das variáveis independentes e a autocorrelação espacial dos erros. Pode ser conhecido como Modelo de Manski (GOLGHER, 2015) ou Modelo de Dependência Espacial Geral – GSM (ALMEIDA, 2012). Todos os outros modelos podem ser então derivados deste. Este modelo engloba o transbordamento global e local e pode ser representado algebricamente pelas equações (25) e (26):

$$y = \rho W_1 y + X\beta + W_1 X\tau + \xi, \quad |\rho| < 1 \quad (25)$$

$$\xi = \lambda W_2 \xi + \varepsilon, \quad |\lambda| < 1 \quad (26)$$

A síntese dos modelos espaciais para dados em corte transversal e seus parâmetros pode ser visualizado no Quadro 1.

Quadro 1 - Síntese dos modelos espaciais para dados *cross section* e seus respectivos parâmetros

Modelos	ρ	τ	λ	γ
GSM	$\neq 0$	$\neq 0$	$\neq 0$	$\neq 0$
SDM	$\neq 0$	$\neq 0$	$= 0$	$= 0$
SDEM	$= 0$	$\neq 0$	$\neq 0$	$= 0$
SAC	$\neq 0$	$= 0$	$\neq 0$	$= 0$
SARMA	$\neq 0$	$= 0$	$= 0$	$\neq 0$
SAR	$\neq 0$	$= 0$	$= 0$	$= 0$
SLX	$= 0$	$\neq 0$	$= 0$	$= 0$
SEM	$= 0$	$= 0$	$\neq 0$	$= 0$
SMA	$= 0$	$= 0$	$= 0$	$\neq 0$

Fonte: Elaborado pelo autor

A estimação dos parâmetros de um modelo econométrico-espacial, caso seja realizada por Mínimos Quadrados Ordinários (MQO), da mesma forma que é feito para o Modelo Clássico de Regressão Linear (MCRL), pode se tornar viesada e/ou ineficiente (TYSZLER, 2006). Por isso, outros métodos são empregados na estimação dos mesmos; a saber: Máxima Verossimilhança (CLIFF; ORD, 1973) ou Método das Variáveis Instrumentais (VI).

Após estimado o modelo de regressão espacial, alguns testes podem ser realizados nos resíduos para verificação de dependência espacial. Florax e Graaff (2004) definiram duas taxonomias para os testes: testes difusos e testes focados. Os testes difusos não indicam o tipo da autocorrelação espacial predominante no modelo, já que a hipótese alternativa diz apenas que existe dependência espacial. Em contrapartida, os testes focados buscam, na hipótese alternativa do teste, indicar o tipo predominante de dependência espacial, tratando-se de um modelo espacial específico.

Cliff e Ord (1973) sugerem a utilização do teste difuso I de Moran. O teste apresenta as hipóteses a serem verificadas como H_0 : independência espacial dos erros; e H_1 : dependência espacial dos erros. A representação matemática do teste é dada pela equação 27.

$$I = \frac{n}{S_0} \left(\frac{e'We}{e'e} \right) \quad (27)$$

Onde:

$$S_0 = \sum_i \sum_j W_{ij};$$

$$e = y - X\hat{\beta};$$

$\hat{\beta}$ é o estimador MQO para β .

Existem na literatura algumas metodologias para especificação dos modelos espaciais: a abordagem de especificação clássica, a abordagem de especificação robusta, a abordagem de especificação híbrida e a abordagem de especificação baseada na metodologia de Hendry (FLORAX; FOLMER; REY, 2003). Entretanto, essas abordagens buscam apenas determinar se a autocorrelação espacial possui como fonte a defasagem espacial ou o erro espacial, deixando de lado outras fontes de dependência espacial, como, por exemplo, a defasagem das variáveis independentes (ALMEIDA, 2012).

Por isso, busca-se adotar uma abordagem de especificação mais geral para representar fielmente o mecanismo gerador de dados, que consiste em (ALMEIDA, 2012; TYSZLER, 2006):

- a) ajustar o modelo MCRL por MQO;
- b) testar se há presença de autocorrelação espacial nos resíduos pelos testes difusos;
- c) caso não haja, a etapa 1 está adequada. Se existir, deve-se seguir para a etapa seguinte;
- d) ajustar diversos modelos econométricos espaciais e escolher o melhor modelo de acordo com os seguintes critérios: não apresentar autocorrelação espacial nos seus resíduos e também apresentar um menor critério de informação AIC e BIC.

Para verificação do ajuste da regressão espacial, alguns critérios podem ser adotados no lugar do coeficiente de determinação, que não é adequado quando se trata de erros que são autocorrelacionados e/ou heterocedásticos; são eles: Critério de informação de Akaike (AIC) e o Critério de Informação Bayseano (BIC). Quanto menor os valores dos critérios, melhor pode ser considerado o modelo. Os critérios podem ser representados algebricamente pelas equações (28) e (29), adaptadas de (MORETTIN; TOLOI, 2006).

$$AIC(k, l) = \ln \sigma_{k,l}^2 + \frac{2(k+l)}{N} \quad (28)$$

$$BIC(k, l) = \ln \sigma_{k,l}^2 + (k+l) \frac{\ln N}{N} \quad (29)$$

Onde:

$\sigma_{k,l}^2$ = variância dos erros do modelo;

N = número de observações.

2.2.2.2 Modelos Econométricos Espaciais para dados em painel

Dados em painel são dados que possuem informações de corte transversal das regiões analisadas em diversos momentos do tempo. Embora haja o pressuposto de que modelos convencionais de dados em painéis possuam independência entre os cortes transversais, quando se trata de modelos de dados em painéis nos quais as observações são unidades do espaço, pode haver dependência entre os cortes (DRISCOLL; KRAAY, 1995).

Existem vantagens e desvantagens ao se trabalhar com modelos espaciais em painéis. Como vantagens, é possível modelar não só a dependência espacial, mas também a heterogeneidade espacial. A heterogeneidade espacial decorre da influência que variáveis omitidas do modelo de regressão espacial podem exercer no intercepto e no componente do erro. A omissão dessas variáveis em modelos *cross-section* pode causar viés na estimação (ALMEIDA, 2012; GOLGHER, 2015).

Além disso, os modelos fornecem mais informações, possuem maior variação dos dados e, como consequência, menor colinearidade entre variáveis; possuem maior grau de liberdade e estimação mais eficiente (ALMEIDA, 2012; GOLGHER, 2015).

Como desvantagem, destaca-se a demanda maior de estimação em termos computacionais, comparada a dados em corte, além da dificuldade de obtenção dos dados (ALMEIDA, 2012; GOLGHER, 2015).

O modelo inicial de regressão em painel, sem acomodação para os efeitos não observáveis, pode ser visualizado matematicamente na equação (30):

$$y_t = X_t\beta + \varepsilon_t, \quad \varepsilon_t \sim N(0, \sigma^2 I_n) \quad (30)$$

Onde:

y_t = vetor da variável dependente representado por $(y_{1t}, \dots, y_{nt})'$;

X_t = matriz das variáveis independentes representada por $(X'_{1t}, \dots, X'_{nt})$;

β = vetor de coeficientes das variáveis explicativas;

ε_t = resíduos representados por $(\varepsilon_{1t}, \dots, \varepsilon_{nt})'$.

Para verificar a presença de efeitos não observáveis no modelo, existe o teste de Breusch-Pagan (1979). Ele é um teste de Multiplicador de Lagrange que investiga a heterocedasticidade em modelos lineares. A hipótese nula do teste é de que a variância dos resíduos, devido aos efeitos não observáveis, é zero; ou seja, há presença de efeitos significativos (GOLGHER, 2015). Caso seja confirmada a presença de efeitos não observáveis, pode-se adicionar ao modelo inicial de regressão em painel termos capazes de captar tais efeitos, em um modelo ainda sem dependência espacial (equação 31).

$$y_t = X_t\beta + \alpha + \varepsilon_t \quad (31)$$

Onde:

α = vetor dos efeitos não observados, representado por $(\alpha_1, \dots, \alpha_n)'$;

$E[\varepsilon_{it}\alpha_i] = 0$.

Utiliza-se, ainda, o teste de Hausman (1978) para verificar se os efeitos não observáveis são fixos ou aleatórios. O teste capta se existe correlação entre os regressores e os efeitos não observáveis. Como hipótese nula, o teste apresenta $E[\alpha X_t] = 0$, ou seja, o modelo de efeitos aleatórios é mais eficiente. Já como hipótese alternativa, $E[\alpha X_t] \neq 0$, devendo-se escolher o modelo de efeitos fixos, já que a estimativa do modelo de efeitos aleatórios apresentaria vies e inconsistência (JOHNSTON; DINARDO, 1997).

Deve-se investigar os resíduos do modelo da equação 31, para verificar se eles são autocorrelacionados espacialmente. O teste baseado em randomização da dependência espacial para modelos de painel, dos autores Croissant e Millo (2008), cumpre tal função. Caso não exista dependência espacial, o modelo da equação (31), com os efeitos indicados pelo teste de Hasuman, está adequado. Caso exista dependência espacial, deve-se incorporar ao modelo termos capazes de capturar esse efeito.

Os modelos de efeitos fixos com dependência espacial, assumem que os interceptos das unidades espaciais são diferentes, mas invariantes no tempo, captando as diferenças entre as regiões e $E[\alpha X_t] \neq 0$; ou seja, α e X_t são correlacionados. Os efeitos não observados são parâmetros a serem estimados (ALMEIDA, 2012; GOLGHER, 2015). A formulação geral do modelo, com as diversas fontes de autocorrelação espacial, pode ser representada matematicamente pelas equações (32) e (33):

$$y_t = \alpha + \rho W_1 y_t + X_t \beta + W_1 X_1 \tau + \xi_t \quad (32)$$

$$\xi_t = \lambda W_2 \xi_t + \varepsilon_t \quad (33)$$

Onde:

$W_1 y_t$ representa a defasagem espacial da variável dependente;

$W_2 \xi_t$ representa a defasagem espacial dos erros;

ρ representa o coeficiente autorregressivo espacial;

λ é o parâmetro do erro autorregressivo espacial;

τ é um vetor de coeficientes espaciais.

A partir de restrições aos parâmetros espaciais da formulação geral, é possível extrair modelos como: SAR com efeitos fixos (equação 34), SAC com efeitos fixos (equações 35 e 36), SLX com efeitos fixos (equação 37), SDM com efeitos fixos (equação 38), SDEM com efeitos fixos (equações 39 e 40) e SEM com efeitos fixos (equações 41 e 42). Os conceitos dos modelos espaciais especificados para dados *cross-section* também se mantêm para dados em painel.

$$y_t = \alpha + \rho W_1 y_t + X_t \beta + \varepsilon_t \quad (34)$$

$$y_t = \alpha + \rho W_1 y_t + X_t \beta + \xi_t \quad (35)$$

$$\xi_t = \lambda W_2 \xi_t + \varepsilon_t \quad (36)$$

$$y_t = \alpha + X_t \beta + W_1 X_1 \tau + \varepsilon_t \quad (37)$$

$$y_t = \alpha + \rho W_1 y_t + X_t \beta + W_1 X_1 \tau + \varepsilon_t \quad (38)$$

$$y_t = \alpha + X_t \beta + W_1 X_1 \tau + \xi_t \quad (39)$$

$$\xi_t = \lambda W \xi_t + \varepsilon_t \quad (40)$$

$$y_t = \alpha + X_t\beta + \xi_t \quad (41)$$

$$\xi_t = \lambda W\xi_t + \varepsilon_t \quad (42)$$

Os modelos de efeitos aleatórios com dependência espacial são aqueles em que os efeitos não observáveis são considerados como variáveis aleatórias, ou seja, os efeitos aleatórios não estão correlacionados ao termo do erro, e sim fazem parte deste termo. Este modelo é mais adequado para painéis de dados provenientes de amostras aleatórias, porém, quando as unidades espaciais representam o todo (população), deve-se, *a priori*, especificar um modelo de efeitos fixos, já que a especificação de modelos de efeitos aleatórios nesse caso, torna-se inapropriada (ALMEIDA, 2012; BEENSTOCK; FELDMAN; FELSENSTEIN, 2012). A especificação geral para o modelo (GSPRE – *Generalized Spatial Random-Effects model*) poder ser visualizada a seguir (equações 43 e 44).

$$y_t = \rho W_1 y_t + X_t\beta + W_1 X_1 \tau + \xi_t \quad (43)$$

$$\xi_t = \alpha + \lambda W_2 \xi_t + \varepsilon_t \quad (44)$$

A partir de restrições aos parâmetros espaciais da formulação geral, é possível extrair os seguintes modelos: SAR com efeitos aleatórios (equações 45 e 46), SAC com efeitos aleatórios (equações 47 e 48), SLX com efeitos aleatórios (equações 49 e 50), SDM com efeitos aleatórios (equações 51 e 52), SDEM com efeitos aleatórios (equações 53 e 54) e SEM com efeitos aleatórios (equações 55 e 56).

$$y_t = \rho W_1 y_t + X_t\beta + \xi_t \quad (45)$$

$$\xi_t = \alpha + \varepsilon_t \quad (46)$$

$$y_t = \rho W_1 y_t + X_t\beta + \xi_t \quad (47)$$

$$\xi_t = \alpha + \lambda W_2 \xi_t + \varepsilon_t \quad (48)$$

$$y_t = X_t\beta + W_1X_1\tau + \varepsilon_t \quad (49)$$

$$\xi_t = \alpha + \varepsilon_t \quad (50)$$

$$y_t = \rho W_1y_t + X_t\beta + W_1X_1\tau + \varepsilon_t \quad (51)$$

$$\xi_t = \alpha + \varepsilon_t \quad (52)$$

$$y_t = X_t\beta + W_1X_1\tau + \xi_t \quad (53)$$

$$\xi_t = \alpha + \lambda W\xi_t + \varepsilon_t \quad (54)$$

$$y_t = X_t\beta + \xi_t \quad (55)$$

$$\xi_t = \alpha + \lambda W\xi_t + \varepsilon_t \quad (56)$$

Adicionalmente, é importante calcular os impactos espaciais diretos e indiretos das variáveis explicativas. Este cálculo pode ser realizado para modelos econométricos espaciais com dados em corte transversal ou dados em painel. Os impactos diretos dizem respeito aos efeitos que a mudança em uma variável explicativa de uma determinada região causa na variável dependente da própria região. Já os impactos indiretos, referem-se ao efeito que a mudança na variável explicativa em uma região gera na variável dependente das regiões vizinhas. Os efeitos totais são a soma dos efeitos diretos e indiretos. (LESAGE; PACE, 2009).

3 MATERIAIS E MÉTODOS

Neste capítulo são abordados a natureza da pesquisa, os procedimentos adotados no intuito de cumprir com os objetivos delineados e a fonte dos dados utilizados.

Assim, com base nos seus objetivos, esta pesquisa se caracteriza como explicativa visto que objetiva verificar a influência ou determinação de variáveis sobre um dado fenômeno. Com base nos procedimentos técnicos utilizados, a pesquisa pode ser classificada *ex-post facto*, já que objetiva verificar a relação entre variáveis a partir de um fenômeno que já ocorreu (GIL, 2002).

Quanto ao tratamento dos dados, ele tem uma abordagem quantitativa, pois foram utilizados procedimentos estatísticos (VERGARA, 2016).

Os resultados foram discutidos na forma de um artigo relacionado aos objetivos e ao tema desta dissertação, que versará sobre a aplicação da AEDE sobre a mortalidade por cânceres de próstata e de mama e, caso verificada a autocorrelação espacial, dar-se-á a Modelagem da Dependência Espacial. A base de dados utilizada bem como os procedimentos referentes à Análise Exploratória de Dados Espaciais (AEDE) e à Modelagem da Dependência Espacial, são especificados nas subseções 3.1, 3.2 e 3.3.

3.1 BASE DE DADOS

Os dados utilizados na pesquisa são classificados como dados em corte transversal. A base de dados utilizada é proveniente do Departamento de Informática do Sistema Único de Saúde do Brasil – DATASUS e do Instituto Brasileiro de Geografia e Estatística – IBGE.

O mapa (*shapefile*) das 558 microrregiões brasileiras, utilizado na pesquisa, foi obtido no site do Instituto Brasileiro de Geografia e Estatística – IBGE.

Baseadas na literatura pertinente e em achados de estudos referenciados nesta pesquisa, foram selecionadas variáveis, de acordo com a disponibilidade por microrregião. O Quadro 2 sintetiza todas as variáveis, bem como suas fontes, que serão utilizadas na análise exploratória de dados espaciais (AEDE) e na modelagem econométrica da dependência espacial.

Quadro 2 - Variáveis utilizadas na pesquisa e suas fontes

(continua)

VARIÁVEL	DESCRIÇÃO	FONTE DE DADOS
Taxa de Mortalidade por Câncer de Mama	Número de óbitos por câncer de mama sobre população de mulheres.	DATASUS (http://tabnet.datasus.gov.br/cgi/deftohtm.exe?sim/cnv/obt10br.def)
Taxa de Mortalidade por Câncer de próstata	Número de óbitos por câncer de próstata sobre população de homens	DATASUS (http://tabnet.datasus.gov.br/cgi/deftohtm.exe?sim/cnv/obt10br.def)
Unidades de Referência (ROCHA-BRISCHILIARI et al., 2018)	Número de unidades para tratamento do câncer pelo SUS	DATASUS (http://tabnet.datasus.gov.br/cgi/tabcgi.exe?cnes/cnv/habbr.def)
Números de médicos (ROCHA-BRISCHILIARI et al., 2018)	Número de médicos oncologistas por microrregião	DATASUS (http://tabnet.datasus.gov.br/cgi/deftohtm.exe?cnes/cnv/prid02br.def)
Taxa de envelhecimento (INSTITUTO NACIONAL DE CÂNCER - INCA, 2020)	Número de pessoas residentes de 60 e mais anos de idade sobre número de pessoas residentes com menos de 15 anos de idade, seguindo a projeção da população do IBGE.	SIDRA/IBGE (https://sidra.ibge.gov.br/tabela/3733)
Taxa de urbanização (BARBOSA et al., 2016)	População urbana por microrregião sobre população total por microrregião, de acordo com a projeção da população do IBGE.	SIDRA/IBGE (https://sidra.ibge.gov.br/Tabela/202)
PIB (GUERRA et al., 2020)	Produto Interno Bruto <i>per capita</i> a preços correntes por microrregião (2010)	SIDRA/IBGE (https://sidra.ibge.gov.br/tabela/5938)

Quadro 2 - Variáveis utilizadas na pesquisa e suas fontes

(conclusão)		
Mamógrafo	Número de mamógrafos por microrregião (2018)	CNES/DATASUS (http://tabnet.datasus.gov.br/cgi/deftohtm.exe?cnes/cnv/equipobr.def)
Pesticidas (SILVA et al., 2015)	Área plantada por microrregião (hectares) vezes a quantidade média de utilização de agrotóxicos (6,7 kg/hec)	SIDRA/IBGE (https://sidra.ibge.gov.br/Tabela/1612)

Fonte: Elaborado pelo autor

Para a manipulação dos dados, foram utilizados dois *softwares*: O GeoDa – para a manipulação dos dados na Análise Exploratória de Dados Espaciais (AEDE); posteriormente, confirmado a dependência espacial, os modelos econométricos espaciais serão ajustados por meio de pacotes do *software* R.

3.2 ANÁLISE EXPLORATÓRIA DE DADOS ESPACIAIS

Na primeira etapa foi verificada se as taxas de mortalidade por câncer de próstata e mama estão distribuídas de forma aleatória no espaço ou estão autocorrelacionados. Para tanto, foi utilizada a estatística *I* de Moran Global. Realizou-se ainda um teste para confirmar a significância estatística do *I* de Moran. Caso não seja significativo, deve ser assumido que não existe autocorrelação espacial nos dados em estudo. Confirmada a dependência espacial, é necessário escolher uma matriz de ponderação espacial que represente fielmente o grau de conexão entre as regiões de interesse. Para tal, as seguintes etapas propostas por Baumont (2004) foram seguidas:

- a) o Modelo Clássico de Regressão Linear (MCRL) por Mínimos Quadrados Ordinários (MQO) foi ajustado;
- b) foi testada a autocorrelação espacial, por meio da estatística *I* de Moran, dos resíduos do melhor modelo ajustado com os diferentes tipos de matrizes de ponderação;
- c) a matriz que gerou o maior valor para o teste *I* de Moran foi a selecionada.

Após selecionada a matriz espacial, gerou-se o diagrama de dispersão de Moran para fornecer o padrão de associação linear e a verificação de existência de *outliers* espaciais. Foi

verificado ainda a autocorrelação espacial local por meio do índice I de Moran Local e gerado o mapa de *clusters* para fornecer os padrões de associação espacial estatisticamente significativos.

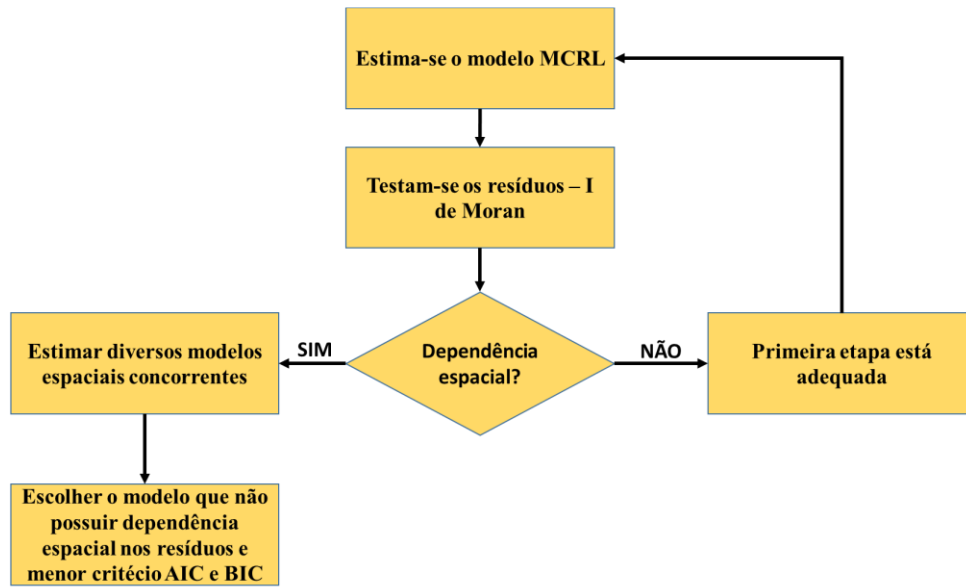
A próxima etapa consiste então na modelagem da dependência espacial.

3.3 MODELAGEM ESPACIAL

Verificada a presença de autocorrelação espacial e definida uma matriz de ponderação espacial, deve-se então modelar a autocorrelação espacial. A modelagem da dependência espacial foi realizada seguindo a metodologia proposta por Almeida (2012) e pode ser visualizada na figura 3:

- a) estima-se o modelo MCRL por MQO para as taxas de mortalidade de câncer de próstata e mama;
- b) verifica-se a presença de autocorrelação espacial nos resíduos pelos testes difusos;
- c) se não houver autocorrelação espacial, a primeira etapa está adequada. Se existir, deve-se seguir para a etapa seguinte;
- d) estimam-se diversos modelos econométricos espaciais concorrentes para cada tipo de mortalidade por câncer e o melhor modelo será aquele que não apresentar autocorrelação espacial nos seus resíduos e também apresentar menor critério de informação AIC e BIC.

Figura 3 – Procedimento de especificação dos modelos espaciais



Fonte: Elaborado pelo autor

4 ARTIGO: ESTUDO ECOLÓGICO ESPACIAL DA MORTALIDADE POR CÂNCER DE PRÓSTATA E MAMA NO BRASIL

RESUMO

No Brasil, o câncer de Próstata (CaP) e Câncer de Mama (CaM) são atualmente os tipos de cânceres mais frequentes em homens e mulheres, respectivamente. Por isso, com o objetivo de determinar a distribuição espacial da mortalidade por CaP e CaM nas microrregiões brasileiras e verificar o seu inter-relacionamento com fatores sociodemográficos, econômicos, ambientais e de infraestrutura, fez-se uso das técnicas da Econometria Espacial: Análise Exploratória de Dados Espaciais – AEDE e Modelagem da dependência espacial. Foram escolhidas variáveis para representar os fatores, de acordo com literatura relevante e disponibilidade de dados, sendo estas: unidades de referência oncológica, número de médicos especialistas em oncologia, taxa de envelhecimento, taxa de urbanização, quantidade média de pesticidas, Produto Interno Bruto e número de mamógrafos. A base de dados utilizada é proveniente do Departamento de Informática do Sistema Único de Saúde do Brasil – DATASUS e do Instituto Brasileiro de Geografia e Estatística – IBGE. Os resultados indicaram autocorrelação espacial positiva para a taxa de mortalidade por câncer de próstata ($I = 0,4537$; $p \leq 0,001$) e câncer de mama ($I = 0,4842$; $p \leq 0,001$). A taxa de mortalidade por câncer de próstata apresentou correlação positiva com a taxa de envelhecimento e taxa de urbanização. A taxa de mortalidade por câncer de mama mostrou-se positivamente correlacionada com número de médicos especialistas, unidades de referência, taxa de urbanização e taxa de envelhecimento, mas negativamente correlacionada com número de mamógrafos. Houve indicativo de que a maior expectativa de vida e o estilo de vida, decorrente da urbanização possuem influência no aumento da mortalidade por câncer de próstata e mama no Brasil.

Palavras-chave: Neoplasias da Próstata. Neoplasias da Mama. Análise Espacial. Regressão Espacial.

1. Introdução

O câncer é uma doença com origem na multiplicação desordenada de células causada por mutações genéticas que codificam as proteínas reguladoras do ciclo celular. Essas alterações provocam diferentes características como a capacidade de multiplicação celular mesmo sem estímulos (BERNARDES et al., 2019). Segundo a Organização Mundial da Saúde (OMS) (ORGANIZAÇÃO MUNDIAL DA SAÚDE, 2018), de algum modo, todas as famílias ao redor do mundo são afetadas pela doença e estima-se que até o ano de 2038 os números de novos casos de câncer devem aumentar em 70%.

Em mulheres, o câncer é a principal causa de morte no mundo, tanto em países desenvolvidos quanto em desenvolvimento. Além disso, a incidência de câncer está aumentando em diversos países devido ao envelhecimento mundial da população (TORRE et

al., 2017). O crescimento na incidência por câncer é esperado principalmente nos países em desenvolvimento, como o Brasil, onde a expectativa de vida está crescendo devido ao avanço nas políticas públicas relacionadas a saúde, como controle de doenças infecciosas e redução na mortalidade infantil (FRANCESCHI; WILD, 2013).

Entre os homens, o segundo tipo de câncer mais frequente e o quinto tipo de câncer que mais mata no mundo é o câncer de próstata (CaP). Para 2018, foi estimada a ocorrência de pelo menos 1,3 milhões de novos casos e 359 mil mortes por CaP entre todos os países (BRAY et al., 2018). No Brasil, entre os homens, excluindo o câncer de pele não-melanoma, o CaP é o tipo de câncer mais comum. Em 2018 houve 15.576 mortes pela doença e espera-se que em 2020 sejam registrados 65.840 novos casos (INSTITUTO NACIONAL DE CÂNCER - INCA, 2020).

O câncer de próstata apresenta como fatores de risco a idade, que possui impacto tanto na incidência quanto na mortalidade; pai ou irmão com a doença antes dos 60 anos de idade; excesso de gordura corporal e exposição a aminas aromáticas, ao arsênio, derivados de petróleo, hidrocarbonetos policíclicos aromáticos (HPA), fuligem e dioxinas (INSTITUTO NACIONAL DE CÂNCER - INCA, 2019a).

O câncer de mama (CaM) é o tipo de câncer mais frequente entre as mulheres tanto no Brasil como no mundo. Em 2018 houve 17.572 mortes por câncer de mama, em mulheres, no Brasil e para 2020 estimavam-se 66.280 novos casos. Excetuando o câncer de pele não melanoma, o CaM é o mais frequente entre as mulheres nas regiões Sul, Sudeste, Centro-Oeste e Nordeste. Este tipo de câncer também acomete homens, porém é considerado raro e o total de casos é de apenas 1% (INSTITUTO NACIONAL DE CÂNCER - INCA, 2020).

A idade é o principal fator de risco para o desencadeamento do câncer de mama, outros fatores que impactam na morbidade da doença são fatores ambientais e comportamentais, como sobrepeso, sedentarismo, consumo de bebidas alcólicas; fatores da história reprodutiva e hormonal, como menopausa, não ter tido filhos, uso de contraceptivos hormonais; e fatores genéticos e hereditários, como histórico familiar e alterações genéticas (INSTITUTO NACIONAL DE CÂNCER - INCA, 2019c).

Para os portadores do câncer de mama, foi observado que um dos fatores que estava associado ao agravamento da doença após o diagnóstico foi o fato de o paciente residir em uma localização geográfica mal assistida por serviços oncológicos. O adequado processo de planejamento e avaliação dos serviços de saúde leva em consideração questões geográficas de localização e o deslocamento necessário para acesso ao tratamento (BARRY; BREEN, 2005; OLIVEIRA et al., 2011a).

A análise dos padrões espaciais do câncer no Brasil ainda é pouco explorada e se configura como uma importante ferramenta para o planejamento e implementação de novos programas de controle de neoplasias. Encontrar correlação da mortalidade da doença com outras variáveis pode ajudar a identificar hipóteses que mereçam futura investigação epidemiológica para contribuir com políticas públicas relacionadas à saúde (HERRMANN et al., 2015; MADHU et al., 2016).

Guerra et al. (2020), Torres-Roman et al. (2019), Amin, Fritsch e Retzlöff (2019), Rocha-Brischiliari et al. (2018), Góes et al. (2018), Núñez et al. (2016), Barbosa et al. (2016), Mueller et al. (2015), Silva et al. (2015) e Vinnakota e Lam (2006) já utilizaram as técnicas de econometria espacial para investigar a mortalidade e/ou incidência dos cânceres de próstata e mama, verificar a existência de dependência espacial e seu inter-relacionamento com outras variáveis. Dentre essas variáveis estão a idade, taxa de urbanização, áreas mais desenvolvidas, exposição a pesticidas e acessibilidade aos serviços de saúde oncológicos. Essas e outras variáveis podem ser testadas para a existência de associação com as taxas de mortalidade por câncer de próstata e mama nas microrregiões brasileiras.

O objetivo desta pesquisa é determinar as microrregiões brasileiras com maior incidência de câncer de próstata e mama em mulheres no ano de 2018 e identificar se a distribuição espacial da taxa de mortalidade por câncer de próstata e mama, em mulheres, nas microrregiões do Brasil está inter-relacionada com as variáveis unidades de referência oncológica, número de médicos especialistas em oncologia, taxa de envelhecimento, taxa de urbanização, quantidade média de pesticidas, Produto Interno Bruto e número de mamógrafos.

2. Metodologia

A pesquisa proposta é um estudo ecológico para investigar a distribuição espacial da taxa de mortalidade por câncer de próstata e mama, em mulheres, e busca determinar o grau de interação entre a doença e as variáveis e se existe um efeito de dependência entre as regiões. Os dados são em corte transversal referentes ao ano de 2018 e foram coletados no portal do Departamento de Informática do Sistema Único de Saúde do Brasil – DATASUS e do Instituto Brasileiro de Geografia e Estatística – IBGE. O mapa (*shapefile*) das 558 microrregiões brasileiras, utilizado na pesquisa, foi obtido no site do Instituto Brasileiro de Geografia e Estatística – IBGE.

Baseados na literatura relacionada ao tema deste estudo, foram selecionadas variáveis das microrregiões, sintetizadas no Quadro 1.

Quadro 1 - Variáveis utilizadas na pesquisa e suas fontes

(continua)

VARIÁVEL	DESCRIÇÃO	FONTE DE DADOS
Taxa de Mortalidade por Câncer de Mama	Número de óbitos por câncer de mama sobre população de mulheres.	DATASUS (http://tabnet.datasus.gov.br/cgi/deftohtm.exe?sim/cnv/obt10br.def)
Taxa de Mortalidade por Câncer de próstata	Número de óbitos por câncer de próstata sobre população de homens.	DATASUS (http://tabnet.datasus.gov.br/cgi/deftohtm.exe?sim/cnv/obt10br.def)
Unidades de Referência (ROCHA-BRISCHILIARI et al., 2018)	Número de unidades para tratamento do câncer pelo SUS.	DATASUS (http://tabnet.datasus.gov.br/cgi/tabcgi.exe?cnes/cnv/habbr.def)
Números de médicos (ROCHA-BRISCHILIARI et al., 2018)	Número de médicos oncologistas por microrregião.	DATASUS (http://tabnet.datasus.gov.br/cgi/deftohtm.exe?cnes/cnv/prid02br.def)
Taxa de envelhecimento (INSTITUTO NACIONAL DE CÂNCER - INCA, 2020)	Número de pessoas residentes de 60 e mais anos de idade sobre número de pessoas residentes com menos de 15 anos de idade, seguindo a projeção da população do IBGE.	SIDRA/IBGE (https://sidra.ibge.gov.br/tabela/3733)
Taxa de urbanização (BARBOSA et al., 2016)	População urbana por microrregião sobre população total por microrregião, de acordo com a projeção da população do IBGE.	SIDRA/IBGE (https://sidra.ibge.gov.br/Tabela/202)
PIB (GUERRA et al., 2020)	Produto Interno Bruto <i>per capita</i> a preços correntes por microrregião (2010)	SIDRA/IBGE (https://sidra.ibge.gov.br/tabela/5938)

Quadro 1 - Variáveis utilizadas na pesquisa e suas fontes

(conclusão)		
Mamógrafo	Número de mamógrafos por microrregião (2018)	CNES/DATASUS (http://tabnet.datasus.gov.br/cgi/deftohtm.exe?cnes/cnv/equipobr.def)
Pesticidas (SILVA et al., 2015)	Área plantada por microrregião (hectares) vezes a quantidade média de utilização de agrotóxicos (6,7 kg/hec)	SIDRA/IBGE (https://sidra.ibge.gov.br/Tabela/1612)

Fonte: Elaborado pelo autor

As análises de dados foram realizadas por meio dos softwares GeoDa e R 3.6.3. As políticas que envolvem estudos com seres humanos foram respeitadas, porém a aprovação por um comitê de ética não foi necessária, pois os dados em análise são de origem secundária.

2.1 Análise Exploratória de Dados Espaciais

A Análise Exploratória de Dados Espaciais (AEDE) pode ser definida como a coleção de técnicas para descrever e visualizar distribuições espaciais, identificar locais atípicos (*outliers* espaciais), descobrir padrões de associação espacial (*clusters*) e sugerir diferentes regimes espaciais e outras formas de instabilidade espacial ou não estacionariedade espacial (ANSELIN, 1997). O objetivo da AEDE é possibilitar ao pesquisador conhecer melhor os dados com que está trabalhando e permitir captar as interrelações dos dados espaciais.

Inicialmente, verifica-se a distribuição espacial das taxas de mortalidade por CaP e CaM nas diferentes microrregiões brasileiras. As taxas de mortalidade por CaP e CaM são calculadas por meio da suavização bayesiana empírica para melhorar a precisão das taxas brutas, minimizando a variabilidade aleatória das taxas por região (ANSELIN, 2005).

A próxima etapa é verificar se as taxas de mortalidade por câncer de próstata e mama estão distribuídas de forma aleatória no espaço ou se estão autocorrelacionadas espacialmente. A estatística *I* de Moran Global é utilizada e definida a matriz *Queen* de ordem 1, como matriz de ponderação espacial.

O resultado para o teste I de Moran varia de -1 a 1 e, quando há autocorrelação espacial positiva, existe a concentração dos dados. Microrregiões com altas taxas de mortalidade estão cercadas por microrregiões vizinhas com altas taxas de mortalidade (Alto-Alto) ou microrregiões com baixas taxas estão cercadas por microrregiões vizinhas com baixas taxas (Baixo-Baixo). A dependência espacial negativa pressupõe dispersão dos dados; localidades com baixas taxas de mortalidade estão cercadas por localidades vizinhas com altas taxas de mortalidade (Baixo-Alto) ou localidades com altas taxas de mortalidade estão cercadas por localidades vizinhas com baixas taxas de mortalidade (Alto-Baixo). Se o resultado do teste for zero, diz-se que as observações são independentes da localização.

Se a autocorrelação espacial for significativa, será verificado os diferentes regimes de associação espacial de cada microrregião por meio do I de Moran Local e os seus resultados expressos na forma de mapa de *clusters* significativos.

2.2 Modelagem Econométrica Espacial

O Modelo Clássico de Regressão Linear (MCRL) por Mínimos Quadrados Ordinários (MQO) deverá ser estimado para as taxas de CaP e CaM (equação 1). Como variáveis independentes atribui-se unidades de referência oncológica, número de médicos especialistas em oncologia, taxa de envelhecimento, taxa de urbanização, quantidade média de pesticidas, Produto Interno Bruto e número de mamógrafos (apenas para a taxa de CaM). Os resíduos dos dois modelos devem ser testados para a presença de autocorrelação espacial por meio da estatística I de Moran, pois não havendo dependência espacial nos resíduos o modelo aplicado pode ser considerado adequado.

$$y = X\beta + \varepsilon, \quad \varepsilon \sim N(0, \sigma^2 I_n) \quad (1)$$

Onde:

X é matriz de variáveis explicativas exógenas;

β representa o coeficiente de regressão;

ε é o erro estocástico.

Se existir dependência espacial significativa nos resíduos, o MCRL pode apresentar viés na estimação por MQO, sendo necessário o ajuste dos modelos econométricos espaciais. Incorpora-se ao modelo de regressão linear convencional termos capazes de capturar esses efeitos, como: Wy , $W\xi$ ou $W\varepsilon$ e WX , que representam a defasagem espacial da variável

dependente, do termo do erro e das variáveis independentes, respectivamente. A formulação do modelo de dependência espacial geral (GSM) pode ser vista nas equações 2 e 3.

$$y = \rho Wy + X\beta + WX\tau + \xi, \quad |\rho| < 1 \quad (2)$$

$$\xi = \lambda W\xi + \varepsilon, \quad |\lambda| < 1 \quad (3)$$

Onde:

ρ representa o coeficiente autorregressivo espacial;

Wy representa a defasagem espacial da variável dependente;

WX representa a defasagem espacial das variáveis independentes;

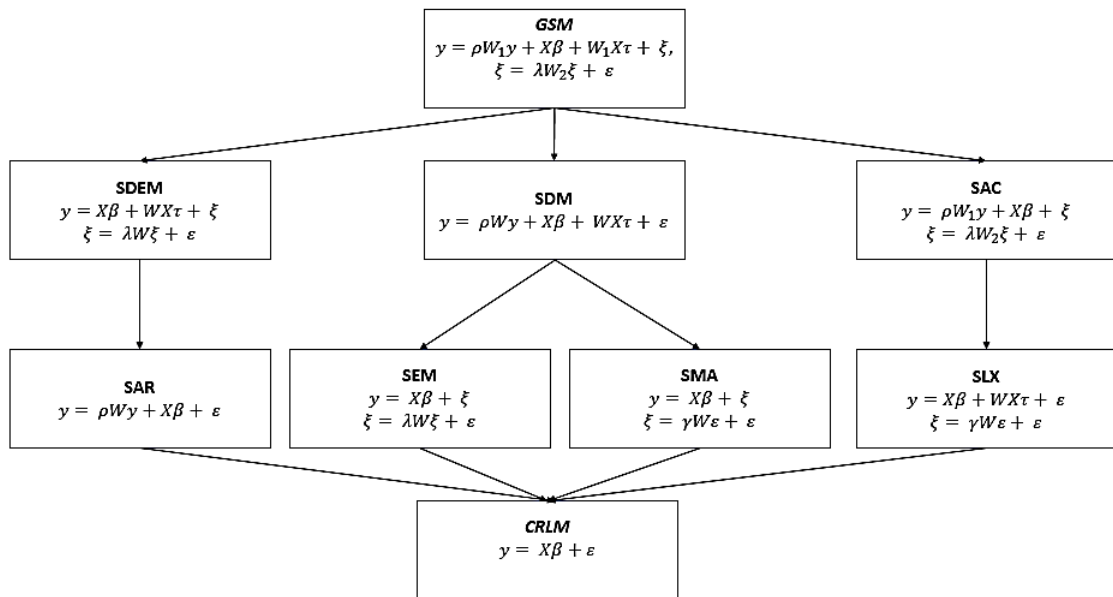
τ é um vetor de coeficientes espaciais.

λ é o parâmetro do erro autorregressivo espacial;

$W\xi$ representa a defasagem espacial do termo do erro.

A partir de restrições nos parâmetros do modelo de dependência espacial geral, podem ser extraídos os seguintes modelos espaciais: SAR (*Spatial Auto Regressive Model*), SEM (*Spatial Error Model*), SAC (*Spatial Autocorrelation Model*), SLX (*Spatially Lagged X Model*), SDM (*Spatial Durbin Model*), SDEM (*Spatial Durbin Error Model*) e SMA (*Spatial Moving Average Model*), conforme figura 1.

Figura 1 – Aninhamento dos modelos espaciais



Fonte: Elaborado pelo autor

O Modelo de *Lag* Espacial (SAR) foi proposto por Whittle (WHITTLE, 1954) e pode ser representado algebricamente pela equação (4):

$$y = \rho W y + X\beta + \varepsilon, \quad \varepsilon \sim N(0, \sigma^2 I_n) \quad (4)$$

Esta especificação implica que cada valor de Y é função de um grupo de variáveis explicativas X e uma média ponderada dos valores da variável dependente (Wy) para as observações vizinhas. O parâmetro ρ indica o grau de correlação espacial entre as observações e, caso seja positivo, há autocorrelação espacial positiva; se ρ for negativo, a autocorrelação espacial é negativa. Se o parâmetro ρ não for significativo, pode-se considerar que o coeficiente é nulo, não existindo indicação de autocorrelação espacial (KLIER; MCMILLEN, 2008).

O Modelo de Erro Autorregressivo Espacial ou SEM foi proposto por Ord (ORD, 1975) e possui a seguinte representação matemática:

$$y = X\beta + \xi \quad (5)$$

$$\xi = \lambda W \xi + \varepsilon \quad (6)$$

Nesse modelo, a dependência espacial se dá apenas pelo erro e não pela variável dependente. No modelo SEM observa-se que o erro estruturado espacialmente (ξ) é uma média dos erros nas regiões vizinhas ($\lambda W\xi$) mais um erro aleatório não explicado pelo modelo (ε). E λ representa o parâmetro do erro autorregressivo que é ponderado por uma matriz de pesos, indicando a defasagem do erro ($W\xi$). Quando não há correlação espacial entre os erros, $\lambda = 0$.

O Modelo de Defasagem Espacial com Erro Autorregressivo Espacial ou SAC representa o modelo tradicional de regressão por MQO acrescido de uma defasagem espacial e um termo de erro autocorrelacionado espacialmente. Nesse modelo a dependência espacial ocorre tanto na variável dependente quanto nos erros (KELEJIAN; PRUCHA, 1998), representado pelas equações (7) e (8):

$$y = \rho W_1 y + X\beta + \xi, \quad |\rho| < 1 \quad (7)$$

$$\xi = \lambda W_2 \xi + \varepsilon, \quad |\lambda| < 1 \quad (8)$$

O Modelo de Erro de Média Móvel Espacial ou SMA foi inicialmente proposto por Haining (HAINING, 1978) e pode ser expresso pelas equações (9) e (10):

$$y = X\beta + \xi \quad (9)$$

$$\xi = \gamma W\varepsilon + \varepsilon, \quad |\gamma| < 1 \quad (10)$$

Por se tratar de um modelo de alcance local, os efeitos espaciais ficam limitados para os vizinhos de primeira e segunda ordem. Observa-se que o erro espacial segue uma dinâmica de média móvel de primeira ordem, onde o erro é composto pelos efeitos na região de origem e pela média dos efeitos nas regiões adjacentes.

O Modelo Regressivo Cruzado Espacial ou SLX é o modelo de regressão linear espacial mais simples capaz de captar o transbordamento espacial local (*spillover effect*) por meio do parâmetro WX . Algebricamente o modelo pode ser expresso pela equação (11):

$$y = X\beta + WX\tau + \varepsilon \quad (11)$$

O Modelo de Durbin Espacial ou SDM possui alcance global por meio do parâmetro da variável dependente defasada e alcance local que se dá pela defasagem das variáveis independentes. O modelo pode ser expresso pela equação (12):

$$y = \rho W y + X\beta + WX\tau + \varepsilon, \quad \varepsilon \sim N(0, \sigma^2 I_n) \quad (12)$$

O Modelo de Durbin Espacial do Erro ou SDEM incorpora componentes espaciais de alcance local por meio das variáveis explicativas exógenas e componente espacial de transbordamento global que afeta o termo de erro. O modelo pode ser representado pelas equações (13) e (14):

$$y = X\beta + WX\tau + \xi \quad (13)$$

$$\xi = \lambda W\xi + \varepsilon, \quad |\lambda| < 1 \quad (14)$$

Após a estimação de todos os modelos espaciais concorrentes, para verificar qual o modelo mais adequado, será escolhido aquele que não apresentar indícios de dependência espacial significativa nos resíduos e apresentar menor valor numérico para o critério penalizador *Akaike Information Criterion* – AIC (Akaike 1974; Almeida 2012).

$$AIC(k, l) = \ln \sigma_{k,l}^2 + \frac{2(k+1)}{N} \quad (15)$$

Onde:

$\sigma_{k,l}^2$ = a variância dos erros do modelo;

N = número de observações.

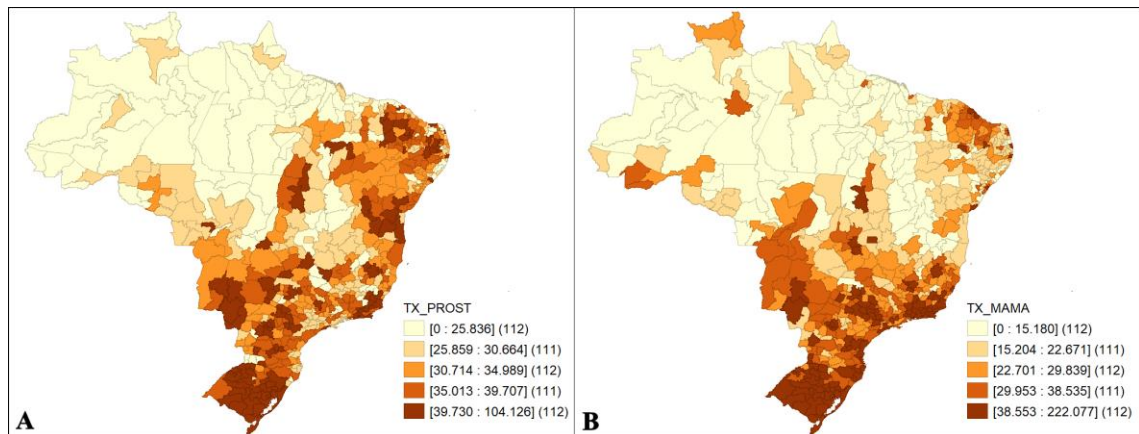
3. RESULTADOS

No ano de 2018, no Brasil, homens com 55 anos de idade ou mais representaram 98,38% dos casos de mortalidade por câncer de próstata. A faixa etária que mais concentrou o número de óbitos foi de homens com 75 anos e mais (62,43% dos casos). A cor de pele branca, foi a

mais predominante, com 50,72% dos casos, seguida da cor parda (35%). Para o câncer de mama, no mesmo ano no Brasil, mulheres a partir dos 45 anos representaram 86,31% dos casos. A maior concentração esteve em mulheres com 75 anos e mais (23,60%). Em relação a cor de pele, mulheres brancas representaram 58,12%, seguida de mulheres pardas (30,52%).

Os resultados para a distribuição espacial da taxa de mortalidade por câncer de próstata e mama, nas microrregiões brasileiras, mostram uma maior concentração dos casos nas regiões centro-sul e nordeste do Brasil (Figura 2).

Figura 2 – Mapa coroplético da taxa de mortalidade por câncer de próstata (A) e mama (B) em 2018 nas microrregiões do Brasil.



Fonte: Elaborado pelo autor

Barretos, microrregião localizada no estado de São Paulo, possui as maiores taxas de mortalidade por 100 mil habitantes de câncer de próstata (104) e câncer de mama (222). O câncer de próstata, destaca-se nas microrregiões de Macau – RN (86), Campanha Central – RS (76), Nova Friburgo – RJ (68) e Campanha Meridional – RS (64). Para o câncer de mama, destacam-se ainda, Muriaé – MG (75), Serrana – RJ (60), Macau-Caceribu – RJ (60) e Jaguarão – RS (60).

A investigação da autocorrelação espacial global da taxa de mortalidade por câncer de próstata e mama, nas microrregiões brasileiras em 2018, pode ser visualizada na Tabela 1.

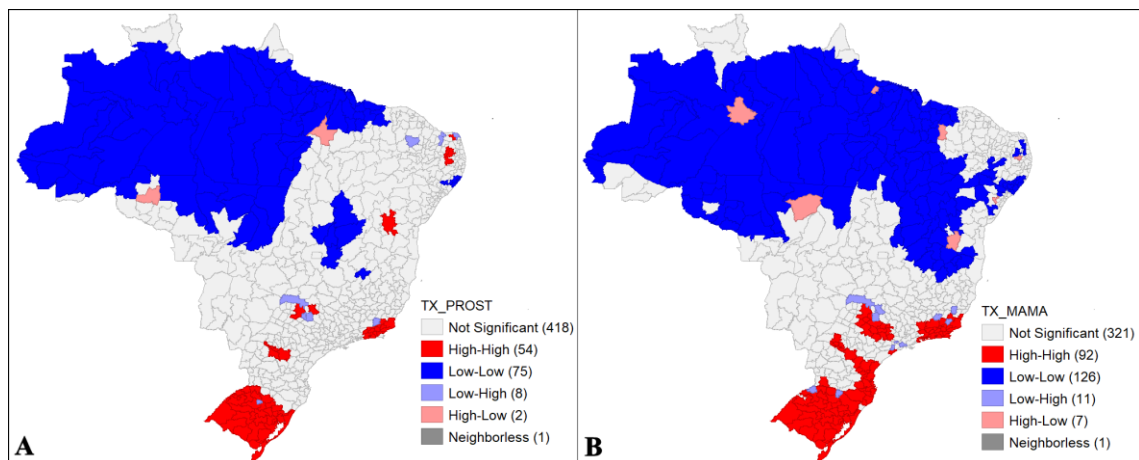
Tabela 1. Resultado para análise da autocorrelação espacial global das taxas de mortalidade por câncer de próstata e mama nas microrregiões brasileiras em 2018.

	Matriz	I de Moran	p-valor
Câncer de Mama	<i>Queen</i>	0,484248298	$p \leq 0,001$
Câncer de Próstata	<i>Queen</i>	0,453742841	$p \leq 0,001$

Fonte: Elaborado pelo autor

Há autocorrelação espacial global positiva e significativa para as taxas de mortalidade por câncer de próstata e mama nas microrregiões brasileiras em 2018. A autocorrelação espacial positiva pressupõe concentração dos dados, indicando predominância de padrões espaciais do tipo Alto-Alto e Baixo-Baixo. O resultado para a análise dos padrões espaciais da mortalidade para os dois tipos de cânceres, nas microrregiões brasileiras, encontra-se na Figura 3.

Figura 3 – Mapa de clusters LISA da taxa de mortalidade por câncer de próstata (A) e mama (B) em 2018 nas microrregiões do Brasil.



Fonte: Elaborado pelo autor

A região norte do país concentra as microrregiões que possuem baixa mortalidade por câncer de próstata e mama e estão cercados por microrregiões vizinhas com baixa mortalidade por CaP e CaM. A região sul do país, o estado de São Paulo e o estado do Rio de Janeiro concentram as microrregiões que possuem elevadas taxas de mortalidade por CaP e CaM e que têm vizinhos com iguais padrões espaciais.

Ajustou-se o modelo clássico de regressão linear para as taxas de mortalidade por CaP e CaM. Como variáveis explicativas encontram-se as unidades de referência oncológica, número de médicos especialistas em oncologia, taxa de envelhecimento, taxa de urbanização,

quantidade média de pesticidas, Produto Interno Bruto e número de mamógrafos. Os resíduos dos dois modelos são testados para a presença de autocorrelação espacial, pelo teste I de Moran (Tabela 2).

Tabela 2 – Resultado da estimação dos parâmetros da regressão linear simples para a mortalidade por câncer de próstata e mama.

Variáveis explicativas	Variável dependente	
	CaP	CaM
Constante	15,94 (8,795)***	-6,571 (-2,67)**
Unidades de referência	$4,359 \times 10^{-1}$ (1,005)	2,361 (4,44)***
Nº de médicos especialistas em oncologia	$-3,611 \times 10^{-3}$ (-0,285)	$8,847 \times 10^{-2}$ (3,108)***
Nº de mamógrafos	-	$-1,42 \times 10^{-1}$ (-3,141)**
Taxa de envelhecimento	36,91 (14,84)***	40,10 (11,908)***
Taxa de urbanização	1,584 (0,538)	25,319 (7,34)***
PIB	$-5,622 \times 10^{-3}$ (-0,207)	$1,442 \times 10^{-1}$ (3,921)***
Pesticidas	$2,77 \times 10^{-7}$ (1,416)	$-4,08 \times 10^{-7}$ (-1,566)
R²	0,3219	0,4235
AIC	4337,439	4337,439
I de Moran	$p \leq 0,01$	$p \leq 0,01$

Entre parênteses está a estatística t
 ***Estatisticamente significativa a 1%
 **Estatisticamente significativa a 5%

Os resultados para o teste I de Moran nos resíduos dos modelos indicam que há presença de autocorrelação espacial estatisticamente significativa, tornando a estimação por MQO viesada (ALMEIDA, 2012). Ajustou-se os modelos espaciais SAR, SEM, SAC, SLX, SDM, SDEM e SMA para as taxas de mortalidade de CaP e CaM. O critério de seleção para o melhor modelo foi aquele que apresentou menor critério penalizador AIC e no qual não houve indícios de autocorrelação espacial nos resíduos, conforme Tabela 3.

Tabela 3 - Estimativa dos parâmetros e impactos espaciais para para a mortalidade por câncer de próstata e mama.

(continuação)

Variáveis	Câncer de Próstata	Câncer de Mama
	SDM	SAR
Constante	9,55669 (4,5841)***	-7,7288 (-3,5654)***
Nº de médicos especialistas em oncologia	-	0,0964 (3,7576)***
Nº de mamógrafos	-	-0,1597 (-3,9147)***
Unidades de Referência	-	2,5776 (4,5421)***
Taxa de Envelhecimento	24,3502 (5,1586)***	20,2711 (5,7736)***
Taxa de Urbanização	8,79588 (3,4704)***	18,5787 (5,7423)***
Wx_Taxa de Envelhecimento	-0,93743 (-0,1587)***	-
Wx_Taxa de Urbanização	-12,43978 (-3,4094)***	-
ρ	0,4901 (75,649)***	0,48 (10,18)***
R2	0,4571	0,5304
AIC	3901,9	4256,4
I de Moran	p-valor = 0,751	p-valor = 0,677

Tabela 3 - Estimativa dos parâmetros e impactos espaciais para para a mortalidade por câncer de próstata e mama.

(conclusão)

Impactos espaciais						
	Direto	Indireto	Total	Direto	Indireto	Total
Nº de médicos especialistas em oncologia	-	-	-	0,1017	0,0836	0,1853
Nº de mamógrafos	-	-	-	-0,1685	-0,1384	-0,3069
Unidades de Referência	-	-	-	2,7189	2,2337	4,9526
Taxa de Envelhecimento	25,6447	20,2330	45,8777	21,3824	17,5665	38,9490
Taxa de Urbanização	7,8389	-14,9566	-7,1177	19,5973	16,10	35,6973

Entre parênteses está a estatística z

*** Estatisticamente significativa a 1%

Fonte: Elaborado pelo autor

O modelo que melhor se adequou à taxa de mortalidade por câncer de próstata foi o modelo de Durbin Espacial. As variáveis significativas foram taxa de envelhecimento e taxa de urbanização, apresentando correlação positiva entre a taxa de mortalidade por câncer de próstata e as variáveis taxa de urbanização e taxa de envelhecimento. Este inter-relacionamento indica que a maior taxa de mortalidade por CaP em uma dada microrregião está relacionada com maiores taxas de urbanização e envelhecimento, na mesma microrregião.

O parâmetro ρ espacial, que determina a correlação espacial, é positivo e significativo, indicando correlação positiva entre a taxa de mortalidade por CaP em uma dada microrregião e nas microrregiões vizinhas. A análise dos impactos espaciais para a taxa de mortalidade por CaP permite inferir que a taxa de envelhecimento gera impactos positivos e tem maior influência sobre a taxa de mortalidade por CaP dentro da mesma microrregião analisada.

A taxa de urbanização gera um impacto espacial direto positivo e um impacto espacial indireto negativo. O aumento em um ponto percentual na taxa de urbanização gera aumento de 7,84% na mortalidade por CaP da mesma microrregião e o aumento em um ponto percentual

na taxa de urbanização da microrregião em análise gera um decréscimo de 14,95% na microrregião vizinha.

Em relação à taxa de mortalidade por câncer de mama, o modelo selecionado foi um SAR. As variáveis significativas ao modelo foram número de médicos especialistas em oncologia, número de mamógrafos, unidades de referência, taxa de urbanização e taxa de envelhecimento. Há correlação positiva entre número de médicos especialistas, unidades de referência, taxa de urbanização, taxa de envelhecimento e a taxa de mortalidade por câncer de mama. Em relação ao número de mamógrafos e a taxa de mortalidade por CaM, a correlação é negativa. Este relacionamento significa que a maior mortalidade por câncer de mama em uma dada microrregião está relacionada ao maior número de médicos especialistas, maior número de unidades de referência para tratamento da doença, maior taxa de envelhecimento populacional, maior taxa de urbanização e menor número de mamógrafos.

O parâmetro ρ espacial é positivo e significativo, o que significa que há correlação positiva entre a taxa de mortalidade por CaM em uma dada microrregião e nas microrregiões vizinhas. Os impactos espaciais das variáveis explicativas possuem maior influência dentro da própria microrregião em análise.

4. DISCUSSÃO

No ano de 2018, a taxa de mortalidade por câncer de próstata e mama em mulheres, no Brasil, esteve concentrada principalmente nas microrregiões da região centro-sul e nordeste. Este resultado é semelhante ao encontrado por Silva et al. (SILVA et al., 2015) que apontou que os estados da região Sul, Centro-Oeste (Mato Grosso e Mato Grosso do Sul), Sudeste (Rio de Janeiro e Espírito Santo) e Nordeste (Pernambuco, Piauí e Sergipe) possuíam as maiores taxas de mortalidade por câncer de próstata no Brasil entre 2005 e 2009.

Para o vice diretor da Associação Brasileira de Cancerologia, a concentração de casos de câncer, principalmente na região sul do Brasil, deve-se à alta urbanização e à longevidade da população (BORTOLIN, 2012). Ou seja, há heterogeneidade na distribuição espacial das doenças devido à diferença nos perfis sociodemográficos entre as microrregiões brasileiras.

A microrregião de Barretos, no estado de São Paulo, apresentou as maiores taxas de mortalidade por CaP e CaM. Este resultado pode ser entendido por estar localizado nesta microrregião o Hospital de Amor de Barretos, a unidade de referência que mais atende casos de câncer pelo Sistema Único de Saúde – SUS (G1, 2016).

Houve correlação positiva entre taxa de mortalidade por CaP e CaM e a taxa de urbanização, nas microrregiões. Este resultado está de acordo com a pesquisa desenvolvida nos municípios brasileiros que indicou que a distribuição espacial das taxas de mortalidade por câncer estão associadas a áreas mais desenvolvidas e urbanizadas (BARBOSA et al., 2016).

Embora não exista um padrão global para a ocorrência dos diversos tipos de câncer, os fatores ambientais relacionados à urbanização como o estilo de vida ocidental possuem importante influência. A utilização de tabaco, consumo de álcool, obesidade, dieta, inatividade física, estresse psicológico, estresse emocional são apontados como fatores do estilo de vida que contribuem com os casos e mortalidade por câncer (DEEPA; KANT, 2016; GARÓFOLO et al., 2004). Por esse motivo, áreas mais urbanizadas tem maior propensão ao surgimento de novos casos e aumento da mortalidade da doença.

Estudo desenvolvido no Estado do Paraná indicou associação linear positiva entre a acessibilidade ao tratamento do câncer e mortalidade por câncer de mama (ROCHA-BRISCHILIARI et al., 2018). Este mesmo inter-relacionamento foi encontrado nesta pesquisa. Microrregiões com maior número de unidades de referência e maior número de médicos especialistas possuem maiores taxas de mortalidade por câncer de mama.

Este fato pode ser explicado devido a existência de gargalos estruturais na rede assistencial do SUS no que diz respeito aos atendimentos, tratamentos e infraestruturas. A Secretaria de Atenção à Saúde já havia admitido a necessidade de pelo menos mais 70 novos hospitais de alta complexidade para o tratamento oncológico no Brasil. Há concentração de unidades de referência e médicos especialistas nos grandes centros, o que provoca a migração dos usuários, filas de espera e aumento das taxas de mortalidade nestas localidades. Estima-se que grande parte dos pacientes resida a mais de 150 km do local de atendimento (BANNA; GONDINHO, 2019; HARTT, 2015; SALES et al., 2011).

Houve correlação negativa entre o número de mamógrafos e a taxa de mortalidade por CaM. Microrregiões com maior número do aparelho responsável pela detecção do câncer de mama, possuem menores taxas de mortalidade da doença. O diagnóstico tardio pode ser considerado determinante para a mortalidade por câncer. Quando realizado precoce, este possibilita um tratamento adequado e melhores prognósticos (OLIVEIRA et al., 2011a).

Foi encontrada correlação positiva entre a taxa de envelhecimento e as taxas de mortalidade por CaP e CaM. A idade, principalmente acima dos 50 anos, é o principal fator de risco para os cânceres de próstata e mama (Instituto Nacional de Câncer - INCA 2020).

O enfrentamento desta doença requer atuação conjunta do poder público, entidades de Saúde, universidades, Organizações não Governamentais (ONGs) e da sociedade civil. Já

existem alguns planos para o enfrentamento da doença, instituídos pelo governo federal. Em 2011, foi instituído pelo governo federal o “Plano de Ações Estratégicas para o Enfrentamento das Doenças Crônicas não Transmissíveis (DCNT) no Brasil 2011-2022” e posteriormente, em 2013 a “Política Nacional para a Prevenção e Controle do Câncer na Rede de Atenção à Saúde das Pessoas com Doenças Crônicas no âmbito do Sistema Único de Saúde (SUS)”. Ambos visam à redução da mortalidade e incapacidade causada pelo câncer. Dentre as metas, podem ser citadas promoção de hábitos alimentares saudáveis, de atividades físicas, enfrentamento do tabagismo, do consumo de álcool, do sobrepeso e detecção precoce da doença (BRASIL, 2011, 2013).

Nota-se, entretanto, a necessidade de maior verificação e controle das metas estabelecidas por parte dos órgãos responsáveis, para que sejam tomadas decisões de enfrentamento da doença conforme o cenário atual.

Parte do serviço oncológico prestado pelo SUS advém de Instituições sem fins lucrativos. A Associação Brasileira de Instituições Filantrópicas de Combate ao Câncer – ABIFICC – conta com 27 associados e em 2017 foi responsável por 25% das aplicações em quimioterapia, 33% em radioterapia, 25% dos procedimentos cirúrgicos e 18% das internações clínicas pelos SUS (ASSOCIAÇÃO BRASILEIRA DE INSTITUIÇÕES FILANTRÓPICAS DE COMBATE AO CÂNCER – ABIFICC, 2017).

Entre seus associados, está o Hospital de Amor de Barretos, localizado na microrregião de Barretos em São Paulo. Este Hospital é o que mais atende casos de câncer pelo SUS. A manutenção financeira de um complexo hospitalar tão importante como este é proveniente parte do Governo Federal, Estado e Município e parte de doações. As doações são captadas pelo site do hospital, campanhas, leilões e vendas de artigo (HOSPITAL DE AMOR, 2020). Este modelo de integração com a participação ativa da sociedade civil merece expansão e pode ser um caminho para o avanço no combate à mortalidade por câncer no Brasil.

5. CONSIDERAÇÕES FINAIS

A metodologia utilizada possibilitou identificar, para a taxa de mortalidade por câncer de próstata, correlação positiva com a taxa de envelhecimento e taxa de urbanização. Para a taxa de mortalidade por câncer de mama, foi encontrada correlação positiva com número de médicos especialistas em oncologia, unidades de referência, taxa de urbanização e taxa de envelhecimento. E correlação negativa com número de mamógrafos.

Houve indicativo que a maior expectativa de vida e o estilo de vida decorrente da urbanização possuem influência no aumento da mortalidade por câncer de próstata e mama no Brasil.

Há necessidade do constante monitoramento das metas contidas na “Política Nacional para a Prevenção e Controle do Câncer na Rede de Atenção à Saúde das Pessoas com Doenças Crônicas no âmbito do Sistema Único de Saúde (SUS)”, que envolve a promoção conscientização da população para adoção de hábitos saudáveis com o objetivo de reduzir os fatores de risco do câncer. Além disso, é necessário que haja ampliação e integração da rede especializada no atendimento e tratamento oncológico, que atualmente está concentrada nos grandes centros urbanos. A identificação precoce da doença é fundamental para a redução das taxas de morbidade e mortalidade.

Observa-se que o sistema de saúde público brasileiro possui dificuldade no refreamento destas doenças, por isso se faz necessário um gerenciamento com maior eficiência destas ações. Também se percebe que o trabalho em conjunto dos órgãos públicos com apoio da sociedade civil, como realizado no Hospital de Amor de Barretos, é uma medida que pode ser mais explorada em outros hospitais públicos do país, vistos os resultados positivos obtidos no combate ao câncer.

Os resultados encontrados nesta pesquisa merecem investigação epidemiológica para subsidiar ações no combate à mortalidade por câncer de próstata e mama em mulheres no Brasil. Como limitação da pesquisa, destaca-se a baixa disponibilidade de dados por microrregião geográfica que limitou a investigação de outras variáveis que pudessem estar correlacionadas com a doença e também não possibilitou a análise dos dados em painel. Destaca-se também a utilização de dados de fontes secundárias. Sugere-se, para trabalhos futuros a previsão espaço-temporal dos casos de câncer e a investigação de outros fatores que influenciem na incidência e mortalidade da doença.

REFERÊNCIAS

ALMEIDA, E. **Econometria Espacial Aplicada**. 1. ed. Campinas, SP: Alínea, 2012

AMIN, R. W.; FRITSCH, B. A.; RETZLOFF, J. E. Spatial Clusters of Breast Cancer Mortality and Incidence in the Contiguous USA: 2000–2014. **Journal of General Internal Medicine**, v. 34, n. 3, p. 412–419, 16 mar. 2019.

ANSELIN, L. **Spatial Econometrics: Methods and Models**. Dordrecht: Springer Netherlands, 1988. v. 4.

- ANSELIN, L. Local indicators of spatial organization - LISA. **Research**, v. 27, n. 2, p. 1–25, 1995.
- ANSELIN, L. **The Moran scatterplot as an ESDA tool to assess local instability in spatial association**. Spatial analytical perspectives on GIS. London: Taylor and Francis, 1996. Disponível em: <<https://www.scienceopen.com/document?vid=b91475ba-1107-46f5-81ff-898e13854593>>. Acesso em: 6 novembro 2019.
- ANSELIN, L. Interactive techniques and exploratory spatial data analysis. **Radiatsionnaia biologiia, radioecologiia / Rossiiskaia akademiia nauk**, v. 37, n. 3, p. 291–296, 1997.
- ANSELIN, L. Spatial Econometrics. In: **A Companion to Theoretical Econometrics**. Malden: Blackwell Publishing Ltd, 2003a. p. 310–330.
- ANSELIN, L. Spatial Externalities, Spatial Multipliers, And Spatial Econometrics. **International Regional Science Review**, v. 26, n. 2, p. 153–166, 26 abr. 2003b.
- ANSELIN, L. **Exploring Spatial Data with GeoDa TM : A Workbook Center for Spatially Integrated Social Science**. Urbana–Champaign: University of Illinois, 2005.
- ANSELIN, L. Thirty years of spatial econometrics. **Papers in Regional Science**, v. 89, n. 1, p. 3–25, 25 mar. 2010.
- ANSELIN, L.; SYABRI, I.; KHO, Y. GeoDa: An Introduction to Spatial Data Analysis. **Geographical Analysis**, v. 38, n. 1, p. 5–22, 1 jan. 2006.
- ANSELIN, L.; SYABRI, I.; SMIRNOV, O. Visualizing multivariate spatial correlation with dynamically linked windows. In: CSISS Workshop on New Tools for Spatial Data Analysis. **Anais...** Santa Barbara, CA: 2002. Disponível em: <<http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.118.7163>>. Acesso em: 5 nov. 2019.
- ARBIA, G.; FINGLETON, B. New spatial econometric techniques and applications in regional science. **Papers in Regional Science**, v. 87, n. 3, p. 311–317, 1 ago. 2008.
- ASSOCIAÇÃO BRASILEIRA DE INSTITUIÇÕES FILANTRÓPICAS DE COMBATE AO CÂNCER – ABIFICC. **Relatório Administrativo 2017**. Disponível em: <https://abificc.org.br/wp-content/uploads/2018/05/relatorio_administrativo_2017_abificc.pdf>. Acesso em: 10 jul. 2020.
- BANNA, S. C.; GONDINHO, B. V. C. Assistência em oncologia no Sistema Único de Saúde - SUS. **Journal of Management and Primary Health Care**, v.11, n. (Sup). 2019.
- BARBOSA, I. R. et al. Desigualdades Socioespaciais Na Distribuição Da Mortalidade Por Câncer No Brasil. **Hygeia**, v. 12, n. 23, p. 122–132, 2016.
- BARRY, J.; BREEN, N. The importance of place of residence in predicting late-stage diagnosis of breast or cervical cancer. **Health and Place**, v. 11, n. 1, p. 15–29, 2005.
- BAUMONT, C. Spatial effects in housing price models. Do housing prices capitalize urban

- development policies in the agglomeration of Dijon (1999)? **Science**, n. 1999, p. 1–26, 2004.
- BEENSTOCK, M.; FELDMAN, D.; FELSENSTEIN, D. Testing for Unit Roots and Cointegration in Spatial Cross-Section Data. **Spatial Economic Analysis**, v. 7, n. 2, p. 203–222, 2012.
- BERNARDES, N. B. et al. Câncer de Mama X Diagnóstico / Breast Cancer X Diagnosis. **ID on line Revista de Psicologia**, v. 13, n. 44, p. 877–885, 27 fev. 2019.
- BORTOLIN, R. Região Sul tem maior incidência de câncer no país, segundo Inca. **Gazeta do Povo**, Curitiba, 16 jan. 2012. Disponível em: <<https://www.gazetadopovo.com.br/saude/regiao-sul-tem-maior-incidencia-de-cancer-no-pais-segundo-inca-7zu61243od8sxig1pfyzbxc9a/>>. Acesso em: 27 jun. 2020.
- BRASIL. **Plano de ações estratégicas para o enfrentamento das doenças crônicas não transmissíveis (DCNT) no Brasil, 2011-2022**. Brasília, Brasil, 2011. Disponível em: <http://bvsms.saude.gov.br/bvs/saudelegis/gm/2013/prt0874_16_05_2013.html>. Acesso em: 9 jul. 2020
- BRASIL. Portaria nº 874, de 16 de maio de 2013. Institui a Política Nacional para a Prevenção e Controle do Câncer na Rede de Atenção à Saúde das Pessoas com Doenças Crônicas no âmbito do Sistema Único de Saúde (SUS). **Diário da União**, 2013.
- BRAY, F. et al. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. **CA: A Cancer Journal for Clinicians**, v. 68, n. 6, p. 394–424, nov. 2018.
- BREUSCH, T. S.; PAGAN, A. R. A Simple Test for Heteroscedasticity and Random Coefficient Variation. **Econometrica**, v. 47, n. 5, p. 1287, set. 1979.
- CAMARA, G. et al. Análise Espacial e Geoprocessamento. In: **Análise espacial de dados geográficos**. Planaltina: Empraba Cerrados, p. 209, 2004.
- CARRIJO, T. B.; DA SILVA, A. R. Modified Moran's I for Small Samples. **Geographical Analysis**, v. 49, n. 4, p. 451–467, 2017.
- CARVALHO, A. X. Y.; ALBUQUERQUE, P. H. M. Tópicos em econometria espacial para dados cross-section. **Economia regional e urbana: teorias e métodos com ênfase no Brasil**, p. 333–364, 2011.
- CLIFF, A. D.; ORD, J. K. **Spatial autocorrelation**. Londres: Pion, 1973.
- CORRADO, L.; FINGLETON, B. Where is the economics in spatial econometrics? **Journal of Regional Science**, v. 52, n. 2, p. 210–239, 1 maio 2012.
- CROISSANT, Y.; MILLO, G. Panel data econometrics in R: The plm package. **Journal of Statistical Software**, v. 27, n. 2, p. 1–43, 29 jul. 2008.
- DEEPA; KANT, S. Impact of western lifestyle on cancer progression. **International Journal of Healthcare Sciences**, v. 4, n. 1, p. 305–310, 2016.

DOMINGUES, J. **Técnicas de processos espaciais e espaço-temporais com aplicações em dados de dengue**. Universidade Estadual Paulista Julio de Mesquita Filho, 2017.

DRISCOLL, J.; KRAAY, A. Spatial correlation in panel data. **Policy Research Working Paper**, p. 28, 1995.

ELHORSTA, J. P.; VEGAA, S. H. **The SLX model : Extensions and the sensitivity of spatial spillovers to W**, 2017. Disponível em: <<https://www.semanticscholar.org/paper/The-SLX-model-%3A-Extensions-and-the-sensitivity-of-W-Elhorsta-Vegaa/3939c2b0ad9873bdcb14560218c3168af6e97296>>. Acesso em: 14 nov. 2019

EPPERSON, B. K.; LI, T. Measurement of genetic structure within populations using Moran's spatial autocorrelation statistics (spatial patterns/population structure /dispersal/gene flow). **Proceedings of the National Academy of Sciences of the United States of America**, v. 93, n. 19, p. 10528–10532, 1996.

FINGLETON, B. A generalized method of moments estimator for a spatial model with moving average errors, with application to real estate prices. **Empirical Economics**, v. 34, n. 1, p. 35–57, 2008.

FISCHER, M. M.; WANG, J. **Spatial data analysis : models, methods and techniques**. Nova York: Springer, 2011.

FLORAX, R. J. G. .; FOLMER, H.; REY, S. J. Specification searches in spatial econometrics: the relevance of Hendry's methodology. **Regional Science and Urban Economics**, v. 33, n. 5, p. 557–579, 1 set. 2003.

FLORAX, R. J. G. M.; GRAAFF, T. DE. **The Performance of Diagnostic Tests for Spatial Dependence in Linear Regression Models: A Meta-Analysis of Simulation Studies**. In: Berlin, Heidelberg: Springer, 2004. p. 29–65.

FORTIN, M. J.; DRAPEAU, P.; LEGENDRE, P. Spatial autocorrelation and sampling design in plant ecology. **Vegetatio**, v. 83, n. 1–2, p. 209–222, 1989.

FRANCESCHI, S.; WILD, C. P. Meeting the global demands of epidemiologic transition - The indispensable role of cancer prevention. **Molecular Oncology**, v. 7, n. 1, p. 1–13, fev. 2013.

G1. Hospital do Câncer de Barretos faz 4.100 atendimentos todos os dias. **Jornal Hoje**, Rio de Janeiro, 17 jun. 2016. Disponível em: <<http://g1.globo.com/jornal-hoje/noticia/2016/06/hospital-do-cancer-de-barretos-faz-4100-atendimentos-todos-os-dias.html>>. Acesso em: 28 jun. 2020.

GARÓFOLO, A. et al. Diet and cancer: An epidemiological view. **Revista de Nutrição**, v. 17, n. 4, p. 491–505, 2004.

GIL, A. C. **Como elaborar projetos de pesquisa**. São Paulo: Atlas, 2002.

GÓES, J. A. P. et al. Trend and spatial analysis of prostate cancer mortality in the state of

Sergipe, Brazil. **Geospatial Health**, v. 13, n. 2, p. 353–358, 2018.

GOLGHER, A. B. **Introdução à Econometria Espacial**. Jundiaí: Paco Editorial, 2015.

GOSS, P. E. et al. Planning cancer control in Latin America and the Caribbean. **The Lancet Oncology**, v. 14, n. 5, p. 391–436, abr. 2013.

GUERRA, M. R. et al. Inequalities in the burden of female breast cancer in Brazil, 1990-2017. **Population Health Metrics**, v. 18, n. S1, p. 8, 30 set. 2020.

HAINING, R. P. The Moving Average Model for Spatial Interaction. **Transactions of the Institute of British Geographers**, v. 3, n. 2, p. 202, 1978.

HARRIS, R.; MOFFAT, J.; KRAVTSOVA, V. **Spatial Economic Analysis**, v. 6, n. 3, p. 249–270, set. 2011.

HARTT, V. [5º FÓRUM] **Acesso ao tratamento no SUS – Barreiras e Defasagens - Instituto Oncoguia**. Disponível em: <<http://www.oncoguia.org.br/conteudo/5-forum-acesso-ao-tratamento-no-sus-barreiras-e-defasagens/7510/8/>>. Acesso em: 9 jul. 2020.

HAUSMAN, J. A. Specification tests in econometrics. **Applied Econometrics**, v. 38, n. 2, p. 112–134, 1978.

HERRMANN, C. et al. 40 years of progress in female cancer death risk: A Bayesian spatio-temporal mapping analysis in Switzerland. **BMC Cancer**, v. 15, n. 1, 9 out. 2015.

HOSPITAL DE AMOR. **Arquivos solidariedade | Hospital de Amor**. Disponível em: <<https://hospitaldeamor.com.br/site/tag/solidariedade/>>. Acesso em: 10 jul. 2020.

INSTITUTO NACIONAL DE CÂNCER - INCA. **Câncer de Próstata | INCA - Instituto Nacional de Câncer**. Disponível em: <<https://www.inca.gov.br/tipos-de-cancer/cancer-de-prostata>>. Acesso em: 21 set. 2019a.

INSTITUTO NACIONAL DE CÂNCER - INCA. **A situação do câncer de mama no Brasil : Síntese de dados dos sistemas**. Rio de Janeiro - RJ: INCA, 2019b.

INSTITUTO NACIONAL DE CÂNCER - INCA. **Câncer de mama | INCA - Instituto Nacional de Câncer**. Disponível em: <<https://www.inca.gov.br/tipos-de-cancer/cancer-de-mama>>. Acesso em: 21 set. 2019c.

INSTITUTO NACIONAL DE CÂNCER - INCA. **Tipos de Câncer | INCA - Instituto Nacional de Câncer**. Disponível em: <<https://www.inca.gov.br/tipos-de-cancer>>. Acesso em: 27 jun. 2020.

JOHNSTON, J.; DINARDO, J. **Econometric methods**. Madison: McGraw Hil, 1997.

KELEJIAN, H. H.; PRUCHA, I. R. A Generalized Spatial Two-Stage Least Squares Procedure for Estimating a Spatial Autoregressive Model with Autoregressive Disturbances. **Journal of Real Estate Finance and Economics**, v. 17, n. 1, p. 99–121, 1998.

KLIER, T.; MCMILLEN, D. P. Clustering of Auto Supplier Plants in the United States: Generalized Method of Moments Spatial Logit for Large Samples, **Journal of Business & Economic Statistics**, Taylor & Francis, Ltd. American Statistical Association, 2008. Disponível em: <<https://www.jstor.org/stable/27639005>>. Acesso em: 9 nov. 2019

LESAGE, J. P.; P., J. An Introduction to Spatial Econometrics. **Revue d'économie industrielle**, n. 123, p. 19–44, 15 set. 2008.

LESAGE, J. P.; PACE, R. K. **Introduction to spatial econometrics**. Ohio: CRC Press, 2009.

MADHU, B. et al. Spatio-Temporal Pattern of Breast Cancer-Case Study of Southern Karnataka, India. **Journal of Clinical and Diagnostic Research**, v. 10, n. 4, p. 20–24, 2016.

MORAN, P. A. P. The Interpretation of Statistical Maps. **Journal of the Royal Statistical Society: Series B (Methodological)**, v. 10, n. 2, p. 243–251, 1 jul. 1948.

MORETTIN, P. A.; TOLOI, C. M. DE C. **Análise de séries temporais**. São Paulo: Edgard Blucher, 2006.

MUELLER, G. S. et al. Manuscript title: Geospatial analysis of Cancer risk and residential proximity to coal mines in Illinois. **Ecotoxicology and Environmental Safety**, v. 120, p. 155–162, out. 2015.

NÚÑEZ, O. et al. Arsenic and chromium topsoil levels and cancer mortality in Spain. **Environmental Science and Pollution Research**, v. 23, n. 17, p. 17664–17675, 30 set. 2016.

OLIVEIRA, E. X. G. DE et al. Acesso à assistência oncológica: mapeamento dos fluxos origem-destino das internações e dos atendimentos ambulatoriais. O caso do câncer de mama. **Caderno de Saúde Pública**, v. 27, n. 2, p. 317–326, 2011a.

OLIVEIRA, E. X. G. DE et al. Acesso à assistência oncológica: Mapeamento dos fluxos origem-destino das internações e dos atendimentos ambulatoriais. o caso do cancer de mama. **Cadernos de Saude Publica**, v. 27, n. 2, p. 317–326, fev. 2011b.

ORD, K. Estimation Methods for Models of Spatial Interaction. **Journal of the American Statistical Association**, v. 70, n. 349, p. 120–126, mar. 1975.

ORGANIZAÇÃO MUNDIAL DA SAÚDE. **Casos de câncer devem aumentar 70% até 2038, calcula OMS | ONU Brasil**. Disponível em: <<https://nacoesunidas.org/casos-de-cancer-devem-aumentar-70-ate-2038-calcula-oms/>>. Acesso em: 5 out. 2019.

PIRAS, G.; LOZANO-GRACIA, N. Spatial J-test: some Monte Carlo evidence. **Statistics and Computing**, v. 22, n. 1, p. 169–183, 6 jan. 2012.

ROCHA-BRISCHILIARI, S. C. et al. Spatial distribution of breast cancer mortality: Socioeconomic disparities and access to treatment in the state of Parana, Brazil. **PloS one**, v. 13, n. 10, p. e0205253, 31 out. 2018.

SALES, C. A. et al. Qualidade de vida sob a ótica de pessoas em tratamento antineoplásico: uma análise fenomenológica. **Revista Eletrônica de Enfermagem**, 2011.

- SANTOS, M. DE O. **Estimativa 2018: Incidência de Câncer no Brasil**. Disponível em: <<https://rbc.inca.gov.br/revista/index.php/revista/article/view/115/55>>. Acesso em: 24 set. 2019.
- SILVA, J. F. S. DA et al. Correlação entre produção agrícola, variáveis clínicas-demográficas e câncer de próstata: um estudo ecológico. **Ciência & Saúde Coletiva**, v. 20, n. 9, p. 2805–2812, set. 2015.
- STAKHOVYCH, S.; BIJMOLT, T. H. A. Specification of spatial models: A simulation study on weights matrices. **Papers in Regional Science**, v. 88, n. 2, p. 389–408, 1 jun. 2009.
- TOBLER, W. R. Cellular Geography. In: **Philosophy in Geography**. Dordrecht: Springer Netherlands, p. 379–386, 1979.
- TORRE, L. A. et al. Global cancer statistics, 2012. **CA: A Cancer Journal for Clinicians**, v. 65, n. 2, p. 87–108, mar. 2015.
- TORRE, L. A. et al. Global cancer in women: Burden and trends. **Cancer Epidemiology Biomarkers and Prevention**, v. 26, n. 4, p. 444–457, 2017.
- TORRES-ROMAN, J. S. et al. Prostate cancer mortality rates in Peru and its geographical regions. **BJU International**, v. 123, n. 4, p. 595–601, 2019.
- TYSZLER, M. Econometria espacial: discutindo medidas para a matriz de ponderação espacial. 19 jan. 2006.
- VERGARA, S. C. **Projetos e relatórios de pesquisa em administração**. 16. ed. São Paulo: Atlas S.A, 2016.
- VIEIRA, R. DE S. A abordagem clássica de Econometria Espacial. In: **Crescimento econômico no estado de São Paulo: uma análise espacial**. São Paulo: Editora UNESP, 2009. p. 103.
- VINNAKOTA, S.; LAM, N. Socioeconomic inequality of cancer mortality in the United States: a spatial data mining approach. **International Journal of Health Geographics**, v. 5, n. 1, p. 9, 15 fev. 2006.
- WHITTLE, P. On Stationary Processes in the Plane. **Biometrika**, v. 41, n. 3/4, p. 434-449, dez. 1954.

5 CONSIDERAÇÕES FINAIS

Esta pesquisa objetivou determinar a distribuição espacial da mortalidade por câncer de próstata e mama nas microrregiões brasileiras e verificar o seu inter-relacionamento com fatores sociodemográficos, econômicos, ambientais e de infraestrutura. Para isso, foi verificado se as microrregiões brasileiras apresentaram autocorrelação espacial pela estatística I de Moran, determinaram-se os padrões espaciais da mortalidade por câncer de próstata e mama no Brasil, escolheram-se as variáveis que representassem os fatores sociodemográficos, econômicos, ambientais e de infraestrutura de acordo com literatura relevante e disponibilidade de dados por microrregião e foi ajustado um modelo espacial para representar a autocorrelação espacial das taxas de mortalidade por câncer de próstata e mama nas microrregiões brasileiras, baseado nas estatísticas de ajustes.

Os resultados mostraram que há uma maior concentração da mortalidade por CaP e CaM nas regiões centro-sul e nordeste do Brasil e há autocorrelação espacial global positiva e significativa para os dois casos. A região norte do país concentra as microrregiões que possuem padrão espacial do tipo Baixo-Baixo e a região sul do país, o estado de São Paulo e o estado do Rio de Janeiro concentram as microrregiões que possuem padrões espaciais do tipo Alto-Alto.

O modelo ajustado para a taxa de mortalidade por CaP foi o modelo de Durbin Espacial. As variáveis significativas foram taxa de envelhecimento e taxa de urbanização, apresentando correlação positiva entre a taxa de mortalidade por câncer de próstata e as variáveis taxa de urbanização e taxa de envelhecimento. Este inter-relacionamento indicou que uma maior taxa de mortalidade por CaP em uma dada microrregião está relacionada com maiores taxas de urbanização e envelhecimento, na mesma microrregião.

Em relação a taxa de mortalidade por câncer de mama, o modelo ajustado foi um SAR. As variáveis significativas ao modelo foram número de médicos especialistas em oncologia, número de mamógrafos, unidades de referência, taxa de urbanização e taxa de envelhecimento. Há correlação positiva entre número de médicos especialistas, unidades de referência, taxa de urbanização, taxa de envelhecimento e a taxa de mortalidade por câncer de mama. Em relação ao número de mamógrafos e a taxa de mortalidade por CaM, a correlação é negativa. Este relacionamento significa que uma maior mortalidade por câncer de mama em uma dada microrregião está relacionada a um maior número de médicos especialistas, um maior número de unidades de referência para tratamento da doença, maior taxa de envelhecimento populacional, maior taxa de urbanização e menor número de mamógrafos.

Houve indicativo que a maior expectativa de vida e o estilo de vida decorrente da urbanização possuem influência no aumento da mortalidade por câncer de próstata e mama no Brasil.

Há necessidade do constante monitoramento das metas contidas na “Política Nacional para a Prevenção e Controle do Câncer na Rede de Atenção à Saúde das Pessoas com Doenças Crônicas no âmbito do Sistema Único de Saúde (SUS)”, que envolve a promoção e conscientização da população para adoção de hábitos saudáveis com o objetivo de reduzir os fatores de risco do câncer. Além disso, é necessário que haja ampliação e integração da rede especializada no atendimento e tratamento oncológico, que atualmente está concentrada nos grandes centros urbanos. A identificação precoce da doença é fundamental para a redução das taxas de morbidade e mortalidade.

Observa-se que o sistema de saúde público brasileiro possui dificuldade no refreamento destas doenças, por isso se faz necessário um gerenciamento com maior eficiência destas ações. Também, percebe-se que o trabalho em conjunto dos órgãos públicos com apoio da sociedade civil, como o realizado no Hospital de Amor de Barretos, é uma medida que pode ser mais explorada em outros hospitais públicos do país, vistos os resultados positivos obtidos no combate ao câncer.

Diante do exposto, considera-se que os objetivos gerais e específicos da pesquisa foram atingidos e a análise espacial da mortalidade por câncer possibilitou encontrar correlações estatísticas que merecem investigação epidemiológica para subsidiar ações no combate da mortalidade por câncer de próstata e mama em mulheres no Brasil. Como limitação da pesquisa, destaca-se a baixa disponibilidade de dados por microrregião geográfica que limitou a investigação de outras variáveis que pudessem estar correlacionadas com a doença e também não possibilitou a análise dos dados em painel. Destaca-se também a utilização de dados de fontes secundárias. Sugere-se, para trabalhos futuros a previsão espaço-temporal dos casos de câncer e a investigação de outros fatores que influenciem na incidência e mortalidade da doença.

REFERÊNCIAS

- ALMEIDA, E. **Econometria Espacial Aplicada**. 1. ed. Campinas, SP: Alínea, 2012
- AMIN, R. W.; FRITSCH, B. A.; RETZLOFF, J. E. Spatial Clusters of Breast Cancer Mortality and Incidence in the Contiguous USA: 2000–2014. **Journal of General Internal Medicine**, v. 34, n. 3, p. 412–419, 16 mar. 2019.
- ANSELIN, L. **Spatial Econometrics: Methods and Models**. Dordrecht: Springer Netherlands, 1988. v. 4.
- ANSELIN, L. Local indicators of spatial organization - LISA. **Research**, v. 27, n. 2, p. 1–25, 1995.
- ANSELIN, L. **The Moran scatterplot as an ESDA tool to assess local instability in spatial association**. Spatial analytical perspectives on GIS. London: Taylor and Francis, 1996. Disponível em: <<https://www.scienceopen.com/document?vid=b91475ba-1107-46f5-81ff-898e13854593>>. Acesso em: 6 novembro 2019.
- ANSELIN, L. Interactive techniques and exploratory spatial data analysis. **Radiatsionnaia biologii, radioecologia / Rossiiskaia akademiia nauk**, v. 37, n. 3, p. 291–296, 1997.
- ANSELIN, L. Spatial Econometrics. In: **A Companion to Theoretical Econometrics**. Malden: Blackwell Publishing Ltd, 2003a. p. 310–330.
- ANSELIN, L. Spatial Externalities, Spatial Multipliers, And Spatial Econometrics. **International Regional Science Review**, v. 26, n. 2, p. 153–166, 26 abr. 2003b.
- ANSELIN, L. **Exploring Spatial Data with GeoDa TM : A Workbook Center for Spatially Integrated Social Science**. Urbana–Champaign: University of Illinois, 2005.
- ANSELIN, L. Thirty years of spatial econometrics. **Papers in Regional Science**, v. 89, n. 1, p. 3–25, 25 mar. 2010.
- ANSELIN, L.; SYABRI, I.; KHO, Y. GeoDa: An Introduction to Spatial Data Analysis. **Geographical Analysis**, v. 38, n. 1, p. 5–22, 1 jan. 2006.
- ANSELIN, L.; SYABRI, I.; SMIRNOV, O. Visualizing multivariate spatial correlation with dynamically linked windows. In: CSISS Workshop on New Tools for Spatial Data Analysis. **Anais...**Santa Barbara, CA: 2002. Disponível em: <<http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.118.7163>>. Acesso em: 5 nov. 2019.
- ARBIA, G.; FINGLETON, B. New spatial econometric techniques and applications in regional science. **Papers in Regional Science**, v. 87, n. 3, p. 311–317, 1 ago. 2008.
- ASSOCIAÇÃO BRASILEIRA DE INSTITUIÇÕES FILANTRÓPICAS DE COMBATE AO CÂNCER – ABIFICC. **Relatório Administrativo 2017**. Disponível em: <https://abificc.org.br/wp-content/uploads/2018/05/relatorio_administrativo_2017_abificc.pdf>. Acesso em: 10 jul. 2020.

BANNA, S. C.; GONDINHO, B. V. C. Assistência em oncologia no Sistema Único de Saúde - SUS. **Journal of Management and Primary Health Care**, v.11, n. (Sup). 2019.

BARBOSA, I. R. et al. Desigualdades Socioespaciais Na Distribuição Da Mortalidade Por Câncer No Brasil. **Hygeia**, v. 12, n. 23, p. 122–132, 2016.

BARRY, J.; BREEN, N. The importance of place of residence in predicting late-stage diagnosis of breast or cervical cancer. **Health and Place**, v. 11, n. 1, p. 15–29, 2005.

BAUMONT, C. Spatial effects in housing price models. Do housing prices capitalize urban development policies in the agglomeration of Dijon (1999)? **Science**, n. 1999, p. 1–26, 2004.

BEENSTOCK, M.; FELDMAN, D.; FELSENSTEIN, D. Testing for Unit Roots and Cointegration in Spatial Cross-Section Data. **Spatial Economic Analysis**, v. 7, n. 2, p. 203–222, 2012.

BERNARDES, N. B. et al. Câncer de Mama X Diagnóstico / Breast Cancer X Diagnosis. **ID on line Revista de Psicologia**, v. 13, n. 44, p. 877–885, 27 fev. 2019.

BORTOLIN, R. Região Sul tem maior incidência de câncer no país, segundo Inca. **Gazeta do Povo**, Curitiba, 16 jan. 2012. Disponível em: <<https://www.gazetadopovo.com.br/saude/regiao-sul-tem-maior-incidencia-de-cancer-no-pais-segundo-inca-7zu61243od8sxig1pfyzbxc9a/>>. Acesso em: 27 jun. 2020.

BRASIL. **Plano de ações estratégicas para o enfrentamento das doenças crônicas não transmissíveis (DCNT) no Brasil, 2011-2022**. Brasília, Brasil, 2011. Disponível em: <http://bvsmis.saude.gov.br/bvs/saudelegis/gm/2013/prt0874_16_05_2013.html>. Acesso em: 9 jul. 2020

BRASIL. Portaria nº 874, de 16 de maio de 2013. Institui a Política Nacional para a Prevenção e Controle do Câncer na Rede de Atenção à Saúde das Pessoas com Doenças Crônicas no âmbito do Sistema Único de Saúde (SUS). **Diário da União**, 2013.

BRAY, F. et al. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. **CA: A Cancer Journal for Clinicians**, v. 68, n. 6, p. 394–424, nov. 2018.

BREUSCH, T. S.; PAGAN, A. R. A Simple Test for Heteroscedasticity and Random Coefficient Variation. **Econometrica**, v. 47, n. 5, p. 1287, set. 1979.

CAMARA, G. et al. Análise Espacial e Geoprocessamento. In: **Análise espacial de dados geográficos**. Planaltina: Empraba Cerrados, p. 209, 2004.

CARRIJO, T. B.; DA SILVA, A. R. Modified Moran's I for Small Samples. **Geographical Analysis**, v. 49, n. 4, p. 451–467, 2017.

CARVALHO, A. X. Y.; ALBUQUERQUE, P. H. M. Tópicos em econometria espacial para dados cross-section. **Economia regional e urbana: teorias e métodos com ênfase no Brasil**, p. 333–364, 2011.

CLIFF, A. D.; ORD, J. K. **Spatial autocorrelation**. Londres: Pion, 1973.

CORRADO, L.; FINGLETON, B. Where is the economics in spatial econometrics? **Journal of Regional Science**, v. 52, n. 2, p. 210–239, 1 maio 2012.

CROISSANT, Y.; MILLO, G. Panel data econometrics in R: The plm package. **Journal of Statistical Software**, v. 27, n. 2, p. 1–43, 29 jul. 2008.

DEEPA; KANT, S. Impact of western lifestyle on cancer progression. **International Journal of Healthcare Sciences**, v. 4, n. 1, p. 305–310, 2016.

DOMINGUES, J. **Técnicas de processos espaciais e espaço-temporais com aplicações em dados de dengue**. Universidade Estadual Paulista Julio de Mesquita Filho, 2017.

DRISCOLL, J.; KRAAY, A. Spatial correlation in panel data. **Policy Research Working Paper**, p. 28, 1995.

ELHORSTA, J. P.; VEGAA, S. H. **The SLX model : Extensions and the sensitivity of spatial spillovers to W**, 2017. Disponível em: <<https://www.semanticscholar.org/paper/The-SLX-model-%3A-Extensions-and-the-sensitivity-of-W-Elhorsta-Vegaa/3939c2b0ad9873bdcb14560218c3168af6e97296>>. Acesso em: 14 nov. 2019

EPPERSON, B. K.; LI, T. Measurement of genetic structure within populations using Moran's spatial autocorrelation statistics (spatial patterns/population structure /dispersal/gene flow). **Proceedings of the National Academy of Sciences of the United States of America**, v. 93, n. 19, p. 10528–10532, 1996.

FINGLETON, B. A generalized method of moments estimator for a spatial model with moving average errors, with application to real estate prices. **Empirical Economics**, v. 34, n. 1, p. 35–57, 2008.

FISCHER, M. M.; WANG, J. **Spatial data analysis : models, methods and techniques**. Nova York: Springer, 2011.

FLORAX, R. J. G. .; FOLMER, H.; REY, S. J. Specification searches in spatial econometrics: the relevance of Hendry's methodology. **Regional Science and Urban Economics**, v. 33, n. 5, p. 557–579, 1 set. 2003.

FLORAX, R. J. G. M.; GRAAFF, T. DE. **The Performance of Diagnostic Tests for Spatial Dependence in Linear Regression Models: A Meta-Analysis of Simulation Studies**. In: Berlin, Heidelberg: Springer, 2004. p. 29–65.

FORTIN, M. J.; DRAPEAU, P.; LEGENDRE, P. Spatial autocorrelation and sampling design in plant ecology. **Vegetatio**, v. 83, n. 1–2, p. 209–222, 1989.

FRANCESCHI, S.; WILD, C. P. Meeting the global demands of epidemiologic transition - The indispensable role of cancer prevention. **Molecular Oncology**, v. 7, n. 1, p. 1–13, fev. 2013.

- G1. Hospital do Câncer de Barretos faz 4.100 atendimentos todos os dias. **Jornal Hoje**, Rio de Janeiro, 17 jun. 2016. Disponível em: <<http://g1.globo.com/jornal-hoje/noticia/2016/06/hospital-do-cancer-de-barretos-faz-4100-atendimentos-todos-os-dias.html>>. Acesso em: 28 jun. 2020.
- GARÓFOLO, A. et al. Diet and cancer: An epidemiological view. **Revista de Nutrição**, v. 17, n. 4, p. 491–505, 2004.
- GIL, A. C. **Como elaborar projetos de pesquisa**. São Paulo: Atlas, 2002.
- GÓES, J. A. P. et al. Trend and spatial analysis of prostate cancer mortality in the state of Sergipe, Brazil. **Geospatial Health**, v. 13, n. 2, p. 353–358, 2018.
- GOLGHER, A. B. **Introdução à Econometria Espacial**. Jundiaí: Paco Editorial, 2015.
- GOSS, P. E. et al. Planning cancer control in Latin America and the Caribbean. **The Lancet Oncology**, v. 14, n. 5, p. 391–436, abr. 2013.
- GUERRA, M. R. et al. Inequalities in the burden of female breast cancer in Brazil, 1990–2017. **Population Health Metrics**, v. 18, n. S1, p. 8, 30 set. 2020.
- HAINING, R. P. The Moving Average Model for Spatial Interaction. **Transactions of the Institute of British Geographers**, v. 3, n. 2, p. 202, 1978.
- HARRIS, R.; MOFFAT, J.; KRAVTSOVA, V. **Spatial Economic Analysis**, v. 6, n. 3, p. 249–270, set. 2011.
- HARTT, V. [5º FÓRUM] **Acesso ao tratamento no SUS – Barreiras e Defasagens - Instituto Oncoguia**. Disponível em: <<http://www.oncoguia.org.br/conteudo/5-forum-acesso-ao-tratamento-no-sus-barreiras-e-defasagens/7510/8/>>. Acesso em: 9 jul. 2020.
- HAUSMAN, J. A. Specification tests in econometrics. **Applied Econometrics**, v. 38, n. 2, p. 112–134, 1978.
- HERRMANN, C. et al. 40 years of progress in female cancer death risk: A Bayesian spatio-temporal mapping analysis in Switzerland. **BMC Cancer**, v. 15, n. 1, 9 out. 2015.
- HOSPITAL DE AMOR. **Arquivos solidariedade | Hospital de Amor**. Disponível em: <<https://hospitaldeamor.com.br/site/tag/solidariedade/>>. Acesso em: 10 jul. 2020.
- INSTITUTO NACIONAL DE CÂNCER - INCA. **Câncer de Próstata | INCA - Instituto Nacional de Câncer**. Disponível em: <<https://www.inca.gov.br/tipos-de-cancer/cancer-de-prostata>>. Acesso em: 21 set. 2019a.
- INSTITUTO NACIONAL DE CÂNCER - INCA. **A situação do câncer de mama no Brasil : Síntese de dados dos sistemas**. Rio de Janeiro - RJ: INCA, 2019b.
- INSTITUTO NACIONAL DE CÂNCER - INCA. **Câncer de mama | INCA - Instituto Nacional de Câncer**. Disponível em: <<https://www.inca.gov.br/tipos-de-cancer/cancer-de-mama>>. Acesso em: 21 set. 2019c.

INSTITUTO NACIONAL DE CÂNCER - INCA. **Tipos de Câncer | INCA - Instituto Nacional de Câncer**. Disponível em: <<https://www.inca.gov.br/tipos-de-cancer>>. Acesso em: 27 jun. 2020.

JOHNSTON, J.; DINARDO, J. **Econometric methods**. Madison: McGraw Hil, 1997.

KELEJIAN, H. H.; PRUCHA, I. R. A Generalized Spatial Two-Stage Least Squares Procedure for Estimating a Spatial Autoregressive Model with Autoregressive Disturbances. **Journal of Real Estate Finance and Economics**, v. 17, n. 1, p. 99–121, 1998.

KLIER, T.; MCMILLEN, D. P. Clustering of Auto Supplier Plants in the United States: Generalized Method of Moments Spatial Logit for Large Samples, **Journal of Business & Economic Statistics**, Taylor & Francis, Ltd.American Statistical Association, 2008. Disponível em: <<https://www.jstor.org/stable/27639005>>. Acesso em: 9 nov. 2019

LESAGE, J. P.; P., J. An Introduction to Spatial Econometrics. **Revue d'économie industrielle**, n. 123, p. 19–44, 15 set. 2008.

LESAGE, J. P.; PACE, R. K. **Introduction to spatial econometrics**. Ohio: CRC Press, 2009.

MADHU, B. et al. Spatio-Temporal Pattern of Breast Cancer-Case Study of Southern Karnataka, India. **Journal of Clinical and Diagnostic Research**, v. 10, n. 4, p. 20–24, 2016.

MORAN, P. A. P. The Interpretation of Statistical Maps. **Journal of the Royal Statistical Society: Series B (Methodological)**, v. 10, n. 2, p. 243–251, 1 jul. 1948.

MORETTIN, P. A.; TOLOI, C. M. DE C. **Análise de séries temporais**. São Paulo: Edgard Blucher, 2006.

MUELLER, G. S. et al. Manuscript title: Geospatial analysis of Cancer risk and residential proximity to coal mines in Illinois. **Ecotoxicology and Environmental Safety**, v. 120, p. 155–162, out. 2015.

NÚÑEZ, O. et al. Arsenic and chromium topsoil levels and cancer mortality in Spain. **Environmental Science and Pollution Research**, v. 23, n. 17, p. 17664–17675, 30 set. 2016.

OLIVEIRA, E. X. G. DE et al. Acesso à assistência oncológica: mapeamento dos fluxos origem-destino das internações e dos atendimentos ambulatoriais. O caso do câncer de mama. **Caderno de Saúde Pública**, v. 27, n. 2, p. 317–326, 2011a.

OLIVEIRA, E. X. G. DE et al. Acesso à assistência oncológica: Mapeamento dos fluxos origem-destino das internações e dos atendimentos ambulatoriais. o caso do cancer de mama. **Cadernos de Saude Publica**, v. 27, n. 2, p. 317–326, fev. 2011b.

ORD, K. Estimation Methods for Models of Spatial Interaction. **Journal of the American Statistical Association**, v. 70, n. 349, p. 120–126, mar. 1975.

ORGANIZAÇÃO MUNDIAL DA SAÚDE. **Casos de câncer devem aumentar 70% até 2038, calcula OMS | ONU Brasil**. Disponível em: <<https://nacoesunidas.org/casos-de->

cancer-devem-aumentar-70-ate-2038-calcula-oms/>. Acesso em: 5 out. 2019.

PIRAS, G.; LOZANO-GRACIA, N. Spatial J-test: some Monte Carlo evidence. **Statistics and Computing**, v. 22, n. 1, p. 169–183, 6 jan. 2012.

ROCHA-BRISCHILIARI, S. C. et al. Spatial distribution of breast cancer mortality: Socioeconomic disparities and access to treatment in the state of Parana, Brazil. **PloS one**, v. 13, n. 10, p. e0205253, 31 out. 2018.

SALES, C. A. et al. Qualidade de vida sob a ótica de pessoas em tratamento antineoplásico: uma análise fenomenológica. **Revista Eletrônica de Enfermagem**, 2011.

SANTOS, M. DE O. **Estimativa 2018: Incidência de Câncer no Brasil**. Disponível em: <<https://rbc.inca.gov.br/revista/index.php/revista/article/view/115/55>>. Acesso em: 24 set. 2019.

SILVA, J. F. S. DA et al. Correlação entre produção agrícola, variáveis clínicas-demográficas e câncer de próstata: um estudo ecológico. **Ciência & Saúde Coletiva**, v. 20, n. 9, p. 2805–2812, set. 2015.

STAKHOVYCH, S.; BIJMOLT, T. H. A. Specification of spatial models: A simulation study on weights matrices. **Papers in Regional Science**, v. 88, n. 2, p. 389–408, 1 jun. 2009.

TOBLER, W. R. Cellular Geography. In: **Philosophy in Geography**. Dordrecht: Springer Netherlands, p. 379–386, 1979.

TORRE, L. A. et al. Global cancer statistics, 2012. **CA: A Cancer Journal for Clinicians**, v. 65, n. 2, p. 87–108, mar. 2015.

TORRE, L. A. et al. Global cancer in women: Burden and trends. **Cancer Epidemiology Biomarkers and Prevention**, v. 26, n. 4, p. 444–457, 2017.

TORRES-ROMAN, J. S. et al. Prostate cancer mortality rates in Peru and its geographical regions. **BJU International**, v. 123, n. 4, p. 595–601, 2019.

TYSZLER, M. Econometria espacial: discutindo medidas para a matriz de ponderação espacial. 19 jan. 2006.

VERGARA, S. C. **Projetos e relatórios de pesquisa em administração**. 16. ed. São Paulo: Atlas S.A, 2016.

VIEIRA, R. DE S. A abordagem clássica de Econometria Espacial. In: **Crescimento econômico no estado de São Paulo: uma análise espacial**. São Paulo: Editora UNESP, 2009. p. 103.

VINNAKOTA, S.; LAM, N. Socioeconomic inequality of cancer mortality in the United States: a spatial data mining approach. **International Journal of Health Geographics**, v. 5, n. 1, p. 9, 15 fev. 2006.

WHITTLE, P. On Stationary Processes in the Plane. **Biometrika**, v. 41, n. 3/4, p. 434–449,

dez. 1954.