



**Universidade Nova de Lisboa
Instituto de Higiene e Medicina Tropical**

**Avaliação Espacial da Influência dos Factores
Socioeconómicos na Mortalidade por Acidente Vascular
Cerebral na População de Portugal Continental**

Miguel André Fouto Pinho de Oliveira

**DISSERTAÇÃO PARA A OBTENÇÃO DO GRAU DE DOUTOR EM
SAÚDE INTERNACIONAL, ESPECIALIDADE EM POLITICAS DE
SAÚDE E DESENVOLVIMENTO**

MAIO, 2015



**Universidade Nova de Lisboa
Instituto de Higiene e Medicina Tropical**

**Avaliação Espacial da Influência dos Factores
Socioeconómicos na Mortalidade por Acidente Vascular
Cerebral na População de Portugal Continental**

Autor: Miguel André Fouto Pinho de Oliveira

Orientador: Professor Doutor António Jorge Rodrigues Cabral

Co-orientadores: Professora Doutora Maria do Rosário de Oliveira Martins
Professor Doutor Jorge Morais Mendes
Professor Doutor Pedro da Costa Brito Cabral

Dissertação apresentada para cumprimento dos requisitos necessários à obtenção do grau de Doutor em Saúde Internacional, Especialidade em Políticas de Saúde e Desenvolvimento

Apoio financeiro da Fundação para a Ciência e Tecnologia (SFRH/BD/63597/2009)

Elementos bibliográficos

Artigos publicados

- Oliveira, A., Cabral, A.J.R., Mendes, J.M., Martins, M.R.O., Cabral, P., 2015. Spatiotemporal analysis of the relationship between socioeconomic factors and stroke in the Portuguese mainland population under 65 years old. *Geospatial Health* 10. doi:10.4081/gh.2015.365
- Oliveira, A., Cabral, A.J.R., Martins, M.F.O., Cabral, P., Mendes, J.M., Carreira, M., 2016. Poor housing and stroke mortality in population over 64 years old at the parish level, in mainland Portugal. *GeoJournal* 1–18. doi:10.1007/s10708-016-9712-4

Comunicações

- IV Jornadas Científicas do Instituto de Higiene e Medicina Tropical (IHMT-UNL), 13 de Dezembro de 2013. Comunicação oral subordinada ao tema: Modelos globais e locais para dados de contagem na análise espacial do risco de Acidente Vascular Cerebral e determinantes socioeconómicos.
- II Reunião Nacional de Estudantes de Doutoramento em Saúde Pública, 14 de Dezembro de 2012, Instituto de Saúde Pública da Universidade do Porto. Comunicação oral subordinada ao tema: Dependência e Heterogeneidade Espacial em Padrões de Mortalidade por Acidente Vascular Cerebral.
- III Jornadas Científicas do Instituto de Higiene e Medicina Tropical (IHMT-UNL), 12 de Dezembro de 2012. Comunicação oral subordinada ao tema: Dependência e Heterogeneidade Espacial em Padrões de Mortalidade por Acidente Vascular Cerebral.
- II Jornadas Científicas do Instituto de Higiene e Medicina Tropical (IHMT-UNL), 12 de Dezembro de 2011. Participação com poster subordinado ao tema: Comparação entre modelos de regressão espaciais e não espaciais na avaliação da influência de factores socioeconómicos na mortalidade concelhia por acidente vascular cerebral.

– I Jornadas Científicas do Instituto de Higiene e Medicina Tropical (IHMT-UNL), 13 de Dezembro de 2010. Participação com poster subordinado ao tema: Avaliação espacial da influência dos factores socioeconómicos na morbilidade e mortalidade associadas a doenças mentais e do aparelho circulatório.

Agradecimentos

Gostaria de demonstrar o meu agradecimento a todas as pessoas que acompanharam a evolução deste trabalho e que de alguma forma me apoiaram ao longo deste percurso.

Antes de mais, agradeço especialmente ao meu orientador, Professor Doutor Jorge Cabral e aos meus co-orientadores Professora Doutora Maria do Rosário Martins, Professor Doutor Jorge Mendes e Professor Doutor Pedro Cabral, pela amabilidade, disponibilidade e partilha de conhecimentos, que muito contribuíram para a prossecução, aperfeiçoamento e finalização deste trabalho.

Agradeço também aos meus amigos Hugo Carrão, pessoa inestimável cujo valor humano dificilmente pode ser descrito em simples palavras, António Nunes e Vasco Nunes, pela constante ajuda e conselhos que sempre se disponibilizaram a prestar. Na Direcção-Geral da Saúde, agradeço ao Dr. Francisco George, pela sua amabilidade e acolhimento, bem como o apoio do Dr. Mário Carreira e do Professor Doutor Paulo Nogueira. Nesta instituição, não posso deixar de referir aqui as minhas colegas e amigas Rosa Leitão e Manuela Correia, do Núcleo de Apoio à Informática que, juntamente com o meu amigo Luís Serra, sempre me proporcionaram simpatia e apoio, bem como tantas outras pessoas desta instituição onde sempre me senti em casa.

Finalmente, mas de forma alguma com menos importância, agradeço à Cristina pelo seu constante carinho, apoio e crítica construtiva ao longo de todo este caminho, e aos meus Pais e à minha Irmã, sem os quais esta caminhada não teria sido possível.

A todos, um muito bem-haja.

Resumo

O Acidente Vascular Cerebral (AVC) constitui um problema de Saúde Pública de grande magnitude, tanto em mortalidade como morbidade. Em 2010, representava, globalmente, a segunda maior causa de morte e a terceira causa de anos de vida perdidos ajustados pela incapacidade. A sua evolução temporal também traz preocupações, pelo aumento estatisticamente significativo, entre 1990 e 2010, verificado em indicadores como os anos de vida potencialmente perdidos e total de sobreviventes a um episódio.

Embora ocorra mais frequentemente em adultos de meia-idade ou em idosos, entre 1990 e 2010 foi registado um aumento global e significativo na incidência em pessoas com idades compreendidas entre 20 e 64 anos de idade, para o qual contribuíram substancialmente países de baixo e médio rendimento.

Esta situação, aliada ao facto, apontado por alguns autores, do risco de AVC apresentar padrões diversos de distribuição geográfica entre países mas também entre regiões do mesmo país, levanta interrogações acerca da importância dos determinantes socioeconómicos para a mortalidade e morbidade por esta doença, bem como acerca da variação espacial da associação entre estes determinantes e o risco de AVC.

Assim, foram aqui operacionalizados dois estudos com o objectivo de avaliar a variação espacial da associação entre mortalidade por AVC e determinantes socioeconómicos, tendo como área de estudo o território de Portugal Continental.

Um dos estudos investigou a variação geográfica e temporal da associação entre determinantes pertencentes a várias categorias representativas do estatuto socioeconómico (por exemplo o nível de escolaridade) e mortalidade por AVC em pessoas abaixo dos 65 anos de idade, tendo como unidade estatística o município. Este coorte populacional suscitou interesse pela presente situação de crise económica e seus potenciais impactos na população em idade laboral.

No segundo estudo abordou-se essencialmente o coorte populacional inverso, a população com mais de 64 anos de idade, tendo como unidade estatística a freguesia. Investigou-se a variação espacial da associação entre condições adversas de habitabilidade (por exemplo a incapacidade de aquecer a habitação no Inverno) e a mortalidade por AVC, pois a evidência existente sugere que este coorte populacional possa ser mais afectado por esta categoria de determinantes socioeconómicos.

Os resultados de ambos os estudos permitiram constatar de forma expressiva a presença de variação espacial na associação entre cada grupo de determinantes socioeconómicos e o risco de AVC nos coortes populacionais investigados. Em ambos os casos, foi possível avaliar quais os determinantes mais relevantes para o risco de AVC e analisar a variação da força e sinal (positivo ou negativo) das associações ao longo do território. O nível de escolaridade surgiu no primeiro estudo como factor protector em geral, mas como factor de risco em certas zonas do território. A incapacidade de aquecer a habitação no Inverno revelou-se no, segundo estudo, como um potencial factor de risco, particularmente na área noroeste do território continental português.

Os resultados obtidos sugerem que as metodologias testadas poderão contribuir para a vigilância epidemiológica do AVC, por exemplo na selecção de áreas de intervenção prioritária para acções de prevenção ou mitigação desta doença.

Palavras-chave

Acidente Vascular Cerebral, Determinantes Socioeconómicos, Análise Espacial

Abstract

Stroke is a major health problem, both in terms of morbidity and mortality. In 2010, this disease was the second major global cause of death, third in terms of Disability Adjusted Life Years (DALY's). Its temporal evolution raises concerns amongst public health authorities, since a statistically significant global increase was noticed between 1990 and 2010 in health indicators such as DALY's and total number of survivors to a stroke episode.

Although it frequently occurs in middle-aged or older persons, a global and significant increase in incidence amongst persons between 20 and 64 years old was noticed between 1990 and 2010, to which low and medium income countries substantially contributed. This, and the fact that stroke risk presents diversified patterns of geographic distribution between countries but also between regions of the same country, raises serious questions regarding the importance of socioeconomic determinants to general stroke morbidity and mortality, but also about the spatial variation in the association between determinants and stroke risk.

Two studies were operationalized here, both aiming to evaluate the spatial variation in the association between stroke mortality and its socioeconomic determinants within a study area comprising the Portuguese mainland territory. The first research study focused on the spatial and temporal variation of the association between determinants from several categories representing socioeconomic status (such as education) and stroke mortality, in persons below 65 years old, using the municipality as statistical unit. The present situation of economic crisis and its potential impacts in working-age population motivated the focus on this population cohort. The second research study was essentially focused on the opposite cohort, namely the population aged 64 years or older, and its statistical unit was the freguesia, an administrative level immediately below municipality and much smaller in size (in 2001 there were 278 municipalities and 4037 freguesias in the Portuguese mainland territory). This second research focused on the spatial variation of associations between poor housing indicators (such as the inability to warm up the dwelling during winter) and stroke mortality, since current evidence suggests that determinants from this category have a stronger effect on this population segment.

Results for both studies signaled the presence of expressive spatial variation in associations between each group of socioeconomic determinants and stroke risk, in both cohorts. It was also possible for both cases to evaluate which determinants revealed more relevance to stroke risk and analyze the strength and signal (negative or positive) of associations across the Portuguese mainland territory. Education emerged as a general protective factor in the first study, but also as a risk factor in some areas of the territory. The inability to warm up the dwelling during winter was assessed as a potential risk factor in the second research study, particularly in the northwest area of the Portuguese mainland.

The accomplished results suggest that the tested methodologies and tools have the potential to contribute to the epidemiological surveillance of stroke. In this context, these could be used to target priority areas for interventions aiming to prevent or mitigate this disease.

Keywords

Stroke, Socioeconomic Determinants, Spatial Analysis

Índice

Índice de figuras.....	xi
Índice de tabelas.....	xviii
Lista de abreviaturas, siglas ou acrónimos.....	xix
1. Introdução e contexto.....	1
1.1 Breve contexto e motivação.....	1
1.2. Objectivos de investigação.....	3
1.3 Breve descrição do conteúdo da tese.....	4
1.4 Acidente Vascular Cerebral – uma ameaça silenciosa.....	6
1.5 Evolução e impacto do AVC a diferentes escalas geográficas.....	12
1.5.1 Uma ameaça global – evolução e distribuição do AVC à escala mundial.	13
1.5.2 Uma ameaça nacional – o AVC no contexto Português.....	20
1.6 Determinantes do Acidente Vascular Cerebral.....	39
1.6.1 A abordagem clínica aos factores de risco individual de AVC.....	42
1.6.2 Factores de risco individual de AVC em crianças e adultos jovens.....	47
1.6.3 Determinantes socioeconómicos do AVC.....	51
1.7 Geografia da Saúde, Epidemiologia Espacial e Doenças Cerebrovasculares.....	58
1.7.1 Geografia da Saúde, Geografia da Doença, Epidemiologia Espacial.....	59
1.7.2 Considerações essenciais em Epidemiologia Espacial.....	66
1.7.3 Aplicações de Epidemiologia Espacial e Doenças Cerebrovasculares.....	77
2. Análise espaçotemporal da relação entre factores socioeconómicos e acidente vascular cerebral na população de Portugal Continental abaixo de 65 anos de idade.....	101
2.1 Introdução.....	101

2.2 Materiais e métodos	104
2.2.1 Área de estudo e fontes de dados	104
2.2.2 Cálculo do risco de AVC, mapeamento e análise espacial inicial	105
2.2.3 Análises de regressão univariada de Poisson	106
2.2.4 Modelos multivariados de regressão binomial negativa	108
2.2.5 Regressão geograficamente ponderada de Poisson, global e local	109
2.3 Resultados	110
2.3.1 Distribuição especial do risco de AVC e os factores socioeconómicos ..	110
2.3.2 Determinantes socioeconómicos da distribuição geográfica do risco de AVC	116
2.3.3 Avaliação comparativa das abordagens de modelação e análise espacial dos coeficientes locais.....	118
2.4 Discussão	124
2.5 Conclusão.....	128
3. Condições precárias de habitação e mortalidade por acidente vascular cerebral ao nível da freguesia em Portugal Continental, na população com mais de 65 anos de idade.....	131
3.1 Introdução	131
3.2 Materiais e métodos	140
3.2.1 Área de estudo e fontes de dados	140
3.2.2 Cálculo do indicador de risco de AVC, variáveis explicativas e análise espacial visual	141
3.2.3 Análise univariada de clustering espacial	144
3.2.4 Análise de clustering multivariada.....	146
3.3 Resultados	154

3.3.1 Distribuição espacial do risco de AVC e dos determinantes relacionados com a qualidade da habitação	154
3.3.2 Análise univariada de clustering espacial do risco de AVC e determinantes relacionados com a qualidade da habitação	162
3.3.3 Análise de clustering multivariada.....	167
3.4 Discussão	171
3.5 Conclusão.....	176
4. Discussão geral e conclusões	179
5. Referências bibliográficas.....	186
6. Anexos.....	201

Índice de figuras

Figura 1 – Principais constituintes do cérebro.....	7
Figura 2 – Controle de algumas funções por partes diferenciadas do cérebro.	8
Figura 3 – Irrigação sanguínea do cérebro.....	9
Figura 4 – Artérias posterior, média e anterior, com vista lateral do exterior e do interior mediano do hemisfério direito do cérebro.	9
Figura 5 – Sequência temporal e potenciais consequências de um AVC agudo.	12
Figura 6 – Importância relativa dos óbitos por doenças cardiovasculares na mortalidade total em países em desenvolvimento, industrializados e no mundo, em 1980..	14
Figura 7 – Evolução da taxa de mortalidade padronizada, por doenças cerebrovasculares, por 100000 habitantes, entre 1971 e 2009, em Portugal..	21
Figura 8 – Evolução da taxa de mortalidade padronizada, por doenças cerebrovasculares, por 100000 habitantes, entre 1980 e 2011, nos países da União Europeia.....	21
Figura 9 – Evolução da taxa de mortalidade padronizada, por doenças cerebrovasculares em Portugal, por 100 000 habitantes (2002-2012).	22
Figura 10 – Número de óbitos por doenças cerebrovasculares em Portugal Continental (2007-2011)..	22
Figura 11 – Evolução do número de utentes saídos dos episódios de internamento por AVC, em Portugal Continental, entre 2000 e 2012.	24
Figura 12 – Anos de vida perdidos ajustados pela incapacidade, por doenças cerebrovasculares, em Portugal (2006-2010).	25
Figura 13 – Taxa de anos de vida perdidos ajustados pela incapacidade, por doenças cerebrovasculares, em Portugal (2002-2012).	25
Figura 14 – Taxa de mortalidade padronizada por doença cerebrovascular em pessoas de idade igual ou superior a 65 ou 70 anos, por 100 000 habitantes, em Portugal Continental (2007-2011).....	26

Figura 15 – Taxa de mortalidade padronizada por doença cerebrovascular em pessoas de idade inferior a 65 ou 70 anos, por 100 000 habitantes, em Portugal Continental (2007-2011).	26
Figura 16 – Anos de vida perdidos ajustados pela incapacidade por doenças cerebrovasculares entre 2002 e 2012, em homens (esquerda) e mulheres (direita), em Portugal.....	27
Figura 17 – Taxa de anos de vida perdidos ajustados pela incapacidade por doenças cerebrovasculares entre 2002 e 2012, em homens (esquerda) e mulheres (direita), em Portugal.....	27
Figura 18 – Taxa de mortalidade padronizada por doenças cerebrovasculares por 100 000 habitantes entre 2002 e 2012, em homens (esquerda) e mulheres (direita).	27
Figura 19 – Taxa de mortalidade padronizada por doenças cerebrovasculares por 100 000 habitantes entre 2009 e 2012, em homens (esquerda) e mulheres (direita) com 65 ou mais anos de idade.	28
Figura 20 – Taxa de mortalidade padronizada por doenças cerebrovasculares por 100 000 habitantes entre 2009 e 2012, em homens (esquerda) e mulheres (direita) com menos de 65 anos de idade.	28
Figura 21 – Evolução do número de utentes saídos dos episódios de internamento por AVC, por sexos, entre 2000 e 2012.	28
Figura 22 – Número de utentes saídos dos episódios de internamento por AVC, em mulheres, em dois escalões etários, entre 2000 e 2012.	29
Figura 23 – Número de utentes saídos dos episódios de internamento por AVC, em homens, em dois escalões etários, entre 2000 e 2012.....	29
Figura 24 – Taxa de mortalidade padronizada por AVC em Portugal Continental em 2002, 2007 e 2012, por 100 000 habitantes (ambos os sexos).....	31
Figura 25 – Taxa de mortalidade padronizada por AVC em Portugal Continental em 2002, 2007 e 2012, por 100 000 habitantes (homens).	32

Figura 26 – Taxa de mortalidade padronizada por AVC em Portugal Continental em 2002, 2007 e 2012, por 100 000 habitantes (mulheres).....	32
Figura 27 – Taxa de mortalidade padronizada por AVC em Portugal Continental em 2009, 2010 e 2012, ambos os sexos, 65 ou mais anos de idade, por 100 000 habitantes.	33
Figura 28 – Taxa de mortalidade padronizada por AVC em Portugal Continental em 2009, 2010 e 2012, ambos os sexos, menos de 65 anos de idade, por 100 000 habitantes.	33
Figura 29 – Taxa de mortalidade padronizada por AVC em Portugal Continental em 2009, 2010 e 2012, homens com 65 ou mais anos de idade, por 100 000 habitantes. ...	34
Figura 30 – Taxa de mortalidade padronizada por AVC em Portugal Continental em 2009, 2010 e 2012, mulheres com 65 ou mais anos de idade, por 100 000 habitantes. .	34
Figura 31 – Taxa de mortalidade padronizada por AVC em Portugal Continental em 2009, 2010 e 2012, homens com menos de 65 anos de idade, por 100 000 habitantes..	35
Figura 32 – Taxa de mortalidade padronizada por AVC em Portugal Continental em 2009, 2010 e 2012, mulheres com menos de 65 anos de idade, por 100 000 habitantes.	35
Figura 33 – Anos de vida perdidos ajustados pela incapacidade por doenças cerebrovasculares em 2002, 2007 e 2012, todas as idades, ambos os sexos.	36
Figura 34 – Anos de vida perdidos ajustados pela incapacidade por doenças cerebrovasculares em 2002, 2007 e 2012, em homens, todas as idades.	36
Figura 35 – Anos de vida perdidos ajustados pela incapacidade por doenças cerebrovasculares em 2002, 2007 e 2012, em mulheres, todas as idades.....	37
Figura 36 – Taxa de anos de vida perdidos ajustados pela incapacidade por doenças cerebrovasculares em 2002, 2007 e 2012, todas as idades, ambos os sexos	37
Figura 37 – Taxa de anos de vida perdidos ajustados pela incapacidade por doenças cerebrovasculares em 2002, 2007 e 2012, homens, todas as idades.....	38

Figura 38 – Taxa de anos de vida perdidos ajustados pela incapacidade por doenças cerebrovasculares em 2002, 2007 e 2012, mulheres, todas as idades..	38
Figura 39 – Cadeia actual de causalidade em doenças cardiovasculares.	41
Figura 40 – Óbitos por doença cerebrovascular em crianças e adultos jovens em 2005.	48
Figura 41 – Modelo de relações entre estatuto socioeconómico e AVC.	52
Figura 42 – Quadro conceptual da análise espacial de dados epidemiológicos.	61
Figura 43 – Mapa de John Snow do surto de cólera de Londres de 1854	63
Figura 44 – Exemplos de representações de dados em SIG	69
Figura 45 – Localizações pontuais de fontes de emissão de poluentes	78
Figura 46 – Modelação de superfícies através de <i>Kernel Smoothing</i>	79
Figura 47 – Superfície raster da densidade da média anual das contagens diárias de tráfego obtida através de <i>Kernel Density</i> .	80
Figura 48 – Concentrações de dióxido de azoto, por áreas censitárias no sul de Londres.	81
Figura 49 – Resultado do cálculo do Índice I de Moran.	85
Figura 50 – Processo de aleatorização no cálculo do Índice I de Moran.	86
Figura 51 – Mapa da taxa média anual de variação nas taxas de mortalidade em homens caucasianos entre 1970 e 1978 nos Estados Unidos da América.	87
Figura 52 – Clusters das taxas de hospitalização por AVC em pessoas com 65 ou mais anos de idade no período 2005-2006	88
Figura 53 – Mecanismo de um kernel espacial.	97
Figura 54 – <i>Kernels</i> espaciais com larguras de banda fixa ou adaptativa.	98
Figura 55 – Taxa de mortalidade padronizada por AVC em municípios de Portugal Continental, 1992-1996, por 100 000 habitantes, por sexos.	111

Figura 56 – Taxa de mortalidade padronizada por AVC em municípios de Portugal Continental, 2002-2006, por 100 000 habitantes, por sexos.....	112
Figura 57 – Pessoas que trabalhavam mais de 45 horas por semana (%), em 1991, em municípios de Portugal Continental, por sexos	113
Figura 58 – Pessoas que trabalhavam mais de 45 horas por semana (%), em 2001, em municípios de Portugal Continental, por sexos.	114
Figura 59 – Pessoas que atingiram o ensino básico (%) em 1991, em municípios de Portugal Continental, por sexos	115
Figura 60 – Pessoas que atingiram o ensino básico (%) em 2001, em municípios de Portugal Continental, por sexos.	115
Figura 61 – Distribuição geográfica dos coeficientes do modelo GWR de Poisson local para homens que trabalhavam mais de 45 horas por semana (%) no período 1992-1996, em municípios de Portugal Continental.....	120
Figura 62 – Distribuição geográfica dos coeficientes do modelo GWR de Poisson local para homens que atingiram o ensino básico (%), no período 2002-2006, em municípios de Portugal Continental.	122
Figura 63 – Distribuição geográfica dos coeficientes do modelo GWR de Poisson local para mulheres que atingiram o ensino básico (%), no período 1992-1996, em municípios de Portugal Continental.	123
Figura 64 – Distribuição geográfica dos coeficientes do modelo GWR de Poisson local para mulheres que atingiram o ensino básico (%), no período 2002-2006, em municípios de Portugal Continental.	124
Figura 65 – Rácios padronizados de mortalidade por AVC em homens com mais de 64 anos de idade e homens de todas as idades, em municípios de Portugal Continental, 1998-2004.	156
Figura 66 – Rácios padronizados de mortalidade por AVC em mulheres com mais de 64 anos de idade e mulheres de todas as idades em municípios de Portugal Continental, 1998-2004.	157

Figura 67 – Rácios padronizados de mortalidade por AVC, ambos os sexos, mais de 64 anos de idade e ambos os sexos, todas as idades, em municípios de Portugal Continental, 1998-2004.	158
Figura 68 – População residente sem qualquer tipo de aquecimento (%), por freguesia.....	159
Figura 69 – Alojamentos sem aquecimento central (%), por freguesia.....	159
Figura 70 – Alojamentos superlotados (%), por freguesia.....	160
Figura 71 – População residente a viver em apartamentos (%), por freguesia.....	160
Figura 72 – Peso dos residentes em alojamentos arrendados nos residentes proprietários (%), por freguesia.....	161
Figura 73 – População residente em edifícios construídos antes de 1960 (%), por freguesia.....	161
Figura 74 – População residente em alojamentos familiares não clássicos (%), por freguesia.....	161
Figura 75 – Clusters de freguesias por SMRi, homens, maiores de 64 anos de idade.....	163
Figura 76 – Clusters de freguesias por SMRi, homens, todas as idades.....	163
Figura 77 – Clusters de freguesias por SMRi, mulheres, maiores de 64 anos de idade....	163
Figura 78 – Clusters de freguesias por SMRi, mulheres, todas as idades.....	163
Figura 79 – Clusters de freguesias por SMRi, todos os sexos, maiores de 64 anos de idade.....	164
Figura 80 – Clusters de freguesias por SMRi, todos os sexos, todas as idades.....	164
Figura 81 – Clusters de freguesias por % de população residente sem qualquer tipo de aquecimento.....	164
Figura 82 – Clusters de freguesias por % de alojamentos sem aquecimento central...	164
Figura 83 – Clusters de freguesias por % de alojamentos superlotados.....	165

Figura 84 – Clusters de freguesias por % de população residente a viver em apartamentos.....	165
Figura 85 – Clusters de freguesias por peso dos residentes em alojamentos arrendados nos residentes proprietários (%).....	165
Figura 86 – Clusters de freguesias por % de População residente em edifícios construídos antes de 1960.....	165
Figura 87 – Clusters de freguesias por % de população residente em alojamentos familiares não clássicos.....	166
Figura 88 – Dendrograma final	168
Figura 89 – Gráfico de coeficientes de fusão, número de clusters, e clusters a adoptar.....	169
Figura 90 – 3 clusters finais multivariados de freguesias: combinação do critério de agregação de Ward com a distância de Gower.....	169

Índice de tabelas

Tabela 1 – Doenças Cerebrovasculares, Classificação Internacional das Doenças, revisão 10.....	11
Tabela 2 – Comparação das taxas de mortalidade padronizadas por Doenças Cérebro-cardiovasculares, por 100 000 habitantes, em Portugal e alguns países Europeus em 2010.	23
Tabela 3 – Estatísticas descritivas dos factores socioeconomicos investigados para potenciais associações com o risco de morte por AVC.....	112
Tabela 4 – Avaliação comparativa de desempenho dos modelos testados.....	118
Tabela 5 – Avaliação da estacionaridade dos coeficientes nos modelos GWR de Poisson locais.....	119
Tabela 6 – Indicadores de má qualidade na habitação e consequências para a saúde física.....	132
Tabela 7 – Indicadores de má qualidade na habitação e consequências para a saúde mental.....	132
Tabela 8 – Estatísticas descritivas dos rácios padronizados de mortalidade por coorte	155
Tabela 9 – Estatísticas descritivas das variáveis explicativas.....	158
Tabela 10 – Médias e medianas dos SMR _i por tipo de cluster de variável explicativa	167
Tabela 11 – Coeficientes de correlação cofenética por teste.....	168
Tabela 12 – Médias dos valores padronizados das variáveis explicativas nos 3 clusters finais.....	170
Tabela 13 – Valores SMR _i médios e medianos por cluster multivariado.....	171

Lista de abreviaturas, siglas ou acrónimos

- ACSS - Administração Central de Sistemas de Saúde
- AIC - Akaike Information Criterion
- AIT - Acidente Isquémico Transitório
- AVC - Acidente Vascular Cerebral
- DALYs - Disability Adjusted Life Years
- DCV - Doenças Cardiovasculares
- DGS - Direcção-Geral da Saúde
- DIC - Doença Isquémica do Coração
- EROS - Earth Resources Observation and Science
- ESDA - Exploratory Spatial Data Analysis
- GBD - Global Burden of Diseases, Injuries, and Risk Factors Study
- GPS - Global Positioning System
- GWR - Geographically Weighted Regression
- IC - Intervalo de Confiança
- INE - Instituto Nacional de Estatística
- LISA - Local Indicators of Spatial Autocorrelation
- MAUP - Modifiable Area Unit Problem
- NUT - Nomenclatura das Unidades Territoriais
- OGC - Open Geospatial Consortium
- SAR - Spatial Autoregressive Model
- SEM - Spatial Error Model
- SES - Socioeconomic Status
- SGBD - Sistema de Gestão de Bases de Dados
- SIG - Sistema de Informação Geográfica

SMR - Standard Mortality Ratio

SOM – Self Organizing Map

SNIG - Sistema Nacional de Informação Geográfica

TIN - Triangular Irregular Network

WHO - World Health Organization

1. Introdução e contexto

1.1 Breve contexto e motivação

O Acidente Vascular Cerebral (AVC) constitui um problema de saúde pública de grande magnitude em todo o mundo, tanto em termos de mortalidade como de morbidade. Em 2010, esta doença representava, globalmente, a segunda maior causa de morte e a terceira causa de Anos de Vida Perdidos Ajustados pela Incapacidade ou *Disability-Adjusted Life Years (DALYs)*, tendo sido identificadas nesse ano 16,9 milhões de pessoas que sofreram um primeiro episódio, e contabilizados 5,9 milhões de óbitos. Mais ainda, em alguns indicadores como os DALYs e o total de sobreviventes a um episódio, registou-se um aumento estatisticamente significativo entre 1990 e 2010 (Krishnamurthi et al., 2013).

E embora esta doença ocorra com mais frequência em adultos de meia-idade ou em idosos (WHO, 2014a), também entre 1990 e 2010 ocorreu um aumento global e significativo na incidência em pessoas com idades compreendidas entre os 20 e os 64 anos, para o qual contribuíram substancialmente os países de baixo e médio rendimento (Krishnamurthi et al., 2013).

Esta situação levanta, entre outras, a interrogação acerca do peso dos determinantes socioeconómicos desta doença. Sendo o AVC uma doença multifactorial, os seus determinantes podem ser abordados tanto ao nível individual, como ecológico. Por um lado, existe já muito conhecimento sobre o peso dos determinantes individuais do tipo clínico no risco de AVC, em particular da hipertensão, classificada como a situação de maior risco entre os determinantes modificáveis (Brainin e Heiss, 2009). Por outro lado, estes determinantes individuais constituem, em geral, a fase final de uma cadeia de eventos, que se inicia numa conjugação de situações desfavoráveis de carácter ambiental e social, que contribuem para o aumento de comportamentos nocivos e conduzem, por sua vez, aos factores de risco individual de tipo clínico (Labarthe, 2011). Esta conjugação de situações sociais e ambientais desfavoráveis pode ser entendida como estatuto socioeconómico, e engloba determinantes como o nível de escolaridade, a ocupação profissional (Cox et al., 2006) ou até a capacidade de aquecer devidamente a

habitação durante o Inverno (Vasconcelos et al., 2011), entre outros. E apesar dos elos de ligação entre estatuto socioeconómico e ocorrência de AVC terem vindo a ser identificados, o peso exacto do estatuto socioeconómico no risco desta doença permanece ainda controverso (Cox et al., 2006); (Addo et al., 2012). Desta forma, alguns autores defendem a necessidade de se investigar e actuar preventivamente sobre estes factores, na medida em que se encontram no início da cadeia de eventos, ou seja, actuar sobre a “causa das causas” (Marmot e Wilkinson, 2005).

Outra questão muito relevante aqui tem a ver com o facto do risco de AVC apresentar padrões diversos no que respeita à sua distribuição geográfica, não só entre países mas também entre regiões do mesmo país (Cox et al., 2006); (Addo et al., 2012); (Krishnamurthi et al., 2013). Olhando, por exemplo, para o caso de Portugal, o risco de morte em consequência de um AVC por NUTs II (Eurostat, 2014a) variava, em 2012, entre 53 óbitos na NUTII do Algarve e 66,7 por 100.000 habitantes na NUTII do Norte (INE, 2014), não sendo conhecidas as razões subjacentes a estas variações regionais. Sendo assim, identificar os determinantes destas disparidades geográficas deve constituir certamente um desígnio em Saúde Pública, no que diz respeito aos esforços de prevenção e mitigação desta doença.

A motivação para a realização do presente trabalho de investigação teve a sua origem no dealbar, em 2008, da mais recente crise económica e financeira global, e no consequente agravamento geral das condições socioeconómicas dos cidadãos portugueses que decorre deste contexto adverso. Este contexto, em conjugação com a discussão prévia sobre a “tríade” epidemiológica composta por risco de AVC, determinantes socioeconómicos e distribuição espacial de ambos, veio despertar o interesse em avaliar o estado da arte da relação entre estes três vectores no nosso país, com recurso a ferramentas e metodologias da Epidemiologia Espacial, visando contribuir para os esforços de Saúde Pública focados na prevenção e mitigação desta doença.

Coloca-se, desta forma, a hipótese da mortalidade por AVC em Portugal Continental ser influenciada por determinantes socioeconómicos e ambientais, e de esta associação não ser geograficamente estática, sendo possível melhorar os modelos a utilizar nos estudos de Saúde Pública através do recurso a técnicas que permitam ter em conta este contexto.

1.2. Objectivos de investigação

Esta investigação tem como objectivo primordial recorrer à utilização de Informação Geográfica, ferramentas relacionadas e metodologias de Epidemiologia Espacial para operacionalizar casos de estudo focados na avaliação da influência que os factores socioeconómicos possam exercer na distribuição espacial da mortalidade por Acidente Vascular Cerebral na população do território de Portugal Continental. Tem também como objectivo pertinente que os estudos efectuados possam contribuir para melhorar a tomada de decisões ao nível das medidas de prevenção e mitigação dos casos de AVC em Portugal.

Estes dois objectivos essenciais desdobram-se em vários objectivos específicos, designadamente:

- 1- Introduzir e contextualizar o Acidente Vascular Cerebral (AVC), começando pela apresentação de conceitos e definições essenciais, relativos a esta doença e seu impacto na saúde humana e qualidade de vida;
- 2- Estabelecer o ponto de situação no que respeita à evolução, impacto e distribuição espacial do AVC a diferentes escalas geográficas de análise, começando pela escala global e passando depois ao contexto específico de Portugal;
- 3- Apresentar o estado do conhecimento acerca dos determinantes do AVC, tanto no que diz respeito às condições de carácter clínico, como do tipo socioeconómico, sendo que ambos podem ser avaliados ao nível individual como ecológico, e também no que concerne à estrutura da relação entre estes dois tipos de determinantes, a qual pode ser igualmente analisada ao nível individual ou ecológico.
- 4- Terminar a fase de contextualização apresentando metodologias, informação geográfica e ferramentas relacionadas, actualmente utilizadas no âmbito da Epidemiologia Espacial. Simultaneamente tenta-se apresentar, sempre que possível, casos de estudo relevantes que analisam a associação espacial entre os determinantes e o risco de AVC. Ao longo de toda esta primeira fase, optou-se por

uma perspectiva ampla, no sentido de se abordar tanto a mortalidade como a morbidade por esta doença.

- 5- Partindo do conhecimento adquirido na fase anterior, e já no âmbito do desenvolvimento de estudos de aplicação, tomar decisões fundamentais relativas ao desenho epidemiológico dos casos de estudo a pôr em prática. Mais especificamente, determinar as metodologias no âmbito da Epidemiologia Espacial a empregar nos diferentes estudos. Decidir também, de forma específica para cada caso de estudo, quais os determinantes socioeconómicos, coortes populacionais de mortalidade por AVC, períodos temporais e unidades estatísticas espaciais a abordar, de acordo com a disponibilidade de dados. Uma vez tomado este conjunto de decisões, passar à operacionalização dos estudos de aplicação.
- 6- Na sequência do desenvolvimento e finalização dos estudos de aplicação, avaliar e retirar conclusões sob vários ângulos de análise. Importa assim aferir a presença e intensidade de variação espacial na associação entre determinantes e risco de morte por AVC. Importa também avaliar os determinantes mais relevantes em cada estudo e a força da sua associação com o risco, em contexto local ou para todo o território continental português. Outra questão relevante prende-se com a avaliação do grau de adequação das opções tomadas no ponto anterior. Finalmente, mas não menos importante, pretende-se determinar o grau de contributo dos estudos efectuados para a prevenção e mitigação do AVC no território de Portugal Continental, e em última análise, da utilidade da adaptação e integração das suas metodologias na vigilância epidemiológica do AVC. Um exemplo possível reside na capacidade de escolha de áreas de intervenção prioritárias no território de Portugal Continental, com base nas análises efectuadas.

1.3 Breve descrição do conteúdo da tese

Esta dissertação é constituída por quatro capítulos, sendo que o presente capítulo introduz, a partir deste ponto, os elementos fundamentais para enquadrar a problemática abordada nesta dissertação. Começa-se por definir o Acidente Vascular Cerebral (AVC) na sua essência, enquanto doença (subcapítulo 1.4). Dado que este estudo se foca em grande parte na componente de distribuição espacial desta doença, procede-se também a

uma análise da sua importância e impacto a diferentes escalas geográficas, começando pela escala mundial (subcapítulo 1.5.1) e passando de seguida para contexto de Portugal (subcapítulo 1.5.2), como ponto de partida para os dois estudos que irão mais à frente integrar esta dissertação (capítulos 2 e 3), os quais têm como área de estudo o território de Portugal Continental. Procede-se também, no subcapítulo 1.6, à introdução ao segundo grande vector deste estudo, os determinantes do AVC, com especial enfoque nos seus determinantes socioeconómicos (subcapítulo 1.6.3). O primeiro capítulo é encerrado com uma abordagem introdutória às metodologias no âmbito da Epidemiologia Espacial (subcapítulo 1.7) que suportam o desenvolvimento dos estudos contidos nos capítulos 2 e 3, sendo essas metodologias ilustradas sempre que possível com casos de estudo no domínio do AVC, constantes da literatura. O capítulo 2 descreve a primeira investigação efectuada no âmbito deste projecto de doutoramento, focada na variação geográfica e temporal da associação entre um conjunto de determinantes socioeconómicos (entre os quais o nível de escolaridade) e a mortalidade por AVC em pessoas abaixo dos 65 anos. Embora este coorte populacional não seja responsável pela maior fatia de mortalidade, suscitou todavia interesse, motivado pela presente situação de crise económica e seus potenciais impactos na população em idade laboral. De seguida, o capítulo 3 aborda essencialmente o coorte populacional inverso, a população com mais de 64 anos. Investiga desta forma a variação espacial da associação entre condições adversas de habitabilidade (entre as quais a incapacidade de aquecer a habitação no Inverno) e mortalidade por AVC, na medida em que a evidência existente parece indicar ser este coorte populacional mais afectado por este tipo específico de determinantes socioeconómicos. Finalmente, o capítulo 4 encerra esta dissertação, discutindo os resultados obtidos nos dois estudos anteriormente descritos, nos seus desenvolvimentos mais relevantes e também nas limitações encontradas, e termina apontando vários rumos para estudos futuros cujo interesse foi sem dúvida suscitado no decurso das duas investigações efectuadas.

1.4 Acidente Vascular Cerebral – uma ameaça silenciosa

O Acidente Vascular Cerebral consiste numa lesão cerebral causada por uma deficiência no fluxo sanguíneo que irriga o cérebro. Tendo em conta que cada parte específica do cérebro é responsável pelo desempenho de uma determinada função (visão, linguagem, coordenação motora...), diferentes sintomas poderão ocorrer, em função da parte do cérebro que for privada do fornecimento de sangue (Rudd et al., 2004). Também a magnitude do episódio pode variar, consoante a artéria do cérebro e o sector da mesma que foi afectado.

O tipo mais comum de AVC (80% dos casos) designa-se por isquémico, e ocorre quando uma das artérias do cérebro é bloqueada por um coágulo sanguíneo, o qual se pode formar localmente, pela acumulação progressiva de colesterol na parede arterial, ou noutra parte do corpo como o coração ou a artéria aorta, sendo que o coágulo irá viajar até se alojar numa parte do sistema circulatório onde a sua dimensão supera a de um determinado vaso sanguíneo, ficando bloqueado nesse ponto. Este bloqueio irá privar a parte do cérebro alimentado pela artéria em causa, de oxigénio e nutrientes, provocando danos nas células cerebrais dependentes. Os restantes 20% dos casos são devidos a hemorragias cerebrais, situação denominada AVC hemorrágico, as quais podem ocorrer no interior do cérebro (hemorragia intracerebral) ou na sua superfície (hemorragia subaracnóide). Os danos deste último tipo de AVC processam-se através do esmagamento dos tecidos cerebrais circundantes pela pressão do sangue libertado pela hemorragia.

Para um melhor entendimento deste fenómeno, importa ter em conta a estrutura do cérebro e da sua irrigação sanguínea. Desta forma, o cérebro divide-se em duas metades, os hemisférios direito e esquerdo (figura 1).

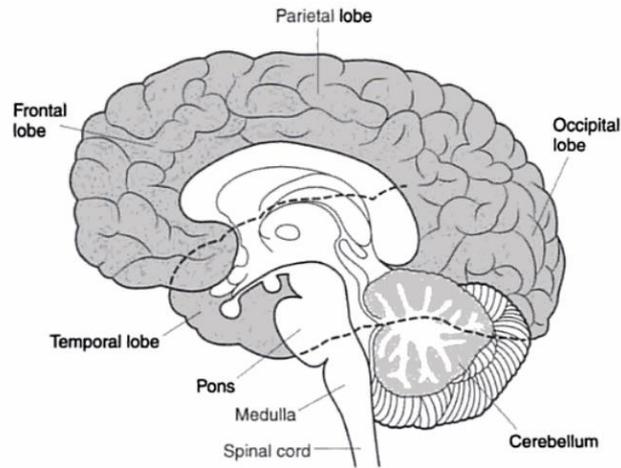


Figura 1 – Principais constituintes do cérebro. Fonte: Rudd et al. (2004).

Na sua base encontra-se o cerebelo, e a ligação entre o cérebro e a espinal medula é assegurada pelo tronco cerebral. A informação proveniente das terminações nervosas do corpo é enviada através da espinal medula e depois através do tronco cerebral para um dos hemisférios. Com base nesta informação, o cérebro decidirá as acções que necessita de executar e enviará, através do mesmo caminho mas na direcção inversa, indicações correspondentes aos músculos do corpo.

A execução de uma determinada função do corpo humano é controlada por uma parte específica do cérebro, sendo que o hemisfério esquerdo controla maioritariamente a metade direita do corpo e o hemisfério direito a metade esquerda. A figura 2 exemplifica a relação entre partes específicas do cérebro e o controlo de determinadas funções.

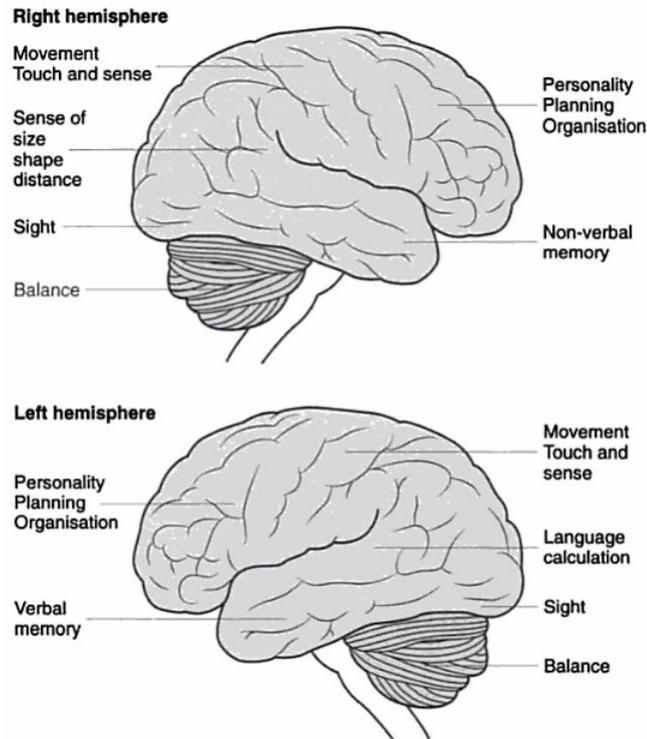


Figura 2 – Controle de algumas funções por partes diferenciadas do cérebro.
Fonte: Rudd et al. (2004).

O sangue flui para o cérebro desde a artéria aorta, passando para uma de quatro artérias que o conduzem ao cérebro. Na parte anterior encontram-se duas artérias carótidas que ladeiam a traqueia e na parte posterior duas artérias vertebrais ladeando a coluna vertebral. Todas as quatro artérias se encontram ligadas no Polígono de Willis após a sua entrada no crânio (figura 3). Este mecanismo tem um papel importante, pois na eventualidade do bloqueio de uma das artérias, o fluxo de sangue proveniente de outras artérias poderá evitar a ocorrência de danos relevantes.

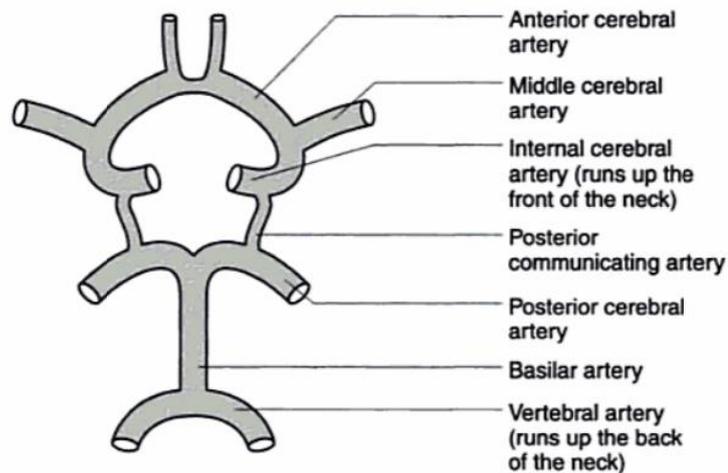


Figura 3 – Irrigação sanguínea do cérebro. Fonte: Rudd et al. (2004).

Saindo do Polígono de Willis, encontram-se as seis artérias mais importantes envolvidas na irrigação cerebral, nomeadamente as artérias anterior, média e posterior, existindo uma para cada lado do cérebro, e que irrigam as partes frontal, média e traseira do cérebro, respectivamente (figura 4). Qualquer uma pode sofrer um bloqueio potencialmente resultante num AVC, embora o caso mais comum seja o bloqueio de uma artéria cerebral média (Rudd et al., 2004).

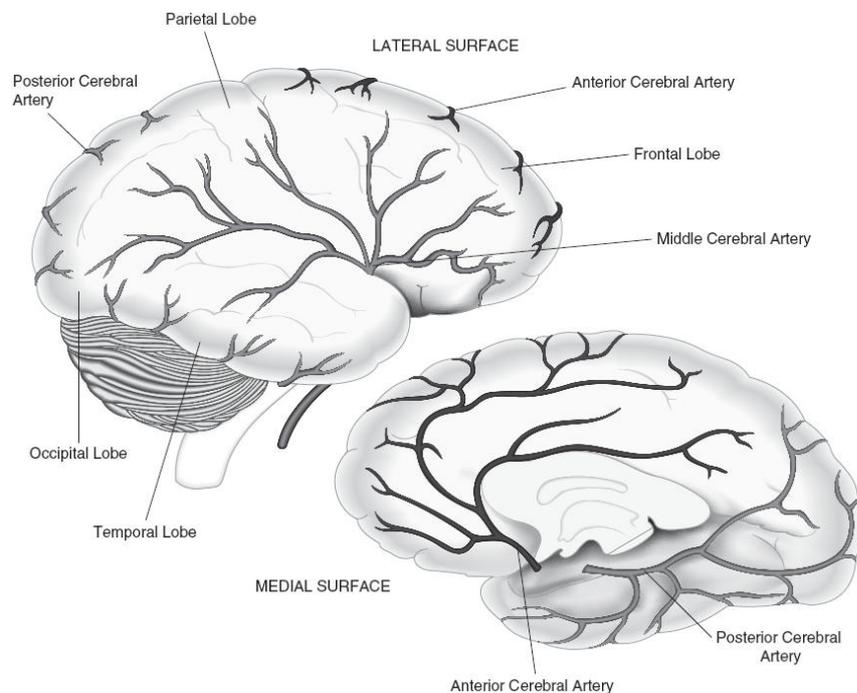


Figura 4 – Artérias posterior, média e anterior, com vista lateral do exterior e do interior mediano do hemisfério direito do cérebro. Fonte: Labarthe (2011).

Na realidade, para lá da classificação dos episódios de AVC em isquémico ou hemorrágico, existem múltiplos caminhos e processos que podem levar à interrupção da circulação sanguínea cerebral, e uma grande variedade de manifestações pode ocorrer, provocadas por diferentes efeitos que derivam da área do cérebro afectada (Labarthe, 2011).

Entre os múltiplos caminhos e processo possíveis, alguns têm a sua origem antes do cérebro, em vasos intermédios, como as artérias carótidas exteriores, que se prolongam em artérias carótidas interiores. Ao serem afectadas por aterosclerose, estas podem perturbar a irrigação sanguínea do cérebro ainda antes de serem atingidos os vasos intracranianos. Estas artérias podem ainda transportar para o cérebro pequenos coágulos sanguíneos, ou trombos, formados no coração e que assim vão progredir através de vasos de diâmetro progressivamente menor até ficarem alojados numa artéria, bloqueando a circulação sanguínea daí para a frente. Outros processos relacionam-se com as artérias vertebrobasilares, que embora de menores dimensões, atingem a parte posterior do cérebro e cuja obstrução pode provocar efeitos nefastos nas funções cerebrais. Ainda outro tipo de processos resulta de um segmento doente e dilatado de uma artéria (aneurisma) que pode romper, causando uma hemorragia no cérebro. Uma variante deste processo pode ocorrer aquando da ruptura de uma artéria, espontânea ou causada por uma lesão traumática do crânio, produzindo-se uma hemorragia subaracnóidea, assim chamada por ocorrer sob a superfície que delimita o crânio, a membrana aracnóide. Finalmente, um trombo pode ainda formar-se dentro da circulação venosa do cérebro (Labarthe, 2011).

Desta forma, não só pelo mecanismo como ocorre, mas também pelos seus efeitos, o termo AVC pode significar diferentes ocorrências, agrupadas sob a designação geral de Doenças Cerebrovasculares nos códigos I60-I69 da Classificação Internacional de Doenças, revisão 10 (ICD-10, 2015), como pode ser observado na tabela 1.

Tabela 1 – Doenças Cerebrovasculares, Classificação Internacional das Doenças, revisão 10.

I60	Hemorragia subaracnóidea
I61	Hemorragia intracerebral
I62	Outras hemorragias intracranianas não-traumáticas
I63	Enfarte cerebral
I64	Acidente vascular cerebral, não especificado como hemorrágico ou isquémico
I65	Oclusão e estenose de artérias pré-cerebrais, não resultando em enfarte cerebral
I66	Oclusão e estenose de artérias cerebrais, não resultando em enfarte cerebral
I67	Outras doenças cerebrovasculares
I68	Transtornos cerebrovasculares em doenças classificadas noutra parte
I69	Sequelas de doenças cerebrovasculares

Fonte: adaptado de ICD-10 (2015)

A investigação em Doenças Cerebrovasculares constitui, nos nossos dias, um campo de investigação muito dinâmico e único, na medida em que envolve especialidades muito diversas como a Neurologia, Medicina Interna, Cirurgia, Radiologia, Epidemiologia, Cardiologia, Hematologia, Psicologia e Reabilitação (ESO, 2013).

No que diz respeito às manifestações da doença, o traço distintivo de um AVC agudo é o facto de ocorrer subitamente, podendo implicar perda repentina de consciência e das funções motoras e sensoriais num dos lados do corpo. Entre as manifestações clínicas mais importantes da interrupção aguda do fornecimento de sangue arterial a uma ou mais áreas do cérebro contam-se o entorpecimento ou enfraquecimento súbito da face, de um braço ou perna, confusão súbita, dificuldade na fala ou no entendimento, dificuldade em andar, tonturas, perda de equilíbrio ou coordenação, dificuldades de visão num olho ou mesmo nos dois, dor de cabeça súbita e severa sem causa aparente (Labarthe, 2011).

A severidade dos episódios pode todavia variar. Os sinais e sintomas podem diminuir ou desaparecer em minutos ou horas sem deixar mesmo problemas clínicos, ou em alternativa podem persistir e progredir, culminando em morte ou incapacidade permanente. Assim, um episódio que se resolve totalmente em 24 horas é designado Acidente Isquémico Transitório (AIT), enquanto se persistir mais de 24 horas será denominado um AVC completo. Um AVC completo seguido de morte num período de 28 dias a contar do episódio é denominado AVC fatal. Esta sequência pode ser observada na figura 5.

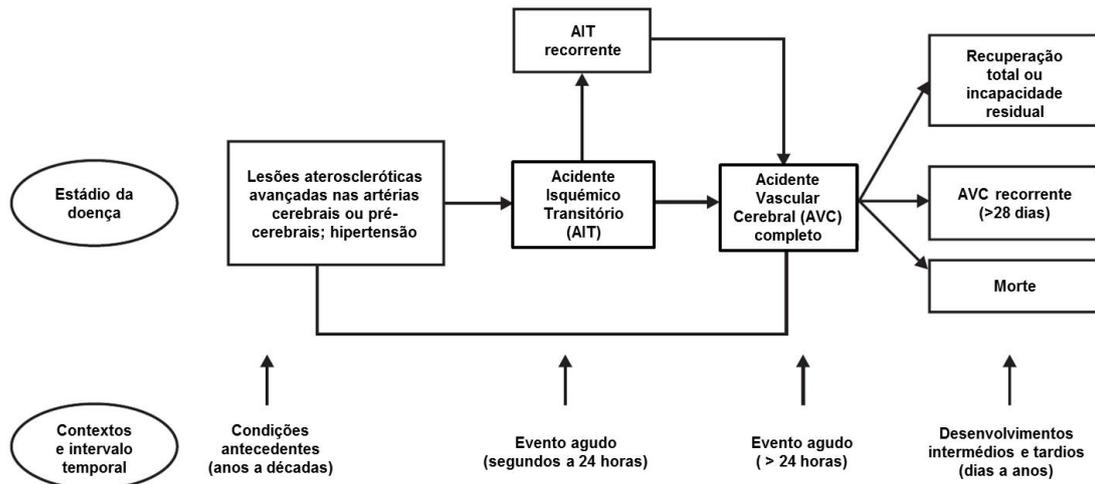


Figura 5 – Sequência temporal e potenciais consequências de um AVC agudo.

Fonte: adaptado de Labarthe (2011).

O carácter abrupto da manifestação desta doença alerta para a necessidade das pessoas atingidas por esta patologia serem assistidas no menor espaço de tempo possível; a este respeito, Labarthe (2011) refere que, nos Estados Unidos da América, quase metade de todas as mortes por AVC (48,3%) em 1999 ocorreram no período que antecede o transporte do paciente para o Hospital ou durante o próprio transporte.

1.5 Evolução e impacto do AVC a diferentes escalas geográficas

O Acidente Vascular Cerebral representa um problema de saúde pública de grande magnitude ao nível global, tanto em termos de mortalidade como de morbilidade, e com padrões de distribuição geográfica diferenciados, não apenas entre países mas também entre regiões do mesmo país (Cox et al., 2006). Estes padrões parecem estar relacionados com a distribuição, igualmente heterogénea, dos determinantes desta doença. Pretende-se assim, no subcapítulo seguinte, introduzir o conhecimento disponível acerca desta doença ao nível global, no que diz respeito ao seu impacto na saúde pública, tendências, distribuição geográfica, relação com o nível socioeconómico dos países e faixas etárias afectadas. No segundo subcapítulo, esta discussão foca-se no contexto Português.

1.5.1 Uma ameaça global – evolução e distribuição do AVC à escala mundial

Em termos de impacto e evolução da morbilidade e mortalidade, o Acidente Vascular Cerebral possui características similares, mas também algumas diferenças, face às doenças cardiovasculares em geral.

Estas doenças estiveram em destaque nas grandes mudanças de causas de morte que ocorreram (particularmente nas sociedades industrializadas) no decurso do século XX, na medida em que durante este período a sua importância foi aumentando progressivamente, ao ponto de se tornarem uma das principais causas de morte num grande número de países e no mundo como um todo. Este não é um processo exclusivo das sociedades industrializadas, sendo as implicações também importantes nos países em vias de desenvolvimento, nos quais o número de idosos aumenta progressivamente e as alterações sociais se sucedem, provocando mudanças nos padrões de ocorrência das doenças. Desta forma, para além da sua importância nos países de elevado rendimento, as doenças cardiovasculares têm vindo a apresentar nas últimas décadas proporções epidémicas em regiões do mundo com níveis baixos e médios de rendimento, no contexto da transição epidemiológica que nelas decorre (Labarthe, 2011). A figura 6 ilustra, já na década de 1980, a importância relativa destas doenças no mundo.

No que diz respeito especificamente ao AVC, a sua ocorrência é actualmente responsável por uma larga proporção da mortalidade e morbilidade global nas doenças cardiovasculares. Globalmente, o AVC representava, em 2010, a segunda causa mais comum de morte e a terceira causa mais frequente de DALYs (Feigin et al., 2014), constituindo uma doença que ocorre predominantemente em adultos de meia-idade ou idades superiores (WHO, 2014a). Em 2010, 16.9 milhões de pessoas sofreram pela primeira vez um AVC, e o número de óbitos relacionados com esta doença atingia já 5.9 milhões, tendo sido ainda responsável por 102 milhões de DALYs e pela existência de 33 milhões de pessoas sobreviventes a um episódio. Mais ainda, todos estes indicadores sofreram um aumento significativo desde 1990 (Feigin et al., 2014).

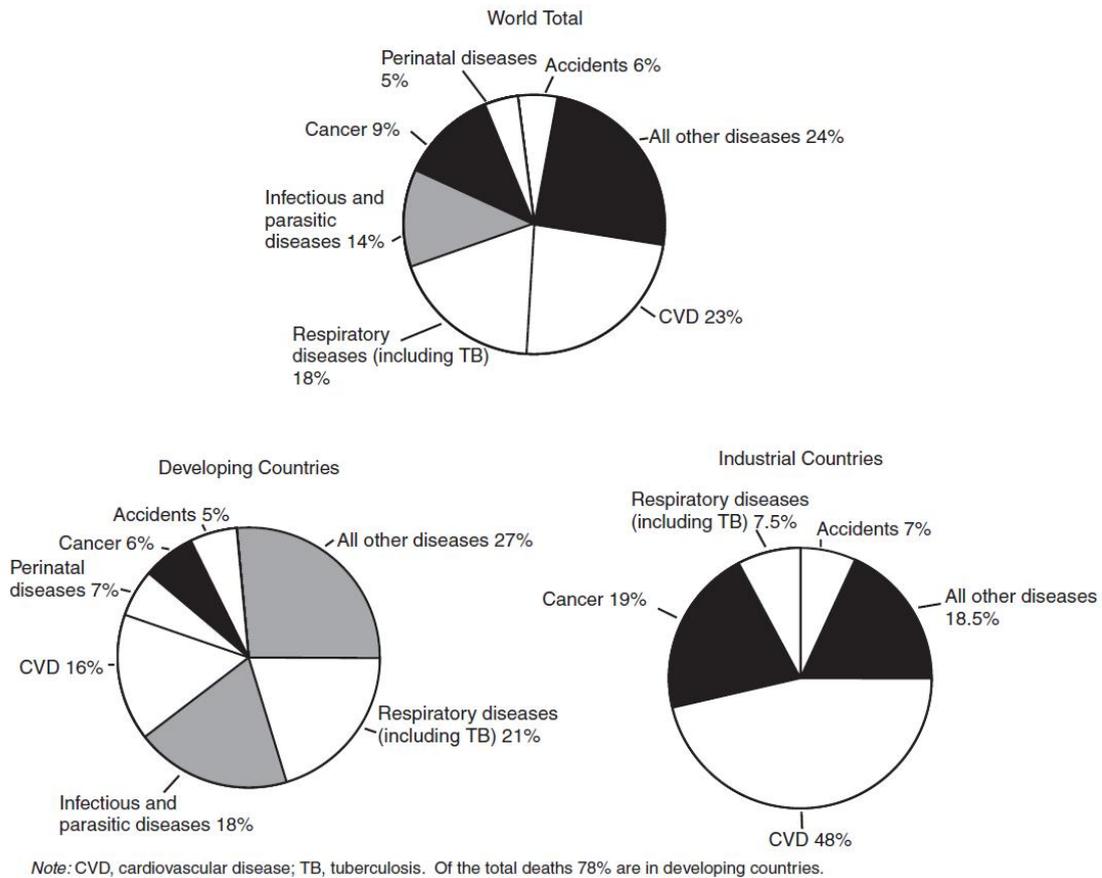


Figura 6 – Importância relativa dos óbitos por doenças cardiovasculares na mortalidade total em países em desenvolvimento, industrializados e no mundo, em 1980. Fonte: Labarthe (2011).

Todavia, para além dos valores globais, é quando se analisa a distribuição geográfica de vários indicadores entre países ou entre territórios pertencentes ao mesmo país que emerge um contexto complexo e heterogéneo.

Neste contexto, emergem duas grandes tendências preocupantes a vigiar atentamente: o aumento da importância do AVC em países de baixo e médio rendimento (e potencialmente em regiões mais deprimidas do mesmo país) e o impacto crescente dos casos ocorridos em pessoas mais jovens.

A este nível, o estudo *Global Burden of Diseases, Injuries, and Risk Factors Study (GBD)* inclui a informação disponível mais recente acerca da distribuição e tendências do AVC no mundo, permitindo estabelecer comparações entre 119 países (dos quais 58 integram os países de rendimento elevado e 61 constituem países de baixo e médio rendimento) no âmbito da mortalidade, incidência, prevalência e DALYs, entre outros indicadores. Estabelece o ponto de situação no ano de 2010 e as tendências de evolução

entre os anos de 1990 e 2010 para estes indicadores, tendo também em conta covariáveis relevantes como a idade das pessoas (Feigin et al., 2014).

No que respeita à incidência, a evolução das taxas padronizadas por idade nos países de rendimento elevado em geral foi no sentido do decréscimo, na medida em que entre 1990 e 2010 sofreram um decréscimo de 12%. Todavia, nos países de baixo e médio rendimento, a incidência sofreu um aumento de 12%. Embora no âmbito do GDB 2010, esta variação não tenha sido estatisticamente significativa, vem reforçar a ideia do aumento da importância deste fenómeno em países menos desenvolvidos, no âmbito da transição epidemiológica que neles tem vindo a decorrer nas décadas mais recentes. Quando analisada ao nível do país, a incidência apresentava em 2010 enormes variações geográficas nos casos por 100000 pessoas-ano (e.g. 60 casos no Kuwait contra 504 casos na Lituânia), o que poderia trazer alguma indeterminação a esta discussão. Todavia, quando se observa o número absoluto de casos incidentes, verifica-se que, de 16.9 milhões de episódios de AVC ocorridos em todo o mundo, a grande maioria (69%) ocorreram em países de baixo e médio rendimento.

Outra discussão relevante diz respeito à idade das pessoas afectadas. Se por um lado mais de 38% dos novos episódios ocorreram em pessoas com 75 anos de idade ou mais (destes, 50% em países de rendimento elevado e 32% nos de baixo e médio rendimento), os episódios ocorridos nas faixas etárias mais jovens começam a assumir proporções alarmantes, mais ainda se se tiver em conta o nível de rendimento dos países. Assim, entre 1990 e 2010 observou-se um aumento global e significativo de 25% da incidência em pessoas entre os 20 e os 64 anos, sendo que grande parte da contribuição para este facto se relaciona com o incremento de 18% verificado nos países de baixo e médio rendimento. Em 2010, as taxas padronizadas de incidência de AVC em países de baixo e médio rendimento excederam as verificadas em países de rendimento elevado em 23%. Por um lado, separando por idades, o excesso foi de 24% em pessoas abaixo dos 75 anos de idade e de 21% com 75 anos ou mais de idade. Mas apesar destes valores serem aparentemente próximos, na verdade o número de casos incidentes de AVC ocorridos em países de baixo e médio rendimento em pessoas com menos de 75 anos apresentou-se três vezes superior aos mesmos casos ocorridos em países de rendimento elevado. E mesmo nos países de rendimento elevado, o decréscimo nas taxas de incidência 1990-2010 foi menos pronunciado em pessoas

abaixo de 75 (9.1%) do que em pessoas de 75 ou mais anos de idade (16%). E se em 2010, os casos incidentes em pessoas abaixo de 64 anos (crianças e jovens com menos de 20 anos e também adultos jovens e de meia-idade entre os 20 e os 64 anos) representaram 31.5% de todos os casos, o peso mais uma vez foi superior em países de baixo e médio rendimento (35.6% dos casos) face aos de rendimento elevado (21.1% dos casos).

A mortalidade apresenta um cenário mais complexo. Por um lado, as taxas padronizadas por idade sofreram um decréscimo global significativo entre 1990 e 2010 de 25%, tanto em países de rendimento elevado (37%) como nos de baixo e médio rendimento (20%), embora este último decréscimo constitua um resultado inesperado e ainda pouco compreendido. Todavia, o número absoluto de mortes por AVC, pelo contrário, aumentou 26% no mesmo período (Feigin et al., 2014).

Tal como na incidência, também entre países a mortalidade apresentava em 2010 enormes variações geográficas nos óbitos por 100000 pessoas-ano, com as taxas padronizadas de mortalidade a variar entre 27 óbitos em França e 264 no Afeganistão, começando-se desde logo aqui a evidenciar o diferencial entre países com diferentes níveis de rendimento. Assim, de 5.9 milhões de óbitos ocorridos no mundo em 2010, 71% ocorreram em países de baixo e médio rendimento.

No que diz respeito à idade das pessoas afectadas, foi nas pessoas com 75 anos de idade ou mais que a variação no número absoluto de óbitos foi maior entre 1990 e 2010, com um aumento de 36%. Também em 2010, 55% das mortes ocorridas globalmente por AVC ocorreram em pessoas nesta faixa etária. Todavia, regionalmente, nos países de baixo e médio rendimento, os óbitos em pessoas com 75 anos de idade contribuíram com apenas 47% (menos de metade dos óbitos), enquanto nos países de rendimento elevado esta faixa etária assume mais importância representando 73% dos óbitos.

E não obstante o decréscimo global das taxas de mortalidade entre 1990 e 2010 já referido, mesmo assim o decréscimo global em pessoas mais jovens (36%) foi inferior ao verificado em pessoas mais idosas (39%).

Verifica-se também que as taxas padronizadas de mortalidade em pessoas com menos de 75 anos ocorridas em países de baixo e médio rendimento em 2010 excederam em 212% os valores dos países de rendimento elevado, enquanto a mesma comparação para pessoas de 75 ou mais anos de idade revelou um acréscimo de apenas 33%.

A reforçar esta situação, o rácio mortalidade/incidência em 2010 evidenciava um valor significativamente mais elevado (37%) em pessoas com menos de 75 anos de idade residentes em países de baixo e médio rendimento face ao mesmo grupo etário residente em países de rendimento elevado. A mesma comparação para o grupo etário de 75 anos ou mais de idade evidenciava apenas um valor 12% mais elevado em países de baixo e médio rendimento. Também em 2010, as mortes por AVC em pessoas com menos de 64 anos representavam uma maior percentagem de todas as mortes por AVC em países de baixo e médio rendimento (27%) do que em países de rendimento elevado (22.7%).

Assim, apesar do grande número de óbitos ocorridos em pessoas mais idosas residentes em países de rendimento elevado, também na mortalidade a magnitude dos óbitos registados nas faixas etárias mais jovens assume proporções alarmantes, particularmente quando se considera o nível de rendimento dos países.

Do ponto de vista dos DALYs, os resultados deste indicador apresentam similaridades com a incidência e mortalidade. Assim, este indicador apresentava também em 2010 enormes variações geográficas, desde 398 DALYs por 100000 pessoas na Austrália até 5227 no Afeganistão.

No que diz respeito aos efeitos nos escalões mais jovens, quando padronizado por idade, este indicador apresentava, em países de rendimento elevado, um decréscimo significativo de 36% em pessoas com menos de 75 anos contra 63% em pessoas com 75 ou mais anos de idade entre 1990 e 2010. Já quando avaliada em países de baixo e médio rendimento, o decréscimo apresenta-se inferior, como esperado, mas a diminuição em pessoas com menos de 75 anos (24%) é mais expressiva do que a ocorrida em pessoas com 75 ou mais anos de idade (17%), o que acresce ao contexto alguma complexidade.

Todavia, quando se observam os valores absolutos, a situação aparenta maior simplicidade. Assim, globalmente, em 2010 foram registados 102.2 milhões de DALYs, com 78% destes a ocorrer em países de baixo e médio rendimento. Mais ainda, 71.7% dos DALYs atribuíveis a AVC ocorreram em pessoas com menos de 75 anos de idade. Finalmente, a variação entre 1990 e 2010 evidencia uma tendência para o agravamento da situação, tendo-se verificado um aumento global do número de DALYs em 16% entre estes dois anos.

Em termos da prevalência, é de notar uma inversão da situação verificada nos indicadores anteriores. No caso específico desta doença, um caso prevalente corresponde a uma pessoa que sobrevive pelo menos 28 dias após um episódio de AVC. Desta forma, geograficamente os valores em 2010 oscilavam entre 82 sobreviventes de AVC por 100000 pessoas no Burundi e 1187 no Canadá. Em termos de tendência, a prevalência aumentou significativamente (27%) nos países de rendimento elevado entre 1990 e 2010 e apenas 8.5% nos países de baixo e médio rendimento (mas não de forma estatisticamente significativa). Quando ajustada à idade, observa-se também uma tendência para o aumento na variação, entre 1990 e 2010, na prevalência em países de rendimento elevado, similar tanto em pessoas com menos de 75 anos de idade (23%) como no grupo contrário (25%). Curiosamente, em países de baixo e médio rendimento, a tendência foi muito semelhante, com aumentos de 9% nos mais jovens e de 10% nos mais idosos, embora não significativa (Feigin et al., 2014).

Adicionalmente, coloca-se a questão de avaliar separadamente a situação e tendências dos casos de AVC hemorrágico e isquémico. Nesse sentido, o conhecimento mais actual indica que, quando analisados em separado, ambos os tipos apresentam na maioria dos indicadores, tendências similares ao AVC quando analisado como um todo, à excepção da incidência. Assim, é de salientar um aumento significativo de 22% nos casos de AVC hemorrágico em países de baixo e médio rendimento entre 1990 e 2010, quando comparado com um aumento não significativo de apenas 6% nos casos de AVC isquémico nesses mesmos países (Krishnamurthi et al., 2013).

Após a observação das diferenças e similaridades entre os diversos indicadores, confirma-se o padrão complexo e heterogéneo já referido. Mas na interpretação deste padrão, as diversas peças encaixam-se, complementando-se entre si, tal como num “puzzle” acabado de construir.

Desta forma, a tendência emergente dos sucessivos estudos GBD realizados em 1990, 2001 e 2010 aponta para o aumento dos valores absolutos do AVC em todo o mundo, mas particularmente em países de baixo e médio rendimento. Embora os valores dos indicadores ajustados para a idade tenham vindo a decrescer, quando se observa a evolução entre 1990 e 2010 dos números absolutos de pessoas afectadas, de pessoas que sobreviveram a um episódio de AVC e vivem agora com as sequelas resultantes, de DALYs e, pior ainda, de mortes relacionadas por AVC, torna-se claro que o peso global

desta doença é grande e com tendência para aumentar, maioritariamente nos países mais desfavorecidos. (Feigin et al., 2014) apontam mesmo uma previsão, referindo que a persistência destas tendências poderá levar a que em 2030 se verifiquem globalmente cerca de 12 milhões de mortes, 70 milhões de casos prevalentes e mais de 200 milhões de DALYs devido a esta doença.

Vários factores contribuem para que as diferenças no peso do AVC em países de diferentes níveis de rendimento sejam substanciais. Nos países de rendimento elevado, os motivos mais prováveis para a diminuição da incidência, mortalidade, rácio mortalidade-incidência e DALYs relacionam-se com a melhor prestação dos serviços de saúde e das estratégias para a prevenção do AVC e assistência a pessoas afectadas, ao contrário do que acontece nos países de baixo e médio rendimento. Por outro lado, neste último grupo de países, também o aumento da prevalência dos factores de risco de AVC (e.g. do tabagismo, entre outros) constitui uma preocupação, na medida em que provocará a continuação da tendência para o aumento do peso desta doença nestes países.

Outra situação aparentemente controversa tem a ver com a significativamente maior prevalência que se verifica em países de rendimento elevado, quanto comparados com os países menos desenvolvidos. Este facto parece relacionar-se com a existência de uma associação inversa entre prevalência *versus* mortalidade e DALYs entre países pertencentes a diferentes classes de rendimento. A explicação para esta situação pode ter a ver com diferenças na gestão clínica dos episódios de AVC agudo e no acesso aos cuidados de saúde. Desta forma, o maior número de episódios fatais e consequentemente de valores inferiores de prevalência pode constituir o reflexo da resposta por vezes menos adequada dos sistemas de saúde em países de baixo e médio rendimento.

Em suma, fenómenos como o envelhecimento da população global e seu progressivo aumento, a redução dos casos fatais e a transição epidemiológica nos países de baixo e médio rendimento, que leva ao aumento da prevalência dos factores de risco, constituem as causas mais prováveis para o aumento do número de casos incidentes, da prevalência e dos DALYs.

O aumento do impacto do AVC nos grupos etários mais jovens (em particular nos países de baixo e médio rendimento) e as razões que poderão explicar este fenómeno merecem também alguma atenção.

Historicamente, o AVC foi sempre considerado como uma doença limitada a pessoas mais idosas. Esta realidade tem tendência a alterar-se, pois como anteriormente referido, os dados mais recentes indicam que a proporção do peso desta doença é já superior em pessoas com menos de 75 anos face ao grupo mais idoso, particularmente nos países mais desfavorecidos. E não obstante a idade média das pessoas com AVC estar a aumentar em países de todos os tipos de rendimento (presumivelmente devido ao envelhecimento da população), a proporção de episódios desta doença em pessoas abaixo dos 65 anos de idade é muito relevante e tem vindo a aumentar, mais uma vez particularmente em países de baixo e médio rendimento.

Vários factores podem contribuir para explicar esta situação. O aumento global e em particular nos países de baixo e médio rendimento da prevalência de diabetes e de outros factores de risco de doença cardiovascular em jovens (e na população em geral) deverá originar previsivelmente um deslocamento gradual do peso desta doença para as faixas mais jovens da população. Assim, as evidências existentes sugerem que o AVC poderá deixar de ser considerado uma doença típica dos idosos (Feigin et al., 2014).

1.5.2 Uma ameaça nacional – o AVC no contexto Português

O AVC constitui um problema *major* de Saúde Pública em Portugal. Não obstante, não é um problema frequentemente invocado nos meios de comunicação social, e alguns autores não hesitam em referir-se a este problema como uma “epidemia silenciosa”. Esta situação terá provavelmente a ver com a sua associação, no entendimento popular, ao quadro das doenças das pessoas idosas (Sá, 2009).

No que diz respeito à mortalidade, e apesar da sua escassa expressão mediática, o AVC constitui uma das principais causas de morte em Portugal, tendo sido em 2011 responsável por 12690 óbitos e por uma taxa padronizada de mortalidade de 61,9 óbitos por 100000 habitantes (DGS, 2013). Mesmo dentro do quadro das doenças cardiovasculares em geral, constitui a mais expressiva, na medida em que a Doença Isquémica do Coração (DIC) foi em 2011 responsável por pouco mais de metade do

número de óbitos (6852) e da taxa de mortalidade padronizada, com 34,9 óbitos por 100000 habitantes.

Em termos de tendências, Portugal (figura 7) tem seguido ao longo das últimas décadas a tendência europeia para a diminuição progressiva das taxas de mortalidade por AVC (figura 8), como aliás se verifica nas doenças cardiovasculares em geral (DGS, 2013).



Figura 7 – Evolução da taxa de mortalidade padronizada, por doenças cerebrovasculares, por 100000 habitantes, entre 1971 e 2009, em Portugal. Fonte: HFA-DB (2011).

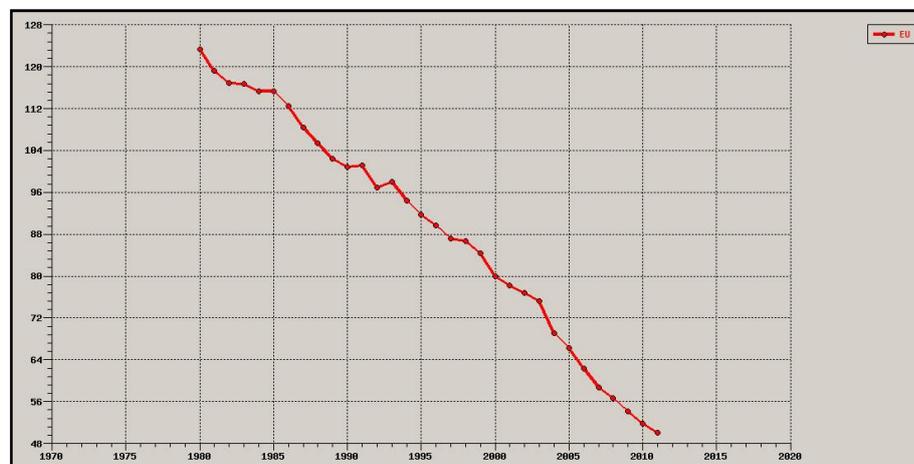


Figura 8 – Evolução da taxa de mortalidade padronizada, por doenças cerebrovasculares, por 100000 habitantes, entre 1980 e 2011, nos países da União Europeia. Fonte: HFA-DB (2011).

Observando-se num intervalo temporal mais curto e próximo da actualidade, entre 2002 e 2012 a taxa de mortalidade padronizada por AVC sofreu uma diminuição progressiva, embora essa diminuição sofra algum abrandamento entre os anos de 2011 e 2012 (Figura 9).

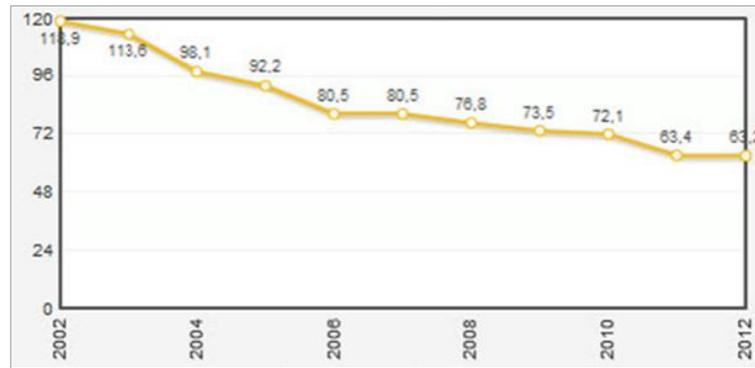


Figura 9 – Evolução da taxa de mortalidade padronizada, por doenças cerebrovasculares em Portugal, por 100 000 habitantes (2002-2012). Fonte: INE (2014).

Esta diminuição foi também uma realidade no que diz respeito ao valor absoluto de óbitos ocorridos, tendo o número total de óbitos diminuído em mais de 1300 casos por ano entre 2007 e 2011 (figura 10).

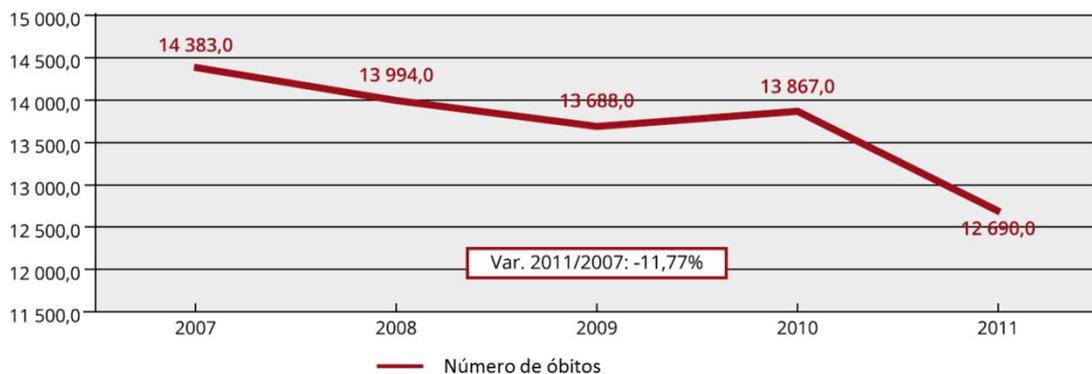


Figura 10 – Número de óbitos por doenças cerebrovasculares em Portugal Continental (2007-2011). Fonte: DGS (2013).

Por outro lado, a taxa de mortalidade padronizada por AVC em Portugal permanece entre as mais altas de entre os países da Europa, mesmo quando comparada com alguns outros países da bacia mediterrânica, sendo que em 2010 apenas a Grécia e a Bulgária apresentavam valores mais elevados (tabela 2). Também apenas Portugal, Grécia e Bulgária apresentavam valores da taxa de mortalidade padronizada por AVC superiores à taxa correspondente por DIC. É de notar igualmente a existência de grandes diferenças entre valores nacionais, com tendência para um gradiente Sul/Oeste - Norte/Leste que ocorre nas regiões europeias (DGS, 2013).

Tabela 2 – Comparação das taxas de mortalidade padronizadas por Doenças Cérebro-cardiovasculares, por 100 000 habitantes, em Portugal e alguns países Europeus em 2010.

	Doenças do Aparelho Circulatório	Doença Isquémica do Coração	Doença Cerebrovascular
Portugal	169,29	39,87	71,07
Espanha	137,58	43,75	33,75
França	118,94	32,06	26,12
Chipre	186,58	62,91	36,16
Itália	167,70	57,37	45,51
Grécia	228,92	62,06	67,84
Alemanha	208,71	80,87	35,86
Reino Unido	164,19	77,25	42,11
Polónia	335,83	90,31	68,50
Finlândia	213,57	120,70	42,95
Bulgária	621,69	114,26	178,62
Letónia	477,62	248,88	131,86

Fonte: (DGS, 2013)

Já no que diz respeito à incidência, este constitui, ao contrário da mortalidade, um indicador para o qual não existem muitas vezes dados disponíveis que permitam uma visão abrangente e completa. Por exemplo, nos Estados Unidos da América, não é conhecida a incidência de AVC no país como um todo, tanto para a doença em geral como para os seus subtipos (Labarthe, 2011). No caso português, uma aproximação possível e disponível consiste em recorrer aos episódios de internamento hospitalar, recolhidos no âmbito da base de dados dos Grupos de Diagnósticos Homogéneos (ACSS, 2014). Assim, a análise da evolução da contagem anual dos utentes saídos dos episódios de internamento entre os anos de 2000 e 2012 parece também ilustrar uma tendência para a diminuição progressiva, embora com flutuações, registando-se o valor mais elevado em 2001 e o mais baixo em 2007 (figura 11). As estimativas da incidência calculadas no âmbito do estudo *Global Burden of Disease* (Feigin et al., 2014) parecem apoiar esta ideia, revelando similaridade com os valores reais.

Assim, estas estimativas apontavam para a ocorrência de cerca de 32129 casos em 1990 (IC 95%: 27025-37833), em 2005 de 33559 casos (IC 95%: 28137-39800) e em 2010 de 37062 casos (IC 95%: 31272-43848). Desta forma, coloca-se a hipótese do número de casos incidentes poder estar a aumentar, não obstante a mortalidade se encontrar em diminuição.

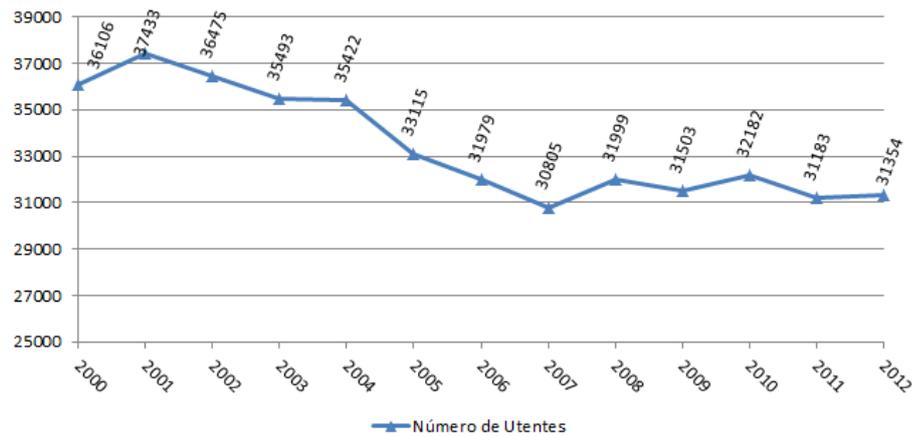


Figura 11 – Evolução do número de utentes saídos dos episódios de internamento por AVC, em Portugal Continental, entre 2000 e 2012. Dados: ACSS (2014).

As estimativas deste estudo para o rácio mortalidade-incidência parecem apoiar esta tendência, apresentando valores de 0.90 (IC 95%: 073-1.09) em 1990, 0.58 (IC 95%: 043-0.82) em 2005 e de 0.54 (IC 95%: 041-0.78). Também o número estimado de casos prevalentes neste estudo aponta uma tendência para o aumento, com 46279 casos (IC 95%: 27258-74080) em 1990, 79610 (IC 95%: 48923-122260) em 2005 e 86892 (IC 95%: 51190-139852) em 2010, o que poderia derivar do aumento das possibilidades de sobrevivência de um crescente número de pessoas que constituem um caso incidente desta doença, até porque a maior parte das pessoas afectadas consegue sobreviver à fase aguda desta doença (Sá, 2009).

Outra vertente relevante para a análise do peso desta doença tem a ver com o impacto que esta doença possui na qualidade e duração de vida das pessoas afectadas. Tendo em conta que, apesar dos valores elevados de mortalidade associados a esta doença, a maior parte das pessoas afectadas consegue sobreviver à fase aguda, o AVC acaba por constituir também, em Portugal, uma das principais causas de dependência e incapacidade (Sá, 2009). Assim, em 2011 registaram-se mais anos de vida perdidos ajustados pela incapacidade (DALYs) por AVC (14990) do que por DIC (14900). Embora neste indicador a diferença seja pouco expressiva, a tendência, tal como na mortalidade, tem vindo a ser para a diminuição dos valores, como se observa na sua evolução entre 2006 e 2010 (Figura 12).



Figura 12 – Anos de vida perdidos ajustados pela incapacidade, por doenças cerebrovasculares, em Portugal (2006-2010). Fonte: DGS (2013).

A análise da evolução da taxa de DALYs entre 2002 e 2012 apresenta um quadro semelhante, embora seja já de notar uma ligeira inflexão da curva entre 2010 e 2012 (figura 13).

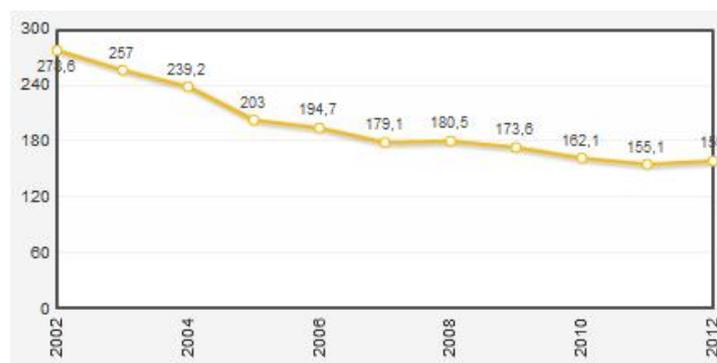


Figura 13 – Taxa de anos de vida perdidos ajustados pela incapacidade, por doenças cerebrovasculares, em Portugal (2002-2012). Fonte: INE (2014).

Passando ao ponto de situação e tendências em idosos e não idosos, e começando pelos idosos, verifica-se pela evolução das taxas de mortalidade padronizadas em pessoas com 65 ou mais anos ou 70 ou mais anos (figura 14), que existe uma tendência para a diminuição, tal como ocorre com a taxa para todas as idades. Na comparação com as variações das taxas incluídas na figura 15, nota-se que, embora a tendência para a diminuição das taxas de mortalidade padronizadas se verifique em idosos e não idosos, esta tendência tem mais expressão na população idosa, particularmente acima dos 70 anos. Esta diferença, embora ténue, pode indiciar um aumento da relevância da ocorrência de AVC junto da população abaixo de 65 anos de idade e a necessidade, tal

como já indicado para o nível global, de vigiar epidemiologicamente e intervir também neste coorte populacional.



Figura 14 – Taxa de mortalidade padronizada por doença cerebrovascular em pessoas de idade igual ou superior a 65 ou 70 anos, por 100 000 habitantes, em Portugal Continental (2007-2011).
Fonte: DGS (2013).

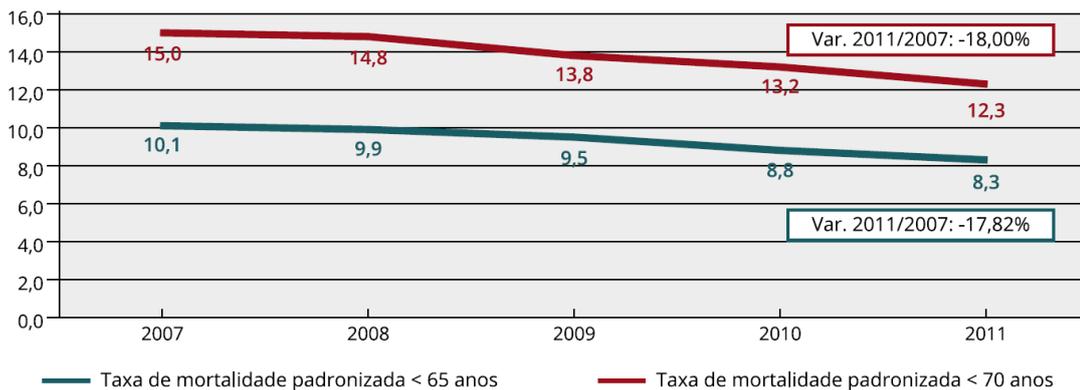


Figura 15 – Taxa de mortalidade padronizada por doença cerebrovascular em pessoas de idade inferior a 65 ou 70 anos, por 100 000 habitantes, em Portugal Continental (2007-2011). Fonte: DGS (2013).

Na análise do peso do AVC na Saúde Pública, é também relevante ter-se em conta o impacto diferencial em homens e mulheres. Assim, este impacto, quando medido em DALYs, parece ser superior nos homens, embora a tendência geral em ambos os sexos seja também para o decréscimo (figura 16). Todavia, é de notar uma ligeira inversão desta tendência, com um acréscimo entre 2011 e 2012 que, embora ainda pouco expressivo, sugere que seja observado atentamente o evoluir deste indicador em anos subsequentes. A evolução da taxa de DALYs para o mesmo período ilustra uma tendência semelhante (figura 17).

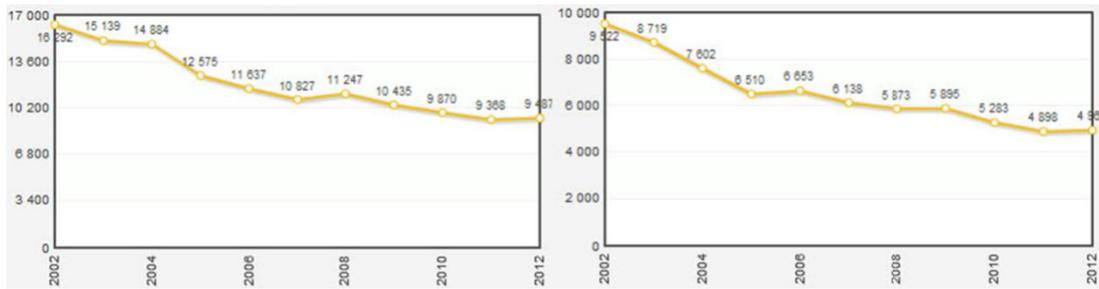


Figura 16 – Anos de vida perdidos ajustados pela incapacidade por doenças cerebrovasculares entre 2002 e 2012, em homens (esquerda) e mulheres (direita), em Portugal. Fonte: INE (2014).

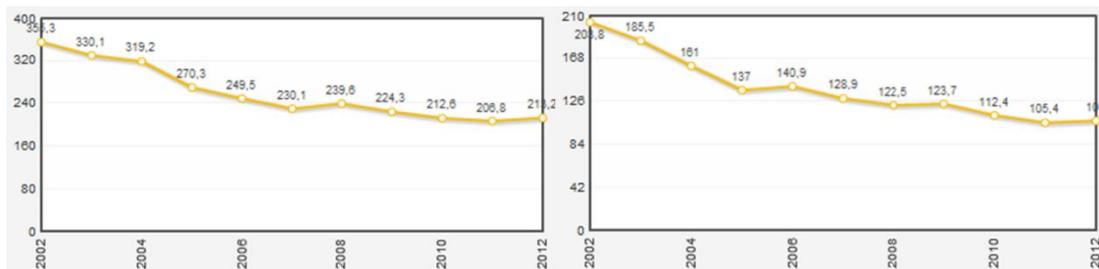


Figura 17 – Taxa de anos de vida perdidos ajustados pela incapacidade por doenças cerebrovasculares entre 2002 e 2012, em homens (esquerda) e mulheres (direita), em Portugal. Fonte: INE (2014).

A tendência na taxa de mortalidade padronizada ilustra um diferencial entre sexos, pois enquanto esta taxa nas mulheres mantém a propensão para a diminuição, já no caso dos homens nota-se também uma inversão entre 2011 e 2012, com um aumento do valor da taxa, embora também ligeiro (figura 18).

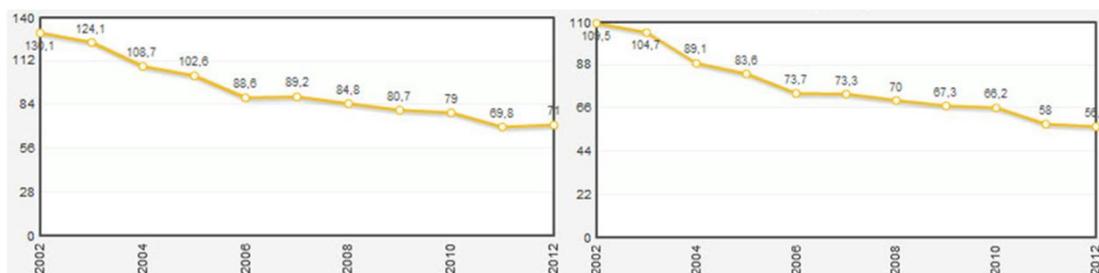


Figura 18 – Taxa de mortalidade padronizada por doenças cerebrovasculares por 100 000 habitantes entre 2002 e 2012, em homens (esquerda) e mulheres (direita). Fonte: INE (2014).

O desdobramento da taxa de mortalidade padronizada em homens e mulheres idosos ou não, entre 2009 e 2012, revela a mesma ligeira tendência para a inversão entre 2011 e 2012 nos homens, tanto na classe etária de 65 e mais anos de idade (figura 19) como nas pessoas com menos de 65 anos de idade (Figura 20).

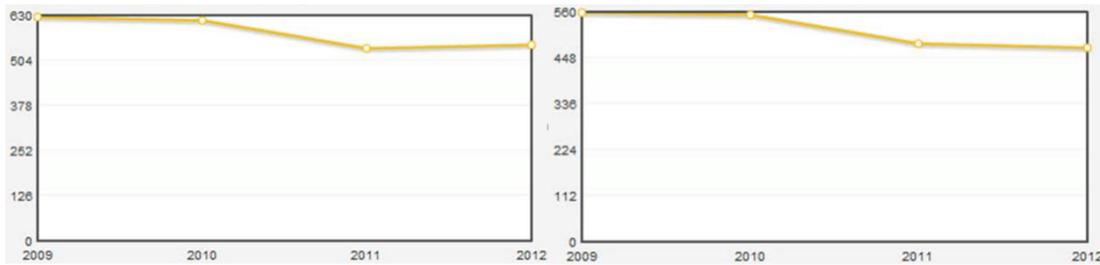


Figura 19 – Taxa de mortalidade padronizada por doenças cerebrovasculares por 100 000 habitantes entre 2009 e 2012, em homens (esquerda) e mulheres (direita) com 65 ou mais anos de idade.
Fonte: INE (2014).

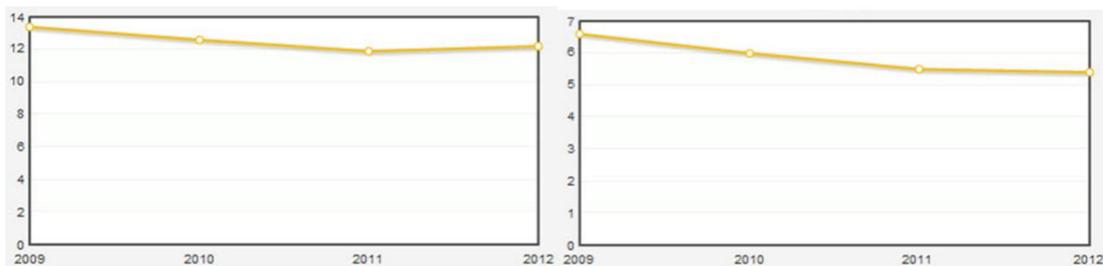


Figura 20 – Taxa de mortalidade padronizada por doenças cerebrovasculares por 100 000 habitantes entre 2009 e 2012, em homens (esquerda) e mulheres (direita) com menos de 65 anos de idade.
Fonte: INE (2014).

Observando-se o diferencial entre homens e mulheres na evolução dos utentes saídos dos episódios de internamento, esta aproximação à incidência apresenta tendências similares, ocorrendo em ambos os sexos uma tendência geral para o decréscimo até 2007 e relativa estabilização a partir desse ano, com a mesma ligeira inversão entre 2011 e 2012 já referida noutros indicadores, para as mulheres (figura 21).



Figura 21 – Evolução do número de utentes saídos dos episódios de internamento por AVC, por sexos, entre 2000 e 2012. Dados: ACSS (2014).

A segmentação deste indicador por escalões etários e sexos mostra que o ligeiro aumento verificado no período 2011-2012 nas mulheres se deve visivelmente ao contributo da faixa etária com 65 ou mais anos (figura 22), ao passo que nos homens, a ligeira diminuição verificada se deve à faixa etária com menos de 65 anos (figura 23).

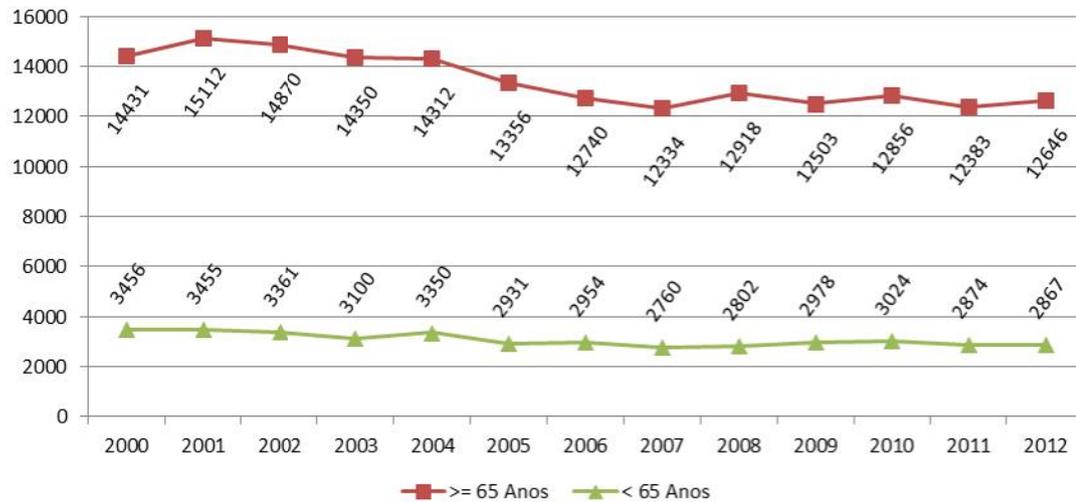


Figura 22 – Número de utentes saídos dos episódios de internamento por AVC, em mulheres, em dois escalões etários, entre 2000 e 2012. Dados: ACSS (2014).

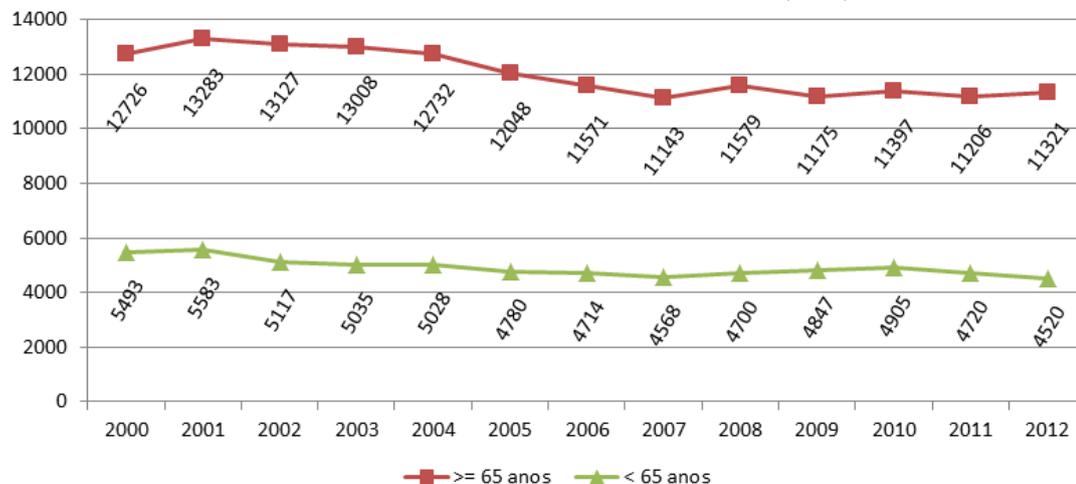


Figura 23 – Número de utentes saídos dos episódios de internamento por AVC, em homens, em dois escalões etários, entre 2000 e 2012. Dados: ACSS (2014).

Poderão assim existir tendências inversas entre, por um lado a mortalidade e os DALYs, e por outro lado a incidência. A diminuição da mortalidade por esta doença tem sido atribuída a vários factores actuando em conjunto, com destaque para os avanços ocorridos no âmbito da intervenção durante a fase clínica do AVC, não apenas nos

fármacos e técnicas inovadoras empregues (tais como a terapêutica fibrinolítica) mas particularmente na vertente organizativa, sendo o exemplo mais expressivo a implementação das Vias Verdes do AVC. Assim, a implementação destas estratégias organizativas veio permitir uma resposta precoce em termos de emergência pré-hospitalar, o correcto encaminhamento para as unidades de saúde que administram os melhores fármacos e tratamentos e ainda o reforço operacional dos meios existentes ao longo do território. A diminuição da mortalidade tem também sido atribuída a um outro conjunto de medidas e estratégias, estas de carácter preventivo, tais como a Lei de Cessação Tabágica, a legislação visando a redução do conteúdo de sal no pão, as campanhas promovidas por várias organizações com o objectivo de fomentar a adesão a estilos de vida saudáveis e ainda os aperfeiçoamentos ocorridos ao nível do diagnóstico e rectificação de factores de risco modificáveis tais como a hipertensão arterial ou a dislipidémia (DGS, 2013).

Por outro lado, na hipótese da tendência estimada para o aumento dos casos incidentes se revelar real, ficaria talvez menos clara a relação entre a incidência e as medidas de carácter preventivo atrás referidas.

Alguns autores, como referido anteriormente, chamam também a atenção para a existência de disparidades não apenas entre países mas por vezes também entre as regiões que compõem um país. Isto chama a atenção para a necessidade de se medirem as disparidades regionais com o objectivo de priorizar intervenções nas regiões onde a situação é mais preocupante. Esta situação é uma realidade no caso de Portugal Continental, quando se observa a distribuição regional dos vários indicadores do AVC. Nesta análise, essencialmente visual, optou-se pelo recurso a mapas coropletos com uma escala de valores discreta (cada região tem associado o seu valor, apenas arredondado à unidade), e a gradação de cores escolhida representa, mais do que a expressão da intensidade dos valores, um *ranking* das regiões face aos vários indicadores analisados. Esta nota cautelar destina-se a evitar interpretações incorrectas da realidade, ou seja “mentir com mapas” (Monmonier,1996), na medida em que os valores se encontram pontualmente muito próximos entre regiões, mas também porque o enfoque essencial desta observação reside em detectar as disparidades regionais existentes nos indicadores em cada ano cartografado.

Começando pela taxa padronizada de mortalidade, todos os sexos e idades, em 2012 verificava-se a existência de uma disparidade norte-sul, com as taxas mais elevadas a ocorrerem na NUTII do Norte e as mais reduzidas no Algarve, embora os valores não estejam muito distanciados do valor nacional (63,2 óbitos por 100000 habitantes). A tendência para as disparidades geográficas é observável também em 2002 e 2007, não obstante a diminuição geral dos valores da taxa de 2002 para 2012. Notam-se algumas flutuações regionais ao longo dos anos, entre as quais o aumento em termos relativos na NUTII do norte e a tendência inversa na NUTII do Algarve (figura 24).

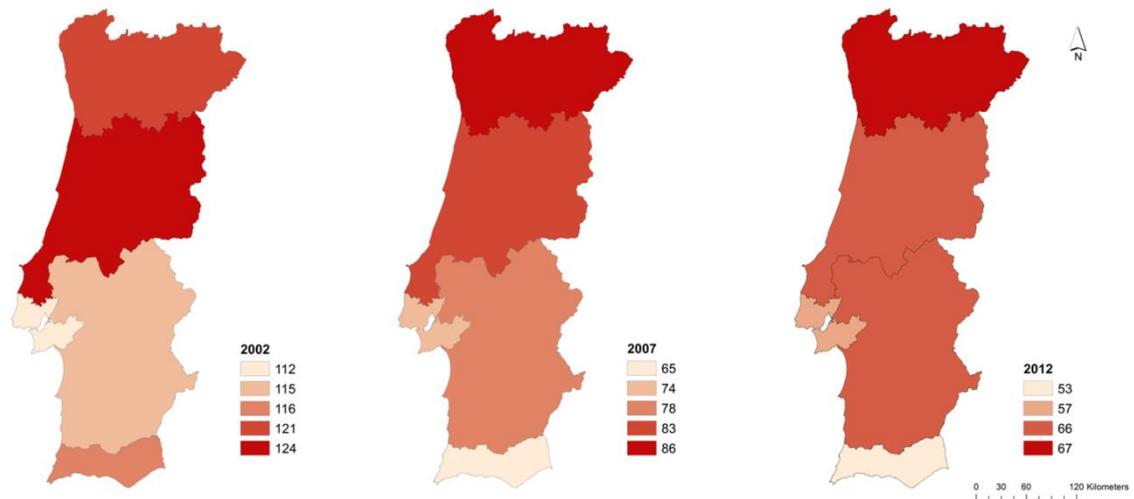


Figura 24 – Taxa de mortalidade padronizada por AVC em Portugal Continental em 2002, 2007 e 2012, por 100 000 habitantes (ambos os sexos). Dados: INE.

Segmentando-se a análise por sexos, todas as idades, observam-se tendências similares, tanto na situação em 2012 como na evolução dos valores regionais da taxa de mortalidade, com disparidades regionais Norte-Sul na taxa calculada para homens (figura 25) e para mulheres (figura 26).

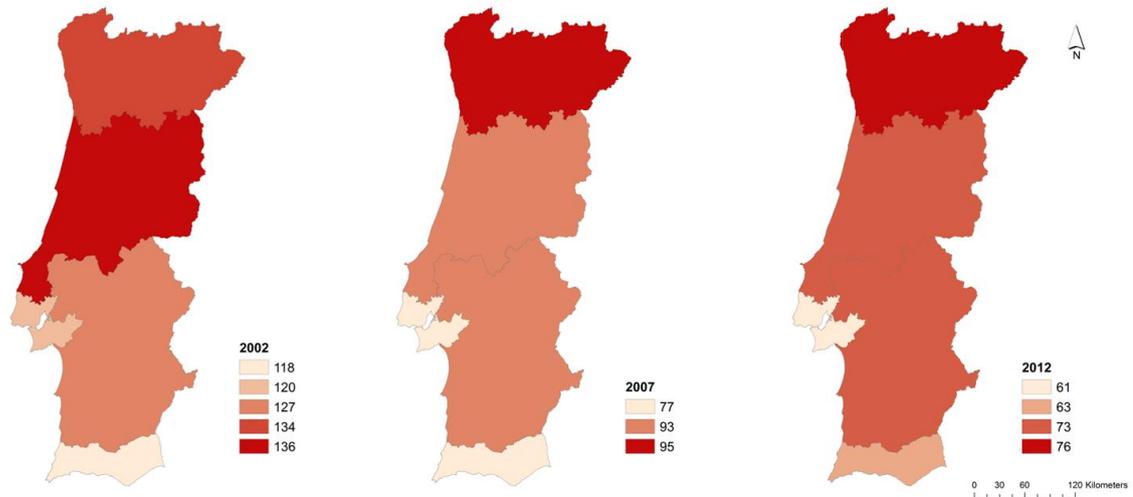


Figura 25 – Taxa de mortalidade padronizada por AVC em Portugal Continental em 2002, 2007 e 2012, por 100 000 habitantes (homens). Dados: INE.

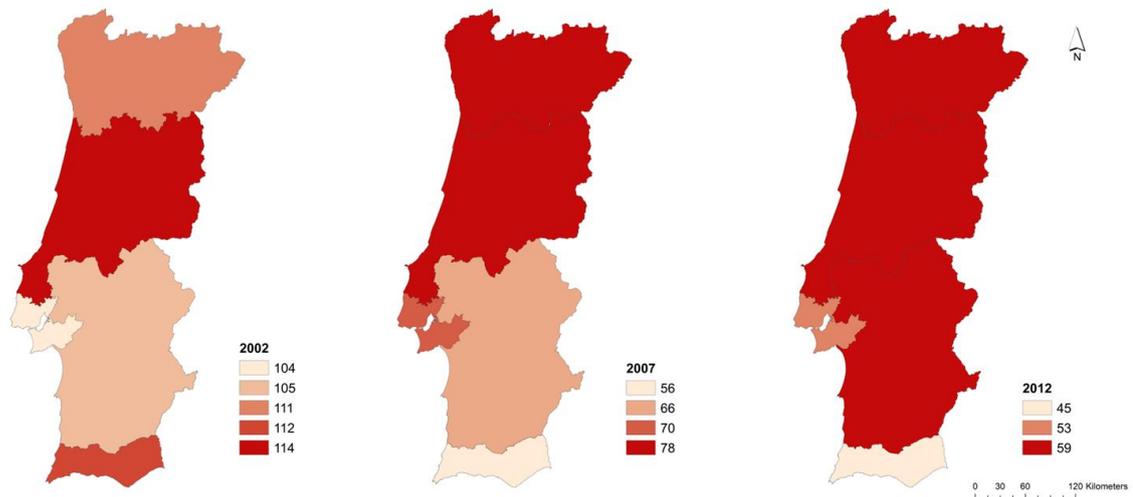


Figura 26 – Taxa de mortalidade padronizada por AVC em Portugal Continental em 2002, 2007 e 2012, por 100 000 habitantes (mulheres). Dados: INE.

Segmentando agora a análise da taxa de mortalidade por dois grandes grupos etários (ambos os sexos) em 2009, 2010 e 2012, verificam-se as mesmas diferenças relativas na variação Norte-Sul no grupo com 65 ou mais anos de idade, e também uma grande estabilidade nos padrões regionais da taxa de mortalidade (figura 27).

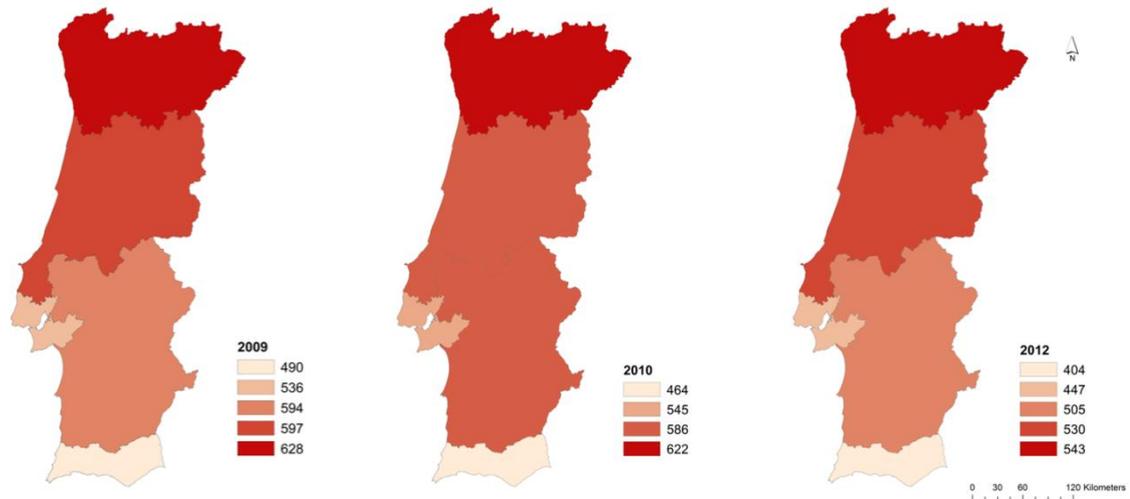


Figura 27 – Taxa de mortalidade padronizada por AVC em Portugal Continental em 2009, 2010 e 2012, ambos os sexos, 65 ou mais anos de idade, por 100 000 habitantes. Dados: INE.

Todavia, a mesma análise, quando efectuada para o grupo etário com menos de 65 anos de idade, revela já um padrão de distribuição regional diferente, configurando-se praticamente a inversão da tendência anteriormente discutida, com valores ligeiramente mais elevados da taxa a ocorrerem mais para Sul e com maior estabilidade dos valores em 2009 e 2012 (figura 28).

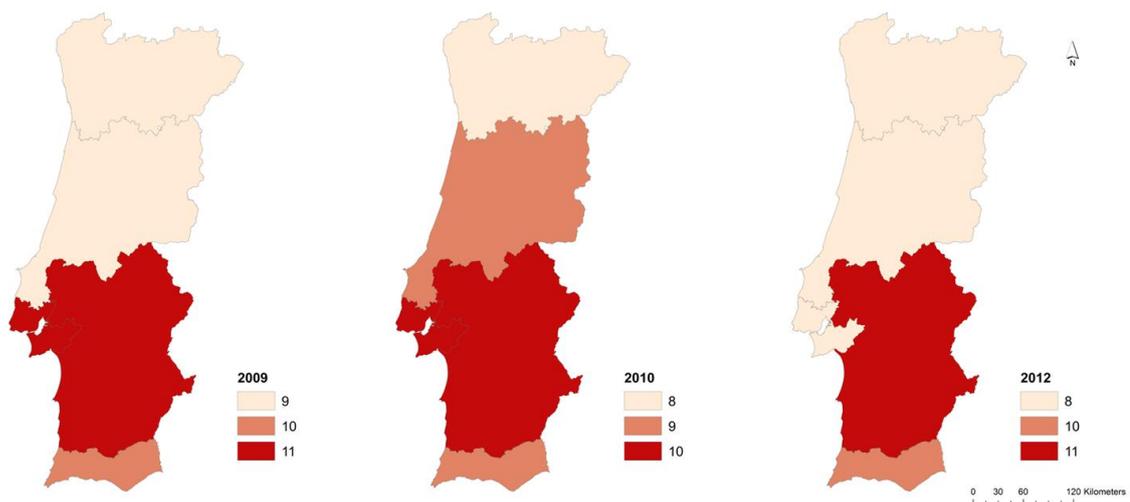


Figura 28 – Taxa de mortalidade padronizada por AVC em Portugal Continental em 2009, 2010 e 2012, ambos os sexos, menos de 65 anos de idade, por 100 000 habitantes. Dados: INE.

Segmentando ainda mais, por grupo etário e sexo, a tendência nos homens com 65 ou mais anos de idade mantém-se muito similar entre 2009 e 2012, voltando a distribuição

regional a evidenciar valores mais altos a Norte (figura 29). Também na distribuição regional da taxa em mulheres para este grupo etário os padrões se revelaram muito similares (figura 30).

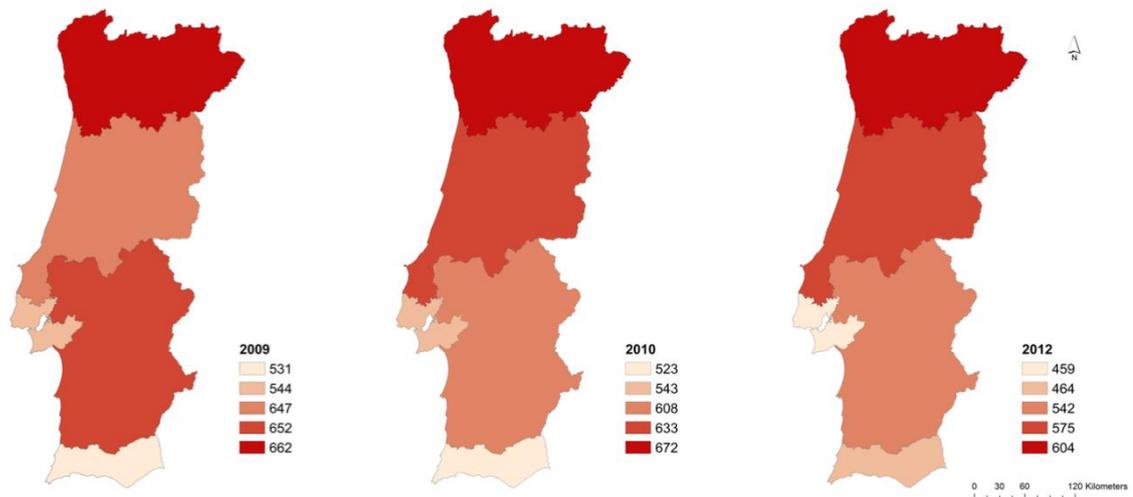


Figura 29 – Taxa de mortalidade padronizada por AVC em Portugal Continental em 2009, 2010 e 2012, homens com 65 ou mais anos de idade, por 100 000 habitantes. Dados: INE.

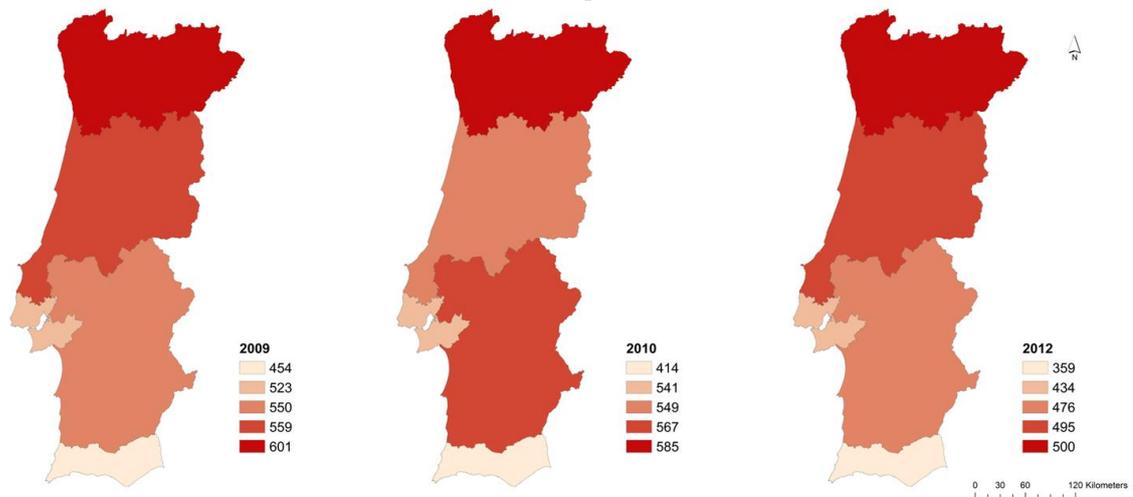


Figura 30 – Taxa de mortalidade padronizada por AVC em Portugal Continental em 2009, 2010 e 2012, mulheres com 65 ou mais anos de idade, por 100 000 habitantes. Dados: INE.

Quanto às tendências da distribuição regional observadas em pessoas com menos de 65 anos de idade, segmentadas por sexo, estas verificam-se semelhantes ao que já tinha sido observado para ambos os sexos mais atrás, surgindo mais uma vez a tendência geral para a ocorrência dos valores mais altos a Sul, nos homens particularmente no Alentejo, nas mulheres destacando-se o Algarve (figuras 31 e 32).

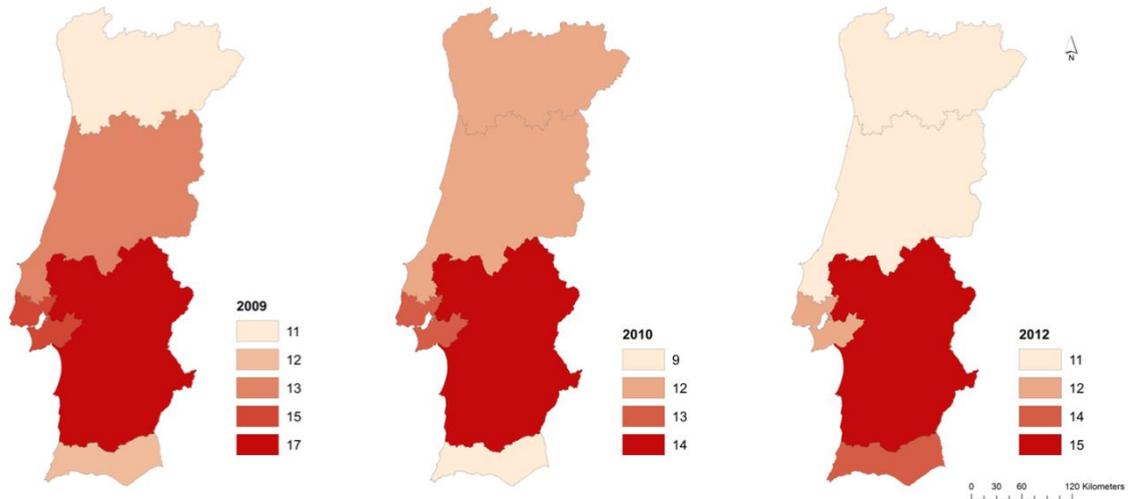


Figura 31 – Taxa de mortalidade padronizada por AVC em Portugal Continental em 2009, 2010 e 2012, homens com menos de 65 anos de idade, por 100 000 habitantes. Dados: INE.

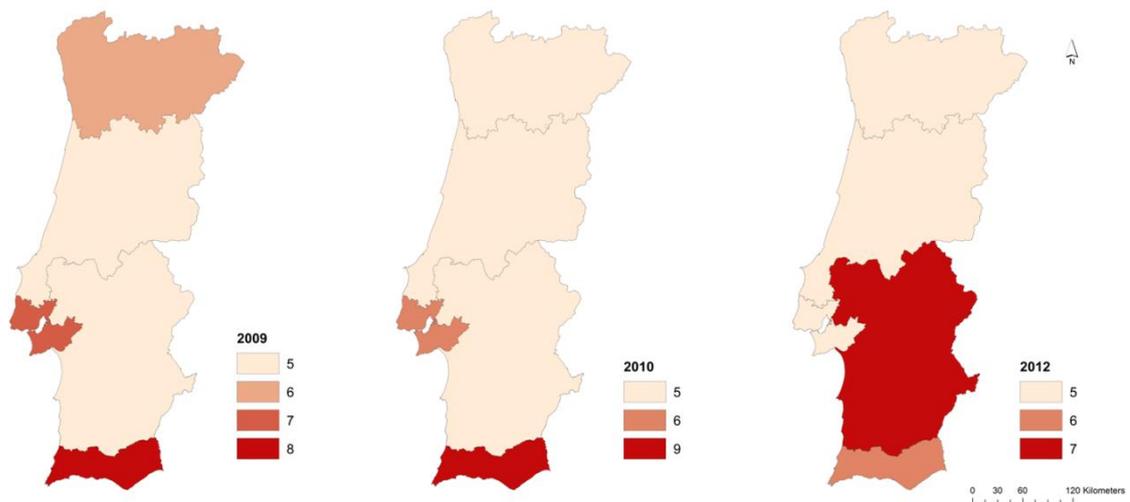


Figura 32 – Taxa de mortalidade padronizada por AVC em Portugal Continental em 2009, 2010 e 2012, mulheres com menos de 65 anos de idade, por 100 000 habitantes. Dados: INE.

Procedendo-se à análise dos padrões regionais dos DALYs (ambos os sexos, todas as idades), surge já uma grande estabilidade no padrão geográfico ao longo dos anos, mantendo-se os valores por NUTII no sentido da sua progressiva diminuição entre 2002 e 2012, bem como a tendência para valores mais altos no Norte do território (Figura 33).

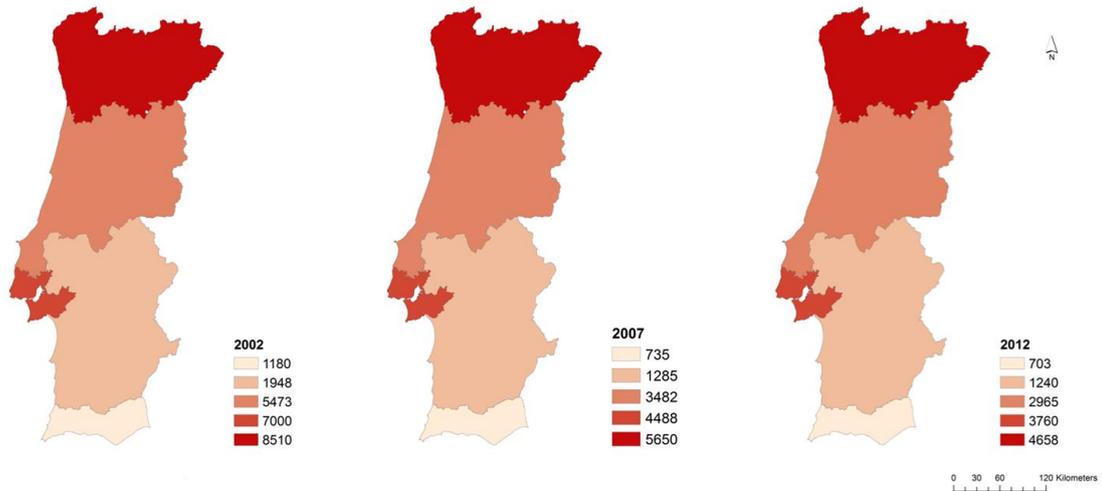


Figura 33 – Anos de vida perdidos ajustados pela incapacidade por doenças cerebrovasculares em 2002, 2007 e 2012, todas as idades, ambos os sexos. Dados: INE.

Passando à análise dos DALYs por sexo (todas as idades), mais uma vez as tendências nos padrões de distribuição regional em homens (figura 34) e mulheres (figura 35) se revelam muito similares aos valores observados para ambos os sexos, mantendo-se as tendências de evolução e de distribuição Norte-Sul, sendo apenas de notar os valores, em geral mais baixos, registados no sexo feminino.

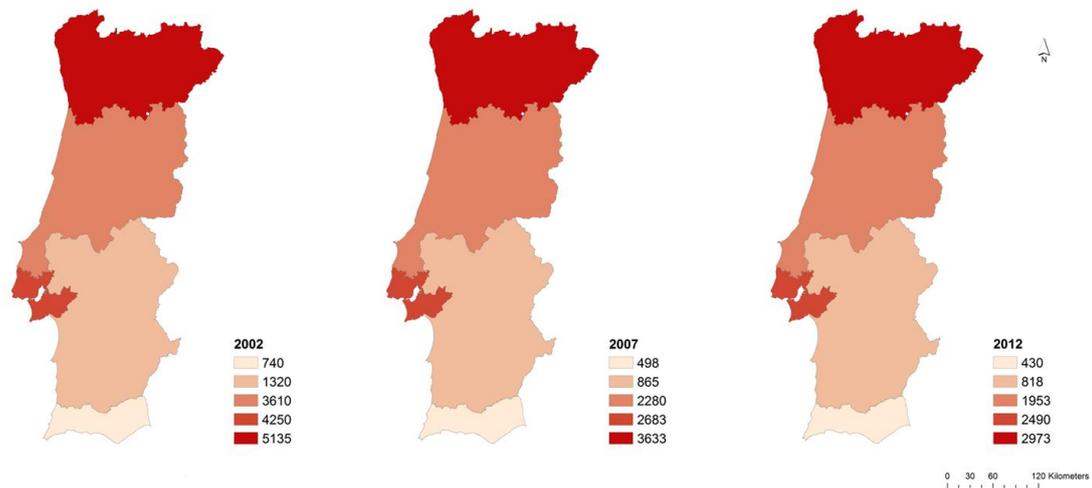


Figura 34 – Anos de vida perdidos ajustados pela incapacidade por doenças cerebrovasculares em 2002, 2007 e 2012, em homens, todas as idades. Dados: INE.

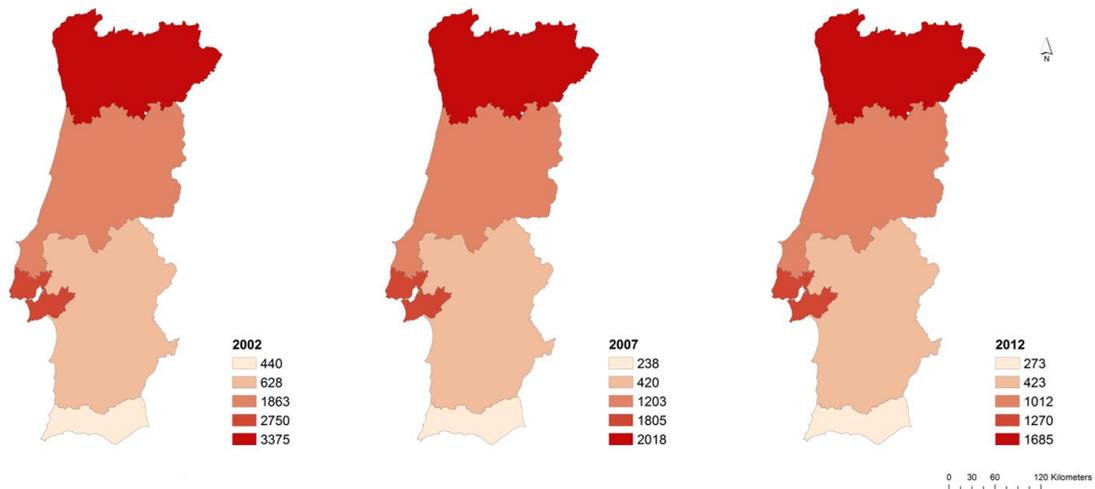


Figura 35 – Anos de vida perdidos ajustados pela incapacidade por doenças cerebrovasculares em 2002, 2007 e 2012, em mulheres, todas as idades. Dados: INE.

Todavia, já na taxa de DALYs (ambos os sexos, todas as idades), verifica-se de novo uma inversão da tendência Norte-Sul, com os valores mais elevados a surgirem no Sul, particularmente Alentejo e Algarve, e valores mais baixos no Norte (figura 36).

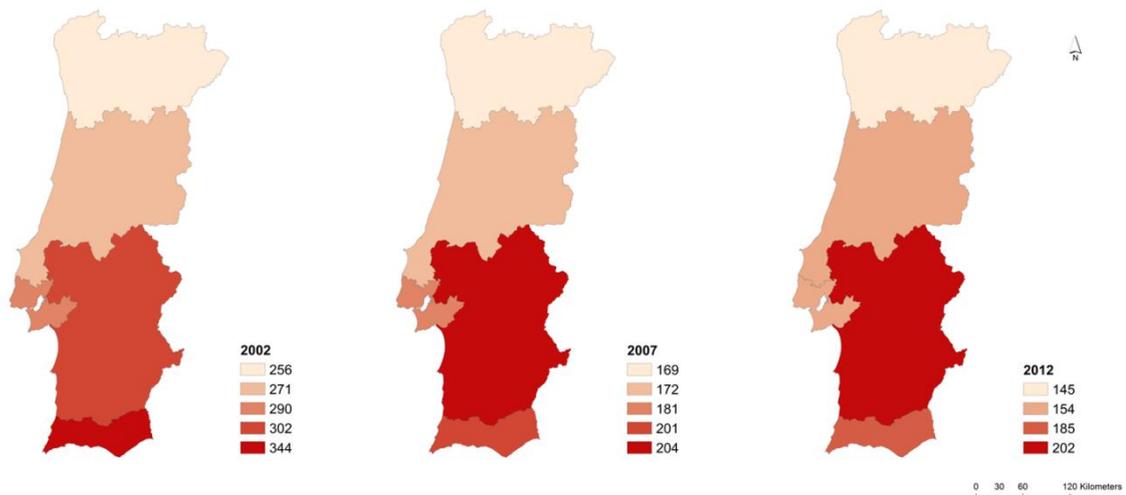


Figura 36 – Taxa de vida perdidos ajustados pela incapacidade por doenças cerebrovasculares em 2002, 2007 e 2012, todas as idades, ambos os sexos. Dados: INE.

Também quando separada por sexos, se mantém na taxa de DALYs, a mesma tendência geral de distribuição Norte-Sul em homens (figura 37) e mulheres (figura 38), embora com maiores variações regionais no caso feminino, ao longo dos três anos representados.

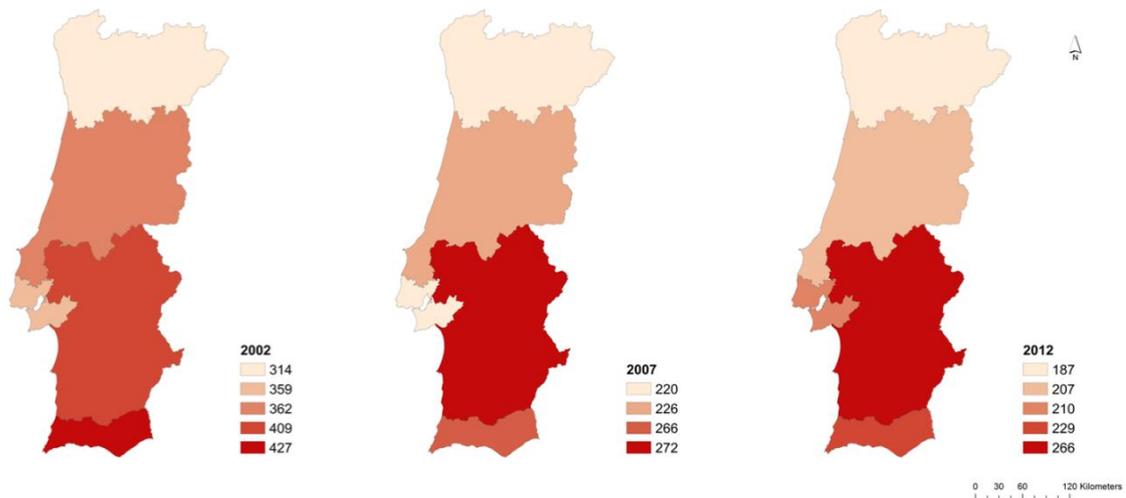


Figura 37 – Taxa de anos de vida perdidos ajustados pela incapacidade por doenças cerebrovasculares em 2002, 2007 e 2012, homens, todas as idades. Dados: INE.

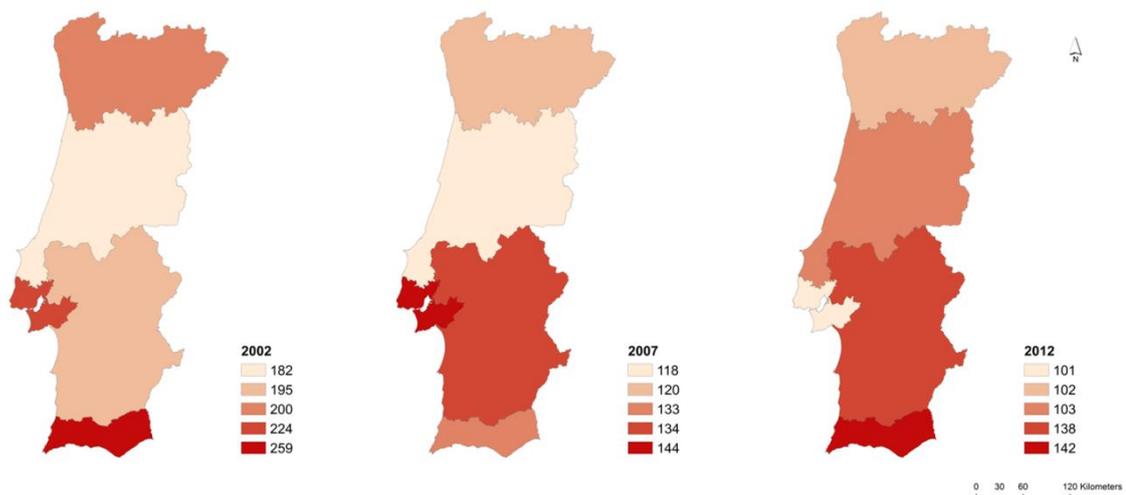


Figura 38 – Taxa de anos de vida perdidos ajustados pela incapacidade por doenças cerebrovasculares em 2002, 2007 e 2012, mulheres, todas as idades. Dados: INE.

Em suma, Portugal segue, de forma geral, a tendência europeia para o decréscimo da mortalidade por AVC, como se observou aqui pela análise da evolução das taxas padronizadas e dos valores absolutos, e também dos DALYs e respectiva taxa padronizada. Todavia, notou-se, no período 2002-2012, abrangido por grande parte das análises efectuadas, a existência de ligeiras inversões da tendência no período 2011-2012, particularmente nas taxas padronizadas de mortalidade e nos DALYs (tanto em valor absoluto como em taxa), cuja evolução a partir de 2012 poderá merecer acompanhamento, pois coloca-se a hipótese de constituírem já um reflexo da presente

crise económica. A evolução dos utentes saídos dos episódios de internamento parece sofrer também uma ligeira subida neste período, embora, tal como em outros indicadores já referidos, seja ainda ténue. As diferenças mais expressivas revelam-se na distribuição regional dos diversos indicadores, com especial relevo para as taxas padronizadas de mortalidade, as quais apresentam persistentemente valores mais elevados nas regiões a norte de Portugal Continente, na maior parte dos casos, com algumas excepções pontuais, não obstante a evolução dos valores regionais tenha em geral seguido a tendência do país e da Europa.

1.6 Determinantes do Acidente Vascular Cerebral

A primeira questão que se levanta consiste em definir o que é um determinante da Saúde. Segundo Porta (2008), um determinante pode ser entendido como “Qualquer factor que desencadeia mudança numa condição de saúde ou outra característica definida”. Esta definição constitui um ponto de partida mas é ainda muito abrangente, impondo-se ainda a necessidade de classificar os determinantes existentes. Ahrens e Pigeot (2004) indicam que os estados de Saúde podem ser influenciados por factores comportamentais, culturais, sociais, psicológicos, biológicos e físicos, sendo todavia esta lista ainda bastante genérica. A este propósito, a Agência de Saúde do Canadá (Health Canada, 2012) fornece uma lista mais extensa, incluindo já doze determinantes da Saúde, nomeadamente rendimento e condição social, emprego, educação, ambientes sociais, ambientes físicos, desenvolvimento saudável da criança, práticas de saúde pessoais e capacidade de lidar com as situações (*coping skills*), serviços de saúde, redes sociais de apoio, herança biológica e genética, sexo, cultura. Outras classificações são mais parcimoniosas, como a do Grupo de Avaliação de Impactos na Saúde da Organização Mundial da Saúde (WHO, 2014b), que começa por agrupar os determinantes em três grandes tipos: ambiente económico e social, ambiente físico, características e comportamentos individuais da pessoa.

Independentemente da classificação escolhida, é também importante determinar qual a importância relativa dos vários determinantes, as vias através das quais eles podem afectar o nível de saúde das pessoas, e as relações entre eles mesmos (e.g. até que ponto

um determinante pode levar à ocorrência de outro), uma vez que a resposta a estas questões influencia geralmente a tomada de decisões por parte das autoridades de Saúde, decisões essas que implicam priorizar intervenções e alocação de recursos que são quase sempre escassos. É de notar também que os determinantes podem ser medidos a um nível individual ou ecológico (com agregação de dados individuais, por exemplo ao nível das NUTII anteriormente referidas ou do município), considerando-se aqui que os resultados se aplicam também à população correspondente ao nível de agregação. Neste sentido, um estudo ecológico (ou agregado) foca-se na comparação entre grupos ao invés de indivíduos (Rothman et al., 2008).

Feita esta distinção, é importante salientar que os determinantes devem ser devidamente enquadrados numa cadeia de causalidade que decorre entre determinante e doença. Neste sentido, Marmot e Wilkinson (2005) começam por segmentar os factores de risco em comportamentos e marcadores biológicos. Estes autores referem, ao discutir o impacto dos factores socioeconómicos na Saúde, que a Epidemiologia se concentra mais frequentemente na identificação e detecção dos factores de risco individuais da doença e que, mais ainda, no âmbito destes factores, se foca geralmente mais na identificação dos marcadores biológicos (e.g. a presença de pressão arterial elevada, associada a um maior risco de doença cardiovascular) do que na detecção dos comportamentos que os antecedem numa cadeia de causalidade. Estes autores afirmam que, numa perspectiva de Saúde Pública, é mais importante encontrar os determinantes dos marcadores biológicos do que os próprios marcadores. Estes últimos são sobretudo relevantes na actuação junto do indivíduo, pois um tratamento com o objectivo de diminuir a hipertensão arterial de uma determinada pessoa levará à diminuição do risco de doença nessa pessoa apenas. Neste contexto, os factores de risco individuais deste tipo dividem-se em modificáveis, tais como o hábito de fumar, ou não modificáveis, como por exemplo a idade de uma pessoa.

Marmot e Wilkinson (2005) salientam assim, numa perspectiva mais ligada à Saúde Pública, a importância do enfoque na “causa das causas”, ou seja, da necessidade de se investigar e actuar junto das fontes de uma cadeia que pode ser longa e complexa e que se desenrola de “montante” para “jusante” ao longo do “rio” da causalidade. Exemplificando, uma situação social em que uma pessoa se encontra sujeita a baixos níveis de rendimento e a forte insegurança laboral pode levar ao desenvolvimento do

hábito de fumar, o qual por sua vez pode determinar a presença de um marcador biológico (por exemplo o aumento da pressão arterial), que por sua vez poderá levar ao desenvolvimento de doença cardiovascular.

Labarthe (2011) sumariza esta cadeia de causalidade para o caso particular das doenças cardiovasculares na actualidade (figura 39). Segundo este autor, esta cadeia inicia-se numa mistura de situações ambientais e sociais desfavoráveis que despoletam padrões de comportamento nocivos. Estes comportamentos (em particular as dietas desequilibradas e a inactividade física) levarão, por sua vez, ao desenvolvimento dos principais factores de risco do tipo clínico, tais como a hipertensão arterial, obesidade, diabetes, tabagismo e presença de níveis elevados de colesterol na circulação sanguínea, entre outros.

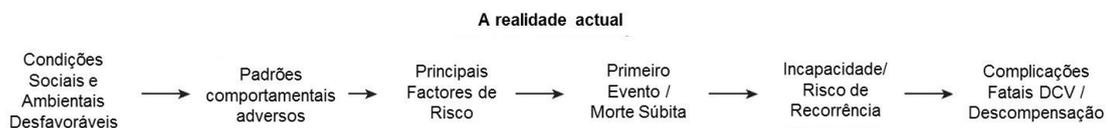


Figura 39 – Cadeia actual de causalidade em doenças cardiovasculares.
Fonte: Adaptado de Labarthe (2011).

Na ausência de medidas de controlo eficazes destes factores de risco, ocorrerá um primeiro evento, o qual resultará em morte súbita ou sobrevivência, esta última muitas vezes acompanhada de incapacidade e de um alto risco de reincidência, sendo que estes pacientes possuem uma alta probabilidade de vir a falecer de complicações cardiovasculares ou descompensação relacionada.

Falando estritamente dos determinantes da Saúde relacionados com a estrutura económica e social, tanto Labarthe (2011) como Marmot e Wilkinson (2005) referem, entre outros, o impacto na Saúde (e em particular nas doenças cardiovasculares) de factores sociais como a ocupação profissional, rendimento e nível educacional, e a alteração temporal das condições sociais (por migrações, mobilidade cultural, entre outros). Exemplificando, uma ocupação profissional de mais baixa remuneração e qualificações pode ser considerada como marca de classe ou estatuto social e relacionada com a ausência de actividade física e aumento dos níveis de *stress*. Por estarem muito ligados à situação social da família, estes autores discutem também condições como o baixo peso à nascença e o deficiente desenvolvimento corporal até à

adolescência, que podem ser considerados precursores de possíveis complicações cardiovasculares em adulto.

No que diz respeito aos determinantes ligados ao ambiente físico, Labarthe (2011) refere-se a factores ambientais em sentido mais “restrito”, como a exposição passiva ao fumo de tabaco ou a partículas atmosféricas poluentes, referindo alguns estudos recentes que evidenciam o impacto negativo da exposição a partículas com diâmetro suficientemente pequeno para atingir vias respiratórias e alvéolos. Estas partículas podem estar disseminadas por grandes áreas ou podem concentrar-se em áreas industriais (onde existam, por exemplo, fundições) ou nos “canhões” urbanos de grandes cidades. Os efeitos destas partículas repercutem-se no sistema cardiovascular pelo facto de os pulmões deixarem de funcionar correctamente, como acontece nas inflamações sistémicas. Por outro lado, este autor refere também factores ambientais em sentido mais “lato”. Um exemplo particularmente relevante tem ver com a distribuição geográfica das unidades hospitalares face à população. Segundo este autor, em 1999, cerca de 48% de todas as mortes por AVC nos EUA ocorreram antes do transporte para o hospital ou durante a chegada a um serviço de urgência hospitalar, sendo esta ocorrência menos frequente em pacientes jovens, e duas vezes mais frequente em casos de AVC isquémico (23% dos casos) do que em situações de AVC hemorrágico (12 a 14% dos casos).

1.6.1 A abordagem clínica aos factores de risco individual de AVC

A discussão iniciada no subcapítulo anterior, embora utilizando exemplos oriundos das doenças cardiovasculares, destinou-se a introduzir os determinantes da Saúde em geral. O presente subcapítulo inicia o debate acerca dos determinantes do AVC, focando-se em particular nos factores de risco individual do tipo clínico, os quais, como referido anteriormente, vão englobar tanto características como comportamentos de cada indivíduo. Apesar do diagnóstico e avaliação destes factores em cada indivíduo ser algo que se encontra no domínio da prática clínica, o seu conhecimento é indispensável também na prática de Saúde Pública, na medida em que constituem os factores imediatamente precursores da ocorrência de acidente vascular cerebral. E isto porque a mitigação destes factores antes de poderem contribuir para o despoletar de doença

cerebrovascular pode constituir também, em última análise, um alvo para as intervenções de Saúde Pública.

Dada a origem multifactorial desta doença, têm sido enumerados em diversos estudos, muitos destes factores de risco, tanto putativos como confirmados (Brainin e Heiss, 2009). Adicionalmente, indicam que estes factores têm também sido classificados por força de evidência, consoante a sua contribuição para o AVC se encontre mais ou menos bem documentada.

Na discussão da importância dos diversos factores de risco, é aqui também referida, quando oportuno, a diferente importância de alguns destes factores nas duas grandes variantes desta doença, o AVC isquémico e hemorrágico (dentro desta segunda variante, é ainda relevante a divisão em hemorragia intracraniana e hemorragia subaracnóidea) pois não obstante a maior parte dos factores contribuir de forma similar para ambos os casos, existem algumas situações em que tal não se verifica.

Sacco et al. (1997) classificam estes factores de risco como modificáveis, potencialmente modificáveis e não modificáveis. Nestes últimos, e apesar do seu carácter imutável, a sua avaliação não deixa de ser importante, pois mesmo que não seja possível tomar medidas no sentido da sua eliminação, a simples detecção da sua presença contribui para a identificação das pessoas em situação de elevado risco, sustentando o recurso urgente a tratamentos no sentido de reduzir os factores de risco modificáveis.

Entre os factores não-modificáveis de risco de AVC mais importantes, contam-se a idade da pessoa, sexo, grupo étnico, factores genéticos e baixo peso à nascença, este último factor comum às doenças cardiovasculares em geral. Entre estes, a idade representa o factor com maior grau de importância, pois em cada 10 anos após os 55 anos de idade, o risco deste tipo de complicação duplica em pessoas de ambos os sexos (Brainin e Heiss, 2009); (ESO, 2013). Todavia, os casos não se encontram apenas concentrados nas faixas etárias mais idosas. Por exemplo, a contribuição deste factor difere um pouco no caso específico do AVC hemorrágico, na medida em que a incidência de hemorragia intracraniana parece aumentar de forma mais moderada com a idade e a faixa etária de maior risco de hemorragia subaracnóide se situa entre os 45 e os 55 anos de idade (Brainin e Heiss, 2009); (Sacco et al., 1997).

No que diz respeito à importância do sexo, a pertença ao sexo masculino acarreta um risco ligeiramente superior. A este respeito, Sacco et al. (1997) referem que as taxas de incidência são 1.25 maiores em homens do que em mulheres. Todavia, sendo que as mulheres têm tendência a viver mais anos do que os homens, o número de mulheres que morrem por AVC por ano acaba por ser superior ao dos homens. Brainin e Heiss (2009) referem também que a ocorrência de AVC é mais frequente em homens na faixa etária entre os 45 e os 84 anos.

Já no que diz respeito ao grupo étnico, Sacco et al. (1997) indicam que a incidência desta doença parece ser menor em indivíduos caucasianos do que em indivíduos não-caucasianos (por exemplo os indivíduos afro-americanos possuem um risco de falecer devido a esta doença mais de duas vezes superior ao dos indivíduos caucasianos). Todavia, a discussão destes autores leva a entender que estas associações, muito baseadas em estudos realizados nos Estados Unidos da América, poderão estar confundidas pela influência de desigualdades socioeconómicas que se verificam sistematicamente entre grupos étnicos. Brainin e Heiss (2009) reforçam este ponto de vista, referindo também estas associações no contexto norte-americano, mas indicando, por outro lado, que a mortalidade por AVC ao nível global não parece seguir nenhum padrão étnico.

No que diz respeito ao historial familiar, Sacco et al. (1997) indicam que este é um factor conhecido há já algum tempo. Entre as potenciais causas apontadas encontram-se a predisposição genética, tanto para o AVC como para os seus factores de risco ou a exposição de toda a família a um determinado factor ambiental ou estilo de vida, ou ainda a interacção entre factores genéticos e ambientais (Brainin e Heiss, 2009). Entre alguns exemplos, Welin et al. (1987) referem um aumento do risco em homens cujas mães tinham falecido desta doença e em mulheres que possuíam já um historial familiar de AVC. Também no âmbito do estudo de Framingham, um dos mais relevantes na área das doenças cardiovasculares, Kiely et al. (1993) associam o registo parental de casos de AVC ou de doença coronária (tanto nos progenitores como nas progenitoras) ao aumento do risco de AVC.

Alguns factores de risco de AVC modificáveis encontram-se bem documentados. Entre estes encontram-se a hipertensão, o tabagismo, a diabetes, a fibrilhação atrial e outras doenças cardíacas, a dislipidémia, a estenose das artérias carótidas, a drepanocitose, a

terapia hormonal pós-menopausa, a prática de uma alimentação com base numa dieta pobre, a inatividade física, a obesidade, bem como a acumulação de gorduras na parte central do corpo.

Por outro lado, entre os factores de risco menos bem documentados ou potencialmente modificáveis encontram-se a síndrome metabólica, o consumo de bebidas alcoólicas em quantidades elevadas, o abuso de substâncias estupefacientes, a utilização de contraceptivos orais, a apneia do sono, as dores de cabeça provocadas por enxaquecas, a hiperhomocisteinemia, lipoproteínas elevadas, as lipoproteínas elevadas associadas a fosfolipase, a hipercoagulabilidade, a presença de inflamação e a presença de infecção (Brainin e Heiss, 2009).

Tendo em conta a multiplicidade de factores referidos, é importante relembrar a natureza multifactorial desta doença, no sentido de que grande parte das vezes a sua ocorrência não é consequência de um só factor de risco individual mas sim da presença e interacção de diversos destes factores (Brainin e Heiss, 2009). Desta forma, os factores modificáveis acima indicados podem ainda ser divididos, por um lado, naqueles que se relacionam com estilos de vida, e por outro, em doenças que muitas vezes derivam desses mesmos estilos de vida, reforçando a ideia de uma complexa rede de factores em interacção.

Neste sentido, muitos estudos têm abordado a associação entre a prevalência de doenças cerebrovasculares e os factores relacionados com os estilos de vida, tais como tabagismo, exercício físico, índice de massa corporal e consumo de álcool. Um exemplo consistiu no uso de dados de dois estudos de coorte de grande magnitude (Estudo da Saúde de Enfermeiras englobando 71234 mulheres; Estudo de Acompanhamento dos Profissionais de Saúde englobando 43685 homens) para definir um índice de baixo risco baseado na adopção de cinco estilos de vida, nomeadamente não fumar, praticar exercício físico moderado (30 minutos ou mais por dia), manter uma dieta saudável, manter um índice de massa corporal igual ou inferior a 25 kg/m^2 e consumir quantidades moderadas de álcool (5 a 30 g/dia para homens e 5 a 15 g/dia em mulheres). Este estudo sinalizou uma associação significativa entre a presença simultânea destes factores e uma redução significativa do risco de AVC em homens e mulheres. Desta forma, Brainin e Heiss (2009) concluem que as modificações no estilo de vida têm o potencial de prevenir o desenvolvimento de certas doenças que

constituem por sua vez factores de risco de AVC tais como a hipertensão, diabetes, dislipidémia e obesidade.

Apesar da multiplicidade de factores apresentados, cerca de 60 a 80% dos casos de AVC isquémico podem ser atribuídos a hipertensão, dislipidémia, tabagismo e diabetes, e ainda a fibrilação atrial e doenças nas válvulas cardíacas (nos casos de AVC isquémico embólico cardiogénico). Por este motivo, e também por simplicidade do texto, a discussão subsequente restringe-se a uma discussão conjunta destes factores modificáveis e bem documentados, podendo uma discussão mais extensa e aprofundada ser encontrada em Sacco et al. (1997) e em Brainin e Heiss (2009).

Começando pela hipertensão, esta representa para a maior parte dos autores o factor de risco modificável mais importante em pessoas de meia-idade e idosas, tanto para o AVC isquémico como para o hemorrágico, tendo a sua presença sido verificada em aproximadamente 70% dos casos (ESO, 2013). Representa igualmente o factor de risco passível de tratamento mais bem documentado (Brainin e Heiss, 2009). Tanto em homens como em mulheres, o risco desta doença parece aumentar proporcionalmente ao aumento da pressão arterial, quase duplicando com um aumento de 7.5 mm Hg na pressão sanguínea diastólica (Collins e MacMahon, 1994). Adicionalmente, a presença de uma pressão sanguínea sistólica superior ou igual a 160 mm Hg e/ou pressão sanguínea diastólica superior ou igual a 95 mm Hg acarreta um risco relativo de AVC aproximadamente 4 vezes superior. Por outro lado, não obstante a grande importância deste factor, incluindo nas pessoas idosas, o impacto da hipertensão parece decrescer com a idade, sendo o rácio de possibilidade (*Odds Ratio*) de 4 aos 50 anos de idade mas decrescendo para 1 aos 90 anos de idade (Sacco et al., 1997); (Brainin e Heiss, 2009).

O tabagismo constitui também um dos maiores factores de risco modificáveis tanto para o AVC isquémico como para o hemorrágico, devendo-se a este factor cerca de 8% das doenças cerebrovasculares (Brainin e Heiss, 2009) e existindo uma diferença de risco relativo em fumadores e ex-fumadores, quando comparada com não-fumadores, de 1.5 e 1.17, respectivamente. Também o número de cigarros fumados por dia parece aumentar proporcionalmente o risco de AVC, sendo que, na comparação entre homens e mulheres, a magnitude desta relação é superior em mulheres. É de salientar que, tal como se verifica com a hipertensão, o risco de AVC atribuível ao tabagismo parece diminuir com a idade (Shinton e Beevers, 1989).

Acerca da dislipidemia, a associação entre o nível total de colesterol sérico e a incidência de AVC (todos os tipos) não é clara. Todavia, estudos recentes sinalizam a existência de diferentes associações entre este factor e o AVC isquémico *versus* AVC hemorrágico (em particular na hemorragia intracraniana), encontrando-se em estudos de coorte prospectivos o risco de AVC isquémico positivamente associado com o colesterol sérico, mas negativamente associado no caso da hemorragia intracerebral. Também a idade, o sexo e os níveis de hipertensão podem modificar esta associação. Em particular, a associação entre mortalidade por AVC e o nível total de colesterol no sangue encontra-se muito dependente da pressão sanguínea (Brainin e Heiss, 2009).

No que respeita à presença de diabetes, as pessoas com esta doença possuem uma maior tendência para aterosclerose, bem como uma maior prevalência de factores aterogénicos, em particular hipertensão, obesidade e níveis anormais de lípidos na corrente sanguínea. Alguns estudos sinalizaram ainda um efeito independente da diabetes, com valores do risco relativo de AVC isquémico em pessoas com esta patologia, situados entre 1.8 e 3.0. Finalmente, acerca da presença de doença cardíaca, é importante destacar a fibrilação atrial, por constituir o mais importante precursor cardíaco de AVC passível de tratamento (Sacco et al., 1997).

Em conclusão, entre os factores de risco individual de AVC modificáveis mais relevantes incluem-se a hipertensão, o tabagismo, a dislipidemia, a diabetes e a presença de doenças cardíacas, notavelmente a fibrilação atrial. Todavia, nem todos os factores contribuem de igual forma para todos os subtipos de AVC. Assim, embora a hipertensão se encontre claramente associada com todos os tipos de AVC, os problemas de carácter anatómico e ao nível do fluxo sanguíneo contribuem mais para o AVC hemorrágico, e os maiores factores de risco de aterosclerose encontram-se mais relacionados com o AVC isquémico (ESO, 2013).

1.6.2 Factores de risco individual de AVC em crianças e adultos jovens

Tendo em conta que a investigação desenvolvida no capítulo 2 desta tese incide sobre a faixa etária que engloba os indivíduos abaixo de 65 anos de idade, é importante também entender a contribuição destes factores de risco do tipo clínico em indivíduos mais jovens, no sentido de ganhar sensibilidade para as semelhanças e diferenças entre estes

casos e os ocorridos em idades mais avançadas. Outras das razões prende-se com o impacto, já anteriormente referido, que esta doença também tem nas faixas etárias mais jovens. De facto, não só a doença cerebrovascular se encontra entre as 10 mais importantes causas de morte durante a infância, como os casos ocorridos em crianças e adultos jovens representam, em países desenvolvidos, 5% a 10% de todos os episódios de AVC. Em países em desenvolvimento, estas proporções atingem valores ainda mais elevados, situando-se entre 19% e 30%. Mais ainda, o impacto do AVC neste grupo etário é devastador para as crianças e adultos jovens, para as suas famílias e para a sociedade (Biller, 2009).

Sendo assim, este autor fornece um ponto de partida para esta discussão, referindo a existência de diferenças na incidência, na apresentação, nos factores de risco e no prognóstico de AVC em indivíduos com menos de 45 anos de idade por comparação com os de mais de 45 anos, e também entre os grandes grupos etários que vão desde neonatais a crianças e destas aos adultos jovens. Alguns destes aspectos serão aqui abordados, dada a sua relevância para esta discussão. Assim, começando pela incidência, embora se verifiquem variações na distribuição mundial destas taxas de incidência em indivíduos jovens, é importante referir que os valores mais elevados se verificam no período perinatal, durante o qual ocorrem 26,4 episódios por cada 100000 nados-vivos em crianças com menos de 30 dias, dos quais 6,7 representam casos de AVC hemorrágico e 17,8 de AVC isquémico. A ocorrência de óbitos segue um padrão similar, na medida em que um dos picos de valores se verifica em crianças com menos de 1 ano, mas já difere um pouco no segundo pico, atingido em adultos dos 35 aos 44 anos (figura 40).

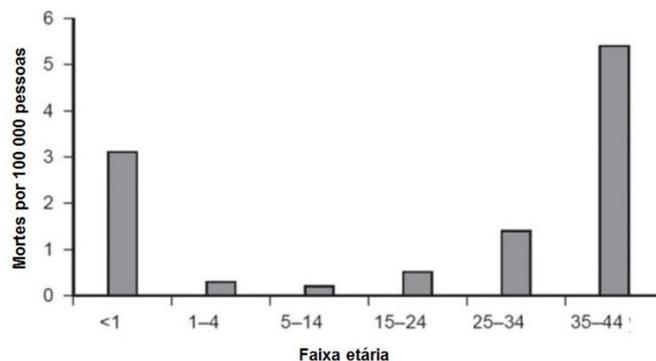


Figura 40 – Óbitos por doença cerebrovascular em crianças e adultos jovens em 2005.
Fonte: Adaptado de Biller (2009).

Tal como se verifica para todas as idades, também nestas faixas etárias o AVC isquémico é mais comum do que o hemorrágico. Todavia, embora esta proporção se mantenha, os casos do tipo hemorrágico assumem uma importância superior em pacientes jovens. Enquanto em adultos o AVC isquémico ocorre em 80% dos casos e o hemorrágico em 20%, já em crianças os casos de AVC isquémico representam 55% das ocorrências e os de AVC hemorrágico 45%. Mais ainda, em pessoas com menos de 45 anos, as ocorrências de hemorragia subaracnóidea e intracerebral representam 42,7% dos casos (enquanto nos pacientes mais idosos representam apenas 15,7% dos casos), estando muitas vezes relacionados com aneurismas e malformações arteriovenosas (Biller, 2009).

Esta diferença de importância relativa nos tipos de AVC pode estar relacionada com a importância relativa dos factores de risco individuais. A este respeito, só para AVC isquémico, existe uma multiplicidade de factores de risco em crianças e adultos jovens, contando-se mais de 100 apenas em crianças, sendo que se referem aqui apenas os mais comuns. Entretanto, no que diz respeito ao AVC em perinatais, esta constitui uma situação distinta, na medida em que têm de ser tidos em conta factores de risco relacionados com a mãe e com o feto, sendo exemplos a infertilidade materna ou as anomalias verificadas ao nível do cordão umbilical, respectivamente, entre muito outros exemplos.

Em termos de factores não modificáveis, também em crianças e jovens adultos são referidas diferenças de carácter étnico no risco de AVC, sendo que nas crianças afrodescendentes o risco de ocorrência de AVC é mais elevado, com um risco relativo de 2.59 no AVC isquémico. Todavia, à semelhança do que foi discutido no subcapítulo anterior, grande parte dos estudos que referem esta disparidade foram realizados no contexto norte-americano, sendo controversa a sua generalização a outras regiões do mundo. Em termos de sexo, a ocorrência de AVC isquémico é mais comum em rapazes do que em raparigas, sendo a probabilidade de sofrer um episódio deste tipo 50% superior em rapazes. No que diz respeito ao historial familiar, não obstante este ser um factor de risco de AVC isquémico na população em geral, a sua contribuição para o risco em indivíduos jovens permanece incerta (Biller, 2009).

Em termos de factores modificáveis relacionados com estilos de vida, o tabagismo representa um factor de risco significativo em indivíduos jovens, sendo que o hábito de

fumar cigarros duplica o risco em jovens adultos. Mais ainda, a presença deste hábito em conjunto com outros factores de risco pode actuar sinergicamente para aumentar o risco de AVC em adultos jovens.

Também o consumo de álcool assume aqui alguma importância. Hillbom et al. (1995) reportam a existência de uma associação entre o AVC isquémico e o consumo de álcool muito recente e em grandes quantidades em adultos jovens entre os 16 e os 40 anos, em casos sem outra etiologia para AVC.

Outro factor de risco relevante para esta discussão é sem dúvida o consumo de drogas estupefacientes, sendo que este hábito representa um aumento de 6.5 vezes no risco de AVC face aos não-consumidores. Entre pacientes com menos de 35 anos, o consumo abusivo destas substâncias representou o factor de risco mais frequente, detectado em 47% dos casos, representando um aumento do risco relativo de 11.7 (Barlow, 1984). Outro estudo mais recente indica que 14% dos episódios de AVC hemorrágico e 14% dos episódios de AVC isquémico tiveram origem no consumo abusivo de drogas tais como anfetaminas e cocaína. O mesmo estudo associa o consumo de anfetaminas com um risco 5 vezes maior de AVC hemorrágico em indivíduos entre os 18 e os 44 anos e refere igualmente que o consumo de cocaína duplica o risco de AVC, tanto isquémico como hemorrágico (Westover et al., 2007).

Finalmente, a obesidade assume aqui um papel muito relevante entre os factores de risco associados a estilos de vida, por constituir presentemente a doença mais prevalente em crianças e jovens adultos. Assim, crianças com excesso de peso encontram-se na eminência de se tornarem igualmente adultos com excesso, e conseqüentemente com risco mais elevado de desenvolvimento de hipercolesterolemia, hipertensão, diabetes, doença coronária e doenças cerebrovasculares (Biller, 2009).

A ideia de que o risco destes últimos factores se manifestará mais tarde, na passagem à idade adulta é confirmada por alguns estudos, sendo que o grau de risco que advém destas patologias diverge do que acontece nas faixas etárias mais idosas, particularmente nos factores de risco ateroscleróticos. Assim, a aterosclerose é em si invulgar como causa de AVC em indivíduos com idades inferiores a 30 anos (Bogousslavsky e Pierre, 1992); (Bendixen et al., 2001). Aliás, a aterosclerose como causa de AVC apenas se verificou em 2% dos doentes entre 16 e 30 anos de idade, sendo que a maioria destes doentes possuía, não obstante, factores de risco clássicos

como hipertensão arterial, diabetes, hábitos de fumo e hiperlipidémia. Todavia, a importância de factores de risco como a hipertensão e a diabetes não deve ser minimizada, sendo que, num estudo de caso-controle, a hipertensão foi detectada em 31% dos pacientes abaixo de 50 anos de idade (Matias-Guiu et al., 1990). Acerca da diabetes, alguns investigadores colocam esta doença em segundo lugar em termos de risco logo a seguir à hipertensão, embora nesta faixa etária das crianças e adultos jovens, o aumento substancial do risco de AVC devido a esta patologia derive da sua combinação com outros factores de risco como a hipertensão, hiperlipidémia, abuso de álcool e consumo de tabaco.

É ainda de referir que existem outras patologias que podem aumentar o risco de aterosclerose em crianças e adultos jovens, tais como alguns distúrbios metabólicos de origem genética (Biller, 2009).

1.6.3 Determinantes socioeconómicos do AVC

Após a discussão dos factores de risco individuais do tipo clínico (proximais), este subcapítulo discute os factores distais, que os antecedem, numa perspectiva mais ligada à Saúde Pública e ao enfoque na “causa das causas” referido por Marmot e Wilkinson (2005).

Um conceito relevante neste ponto é o grau de exposição de uma pessoa a um conjunto de factores socioeconómicos, no que pode ser entendido como o seu estatuto socioeconómico. Este estatuto engloba um certo número de recursos (tais como dinheiro, conhecimento, prestígio, poder e ligações sociais vantajosas) que protegem a saúde da pessoa, independentemente dos mecanismos que actuam de forma mais relevante numa dada altura do tempo (Labarthe, 2011).

Todavia, o debate dos factores socioeconómicos é mais complexo e apresenta geralmente mais desafios do que aquele que se refere aos factores individuais de risco. Assim, apesar da existência de associações entre a saúde e os factores socioeconómicos se encontrar já estabelecida, as vias de causalidade através das quais o efeito destes fenómenos se manifesta não são ainda bem compreendidas, podendo existir uma variedade de mecanismos que interagem de forma diferencial originando diferentes resultados. Não obstante, o desenvolvimento de modelos conceptuais que tentam

explicar estas vias de causalidade contínua. Neste contexto, a relação entre AVC e factores socioeconómicos/estatuto socioeconómico não constitui excepção à regra. E também aqui a complexidade impera, resultante em parte da dissemelhança entre estudos publicados, na medida em que possuem metodologias e objectivos bastante diferenciados, o que dificulta grandemente a comparação entre estudos, mas não só.

A figura 41 exemplifica a complexidade das relações entre estatuto socioeconómico e AVC, recorrendo a um modelo das vias de causalidade, através das quais este estatuto pode influenciar o risco de AVC e a severidade das suas consequências (Cox et al., 2006).

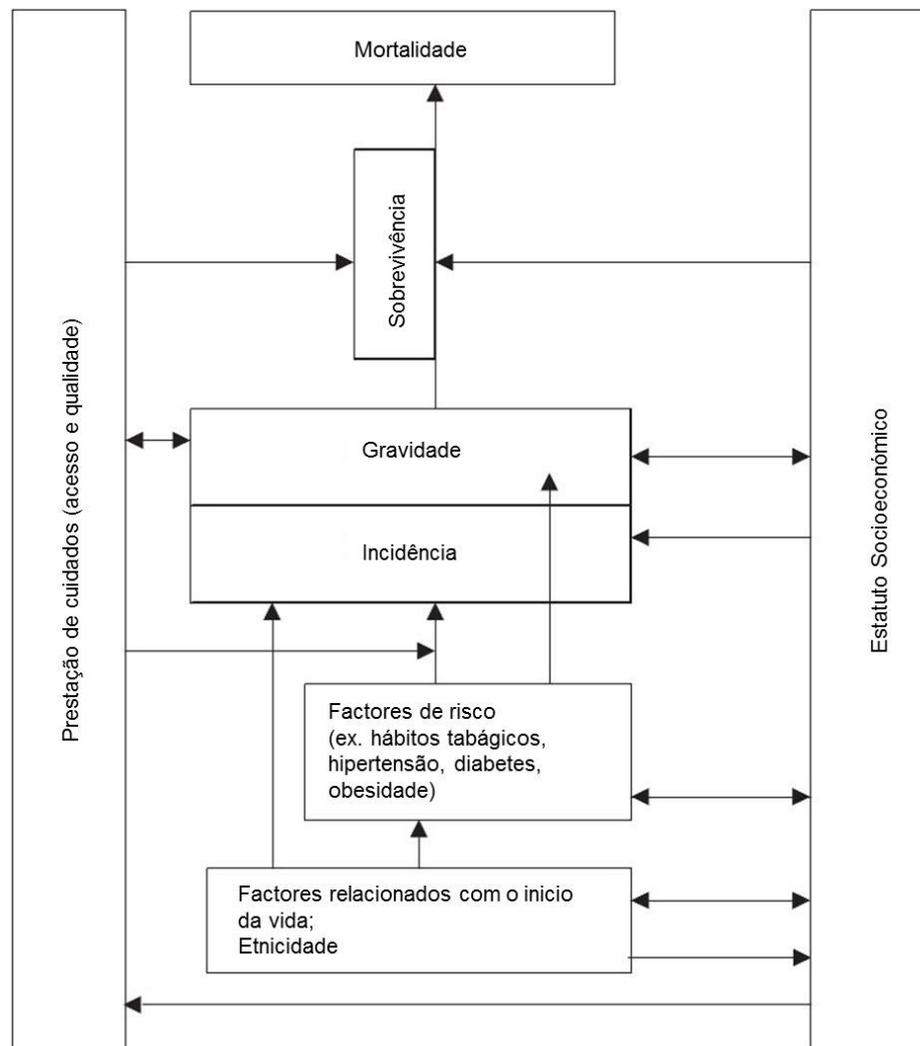


Figura 41 – Modelo de relações entre estatuto socioeconómico e AVC.
 Fonte: Adaptado de Cox et al. (2006).

Para começar, os factores socioeconómicos que afectam a saúde têm sido estudados tanto ao nível individual como ao nível ecológico. Ambos os níveis de estudo apresentam vantagens e desvantagens; enquanto os estudos ao nível individual fornecem uma visão mais restrita, os estudos ecológicos fornecem uma visão mais abrangente mas os resultados das associações entre variáveis podem não ser aplicáveis ao nível individual.

Partindo do nível individual, estes factores têm sido classificados como materiais (e.g. rendimento e outras possessões materiais), comportamentais (e.g. dieta, consumo de tabaco, exercício, sendo que a maior parte destes últimos já foram discutidos no subcapítulo anterior) e psicossociais (e.g. a percepção de desigualdades ou injustiças). Ao nível ecológico, os estudos têm muitas vezes incluído influências do meio ambiente (em sentido lato) em que a pessoa se encontra inserida (tais como a disponibilidade de serviços de saúde ou o acesso diferenciado a estes) e factores psicossociais de carácter mais abrangente (tais como o nível de suporte social existente numa determinada região). Por outro lado, a tentativa de classificação destes factores não exclui o facto de que estes são provavelmente interactivos e se acumulam no decurso da vida; aliás a saúde de cada pessoa pode ser simultaneamente um determinante e um resultado do seu estatuto socioeconómico. Uma lista de indicadores usados na medição deste estatuto, bem como a sua discussão e categorização, pode ser observada na tabela apresentada no anexo 1 (Cox et al., 2006). Entre as categorias mais frequentes encontram-se a educação, a ocupação profissional, o nível dos rendimentos auferidos, a posse de diversos tipos de bens materiais e ainda vários índices complexos que tentam captar o nível de privação socioeconómica de uma determinada área. Tendo em conta que cada um destes indicadores possui as suas próprias limitações, e na verdade mede aspectos distintos do estatuto socioeconómico, boa parte dos estudos optam por utilizar vários indicadores em conjunto. E isto porque, não obstante os vários indicadores se encontrarem muitas vezes associados entre si, eles não são intermutáveis, no sentido em que captam diferentes aspectos do estatuto socioeconómico, com diferentes associações e implicações.

Os estudos que abordam a relação entre determinantes socioeconómicos e AVC focam-se, em grande parte das vezes, na análise da associação entre diversos aspectos do

estatuto socioeconómico e mortalidade e/ou incidência, mas também noutros desfechos possíveis, tais como a sobrevivência ou o grau de severidade da ocorrência.

Em geral, os estudos já efectuados reforçam a existência de associações mais fortes no caso da mortalidade e da incidência, tendo sido encontradas de forma consistente, taxas elevadas destes dois indicadores em grupos de pessoas com baixo nível socioeconómico. Já no que diz respeito à associação com a sobrevivência e o grau de severidade, as evidências têm-se revelado mais ténues (Cox et al., 2006). Parece existir uma relação inversa entre a mortalidade por AVC e o estatuto socioeconómico, a qual tem sido particularmente estudada nos EUA, Japão e Europa Ocidental (Fukuda et al., 2005); (Steenland et al., 2004). Verificam-se inclusive nestas zonas do planeta algumas das mais elevadas associações encontradas, sendo que a força desta relação parece sofrer um aumento com a idade.

No que diz respeito à incidência, Cox et al. (2006) realizaram uma análise comparativa entre vários estudos que analisam a sua ligação aos factores socioeconómicos. Segundo estes autores, embora os vários estudos apresentem características diversas no que diz respeito aos respectivos desenhos e indicadores utilizados, todos reportavam uma relação inversa entre estatuto socioeconómico e incidência de AVC. Alguns destes estudos realizavam também a análise por tipo de AVC, sendo que nestes, a relação entre factores socioeconómicos e incidência de AVC hemorrágico apresentava-se semelhante ou mais forte do que a verificada para todos os tipos desta doença.

Para além da relação entre AVC e estatuto socioeconómico, é ainda relevante ter-se em conta a relação conjunta deste estatuto e dos factores de risco clássicos (clínicos), sendo que as evidências vêm reforçar a cadeia de causalidade previamente debatida. Nesse sentido, estudos realizados na população geral em países desenvolvidos, evidenciaram uma alta prevalência de muitos dos factores clássicos de risco em grupos de estatuto socioeconómico mais baixo, tendo sido encontradas associações inversas com hipertensão, tabagismo, diabetes, inactividade física e obesidade, embora a ligação entre colesterol e estatuto socioeconómico se tenha revelado inconclusiva (Cox et al., 2006). Também nos países em desenvolvimento, não obstante alguns estudos menos recentes referirem uma associação directa entre estatuto socioeconómico e certos factores de risco como a hipertensão e a obesidade (Colhoun et al., 1998), estudos posteriores evidenciaram a alteração deste padrão, com a situação a tender para se aproximar do

que acontece nos países desenvolvidos. Em 2004, um relatório da Organização Mundial de Saúde veio reforçar esta ideia, referindo a deslocação progressiva do peso dos factores de risco para os grupos socioeconómicos mais desfavorecidos, acompanhando o aumento do produto interno bruto (Monteiro et al., 2004). Um exemplo desta situação foi reportado num estudo mais recente realizado na Índia, no qual baixos níveis de escolaridade e de rendimento foram já associados a altos níveis de tabagismo e alta prevalência de diabetes, especialmente em áreas urbanas (Srinath Reddy et al., 2005).

Mais recentemente ainda, Addo et al. (2012) confirmam esta mudança, afirmando que as alterações epidemiológicas e demográficas resultantes não só do envelhecimento da população, mas também da distribuição dos factores de risco nesta, levavam a que o AVC constituísse já uma das grandes preocupações de saúde pública em países de baixo e médio rendimento (esta questão foi já desenvolvida no subcapítulo 1.5.1 referente à distribuição e evolução do AVC no mundo).

Não obstante, alguma cautela se impõe aqui, na medida em que, na avaliação da associação entre estatuto socioeconómico e factores de risco clínicos, o ajuste para a prevalência destes factores de risco tem produzido resultados por vezes heterogéneos. Por exemplo, num estudo do risco de AVC realizado em Inglaterra e focado em homens entre os 40 e os 59 anos de idade e empregados em profissões manuais, o ajuste para hábitos de fumo e hipertensão reduziu para metade o excesso de risco (Shaper et al., 1991). Todavia, noutro estudo realizado em Roterdão, os níveis de associação entre estatuto socioeconómico e AVC praticamente não sofreram alterações após o ajuste para factores de risco tradicionais (Rossum et al., 1999).

Detalhando agora um pouco mais a influência de cada componente do estatuto socioeconómico, e tendo como base a classificação de Cox et al. (2006) apresentada no anexo 1, vários estudos evidenciam a existência de associações entre os vários componentes do estatuto socioeconómico e a mortalidade e/ou morbilidade por AVC. Num estudo focado na mortalidade por níveis de escolaridade em dez países europeus, foram encontradas taxas de mortalidade mais elevadas em grupos de pessoas com um nível de escolaridade abaixo dos níveis mais elevados do ensino secundário ou equivalente em todos os países e grupos etários analisados (Avendaño et al., 2004). Alguns estudos referem também a importância do nível de educação sob a perspectiva de uma maior literacia das pessoas sobre a saúde, o que por sua vez se encontra

relacionado com um menor risco de AVC (Wardle e Steptoe, 2003); (Müller-Nordhorn et al., 2006). Num outro estudo, englobando vários tipos de desigualdades socioeconómicas nos EUA e também em países da Europa, foram encontrados altos níveis de mortalidade por AVC em pessoas com profissões manuais, quando comparados os valores com as taxas nacionais. Exemplificando, as taxas de mortalidade apresentavam valores mais elevados em pessoas com profissões manuais (todas as idades) em Inglaterra, Finlândia, Noruega, França e Portugal, do que em pessoas com profissões não manuais (Kunst et al., 1998). Num estudo focado na ocorrência de AVC precoce (pessoas entre os 20 e os 59 anos de idade), Falcão e Carvalho (2004) indicam que em muitos dos casos os indivíduos atendidos desempenhavam as suas tarefas profissionais no “sector informal”, no que constitui um indício de situações de pobreza e baixas oportunidades, anteriores ao episódio de AVC.

Outro aspecto relacionado com a ocupação profissional é o número de horas de trabalho por unidade de tempo (geralmente a semana ou o mês). Foram já realizados vários estudos relacionados com este indicador, particularmente no Japão, país onde é notória a tendência para o excesso de horas de trabalho (Landsbergis, 2004). Entre os estudos mais credíveis, Hayashi et al. (1996) detectaram níveis mais altos de hipertensão arterial num grupo de trabalhadores de escritório que trabalhavam mais de 60 horas extraordinárias por mês, quando comparados com outros profissionais que não excediam metade destas horas de trabalho. Notaram igualmente, num grupo de trabalhadores com mais de 96 horas extraordinárias por mês, que os valores da tensão arterial se encontravam significativamente altos e as horas de sono bastante mais reduzidas. Num estudo realizado no estado da Califórnia (EUA), Yang et al. (2006) reportaram também o aumento da hipertensão arterial com o aumento das horas de trabalho na população trabalhadora deste estado. Num estudo realizado no Brasil, Mendes (1988) refere a existência de valores mais elevados de hipertensão em pessoas que cumpriam mais de 48 horas de trabalho semanal.

Outro indicador relacionado com a ocupação profissional ou, mais ironicamente, com a ausência desta, refere-se ao efeito potencial da situação de desemprego no AVC. Franks et al. (1991) reportam, num estudo efectuado em Londres, a existência de uma associação forte e significativa entre a mortalidade por AVC e o desemprego em homens entre os 45 e os 74 anos. Noutro estudo mais recente, desta vez realizado na

Argentina, o desemprego encontrava-se associado a um alto risco de mortalidade intra-hospitalar (Sposato et al., 2012). Uma questão relevante e transversal a muitos estudos que abordam a relação do AVC com o número de horas de trabalho, mas também aos que analisam a relação com o tipo de ocupação da pessoa é o facto de incidirem sobre pessoas em idade laboral, o que vem reforçar a importância da análise do AVC em pessoas não idosas. Um caso ilustrativo pode ser encontrado em (Tsutsumi et al., 2011), num estudo prospectivo focado especificamente nos efeitos do *stress* ocupacional numa população japonesa de pessoas em exercício de funções laborais (3190 homens e 3363 mulheres), com 65 ou menos anos de idade e sem historial prévio de doenças cardiovasculares. Entre os resultados deste estudo, observou-se que os indivíduos do sexo masculino em situações profissionais com grande exigência e baixo nível de controlo possuíam quase 3 vezes mais probabilidade de sofrer um episódio de AVC do que indivíduos em situação similar mas com altos níveis de controlo, embora o mesmo resultado não tenha sido observado no sexo feminino. Marmot e Wilkinson (2005) referem também resultados similares, no contexto dos gradientes sociais analisados no âmbito do estudo de Whitehall II, realizado sobre uma população de funcionários públicos britânicos com diferentes posições na hierarquia, embora este estudo seja focado no enfarte agudo do miocárdio.

Entre os componentes do estatuto socioeconómico encontra-se também o nível de rendimento. Este factor, que remete directamente para os níveis de privação materiais, constitui uma boa forma de avaliar a capacidade de aquisição (de bens materiais, educação e serviços de saúde), prestando-se à medição com recurso a vários tipos de indicadores. No âmbito de um estudo realizado na Finlândia, foram feitas várias análises do efeito dos níveis de rendimento na incidência e mortalidade por tipo de AVC, com resultados bastante afirmativos. Assim, na associação entre incidência de AVC isquémico e níveis de rendimento dos pacientes, a incidência padronizada por idades revelava-se 2 vezes mais elevada em pessoas de baixo e médio rendimento, tendo-se mantido estes resultados, independentemente da análise por grupos etários ou sexos (Jakovljević et al., 2001). Na mesma análise, mas para casos de AVC hemorrágico (mais especificamente de hemorragia subaracnóidea), a incidência padronizada por idades revelava-se 3 vezes mais elevada em pessoas de baixos rendimentos (particularmente no grupo etário entre os 25 e os 44 anos de idade, em

ambos os sexos) quando comparadas com o grupo correspondente de rendimentos elevados. Noutro estudo abordando também casos de morte por hemorragia subaracnóidea, mas desta vez realizado nos EUA e Canadá, com base num coorte de 31631 pacientes americanos e 16531 canadianos com idade média de 58 anos, foi reportada uma associação relevante entre o nível de rendimento (medido com recurso ao rendimento médio do agregado familiar para a área correspondente a um determinado código postal) e a mortalidade por AVC, com um *odds ratio* de 0.77 (IC 95%: 0.65-0.93) (Blessing et al., 2013). Voltando ao AVC isquémico, Hanchate et al. (2013) reportaram também um excesso significativo de mortalidade intra-hospitalar em pacientes oriundos de áreas de baixo rendimento, na comparação com pacientes residentes em áreas de rendimento elevado, com um *odds ratio* de 1.08 (IC 95%: 1.02-1.15). Finalmente, outro conjunto importante de factores socioeconómicos diz respeito àqueles relacionados com o ambiente, tanto em sentido restrito como em sentido lato, referidos por Labarthe (2011), discussão atrás desenvolvida no subcapítulo 1.6.

1.7 Geografia da Saúde, Epidemiologia Espacial e Doenças Cerebrovasculares

O presente subcapítulo introduz e enquadra a utilização de técnicas de análise espacial em Epidemiologia, com o objectivo de fornecer uma abordagem às metodologias experimentadas nos casos de estudo, cuja operacionalização se encontra nos capítulos 2 e 3. Para tal, começa-se por fornecer uma breve introdução à Epidemiologia Espacial e ao seu modo de funcionamento típico. Num segundo momento, referem-se algumas considerações de base que devem ser ponderadas nestes estudos, particularmente acerca dos dados a utilizar e de potenciais efeitos que decorrem da espacialização dos fenómenos. Por fim, aprofundam-se os principais tipos de métodos analíticos utilizados neste domínio científico, ilustrados também, sempre que possível, através de casos de estudo no âmbito das doenças cerebrovasculares.

1.7.1 Geografia da Saúde, Geografia da Doença, Epidemiologia Espacial

As aplicações de Ciência da Informação Geográfica na prática de Saúde Pública são uma realidade desde há algum tempo. De forma muito abrangente, estas intervenções podem ser enquadradas no âmbito da Geografia da Saúde (Moon, 2009) ou ainda no da Geografia Médica (Meade e Emch, 2010), designação surgida anteriormente mas igualmente presente em muita da literatura disponível, existindo inclusive controvérsia entre autores acerca de qual destas designações deve ser adoptada. Por um lado, (Santana, 2005) indica que, após abordar meramente o estudo das topografias médicas ou sanitárias, a Geografia Médica evoluiu para Geografia da Saúde, o que lhe veio permitir expressar melhor a sua riqueza metodológica e de objectivos e também a abrangência temática da sua intervenção. Por outro lado, Moon (2009) indica que, na prática, existe muito pouco que permita distinguir o trabalho feito numa ou noutra disciplina. Meade e Emch (2010) caracterizam ainda a Geografia da Saúde como uma disciplina mais focada na dimensão social da Saúde, que se foi tornando cada vez mais uma preocupação fulcral nos países economicamente desenvolvidos. Por simplicidade, adopta-se aqui a visão da Geografia da Saúde. Este termo surge por volta de 1980 com o objectivo de descrever o cruzamento entre Geografia, morte, doença e aspectos relacionados com a provisão de serviços de saúde (Moon, 2009). Por seu lado, Boulos et al. (2001) desdobram esta disciplina em duas subdisciplinas, a Geografia da Doença e a Geografia dos Sistemas de Cuidados de Saúde. Acerca da primeira, este autor apresenta uma definição abrangente para este ramo de aplicações; “A geografia da doença, [...] cobre a exploração, descrição e modelação da incidência espaçotemporal das doenças e fenómenos ambientais relacionados, a detecção e análise de padrões e *clusters* das doenças, a análise das causas e a geração de novas hipóteses”. O termo *cluster* refere-se geralmente a regiões (na maior parte das vezes definidas num espaço bidimensional) nas quais são observadas contagens ou leituras de fenómenos superiores ao esperado (Lawson e Denison, 2002).

Esta definição aproxima-se dos estudos de Geografia Médica das Doenças. Segundo Winston e Emch (2013), estes estudos são muitas vezes enquadrados na conjuntura teórica da ecologia das doenças e, particularizando um pouco mais, os que se focam em métodos espaciais são enquadrados na Epidemiologia Espacial. A este respeito, Elliott e

Wartenberg (2004) referem que esta disciplina se dedica à descrição e análise das variações geográficas na doença face a factores de risco demográficos, ambientais, comportamentais, socioeconómicos, genéticos e infecciosos.

Desta forma, pode-se considerar que a Epidemiologia Espacial e as suas metodologias e técnicas de análise espacial fornecem o enquadramento mais apropriado ao presente estudo. É relevante neste ponto referir que a análise espacial (neste caso, de forma mais precisa, geoespacial) pode ser definida como a que engloba as técnicas aplicáveis quando os dados podem ser localizados num referencial bidimensional (pelo menos) e dizem respeito a actividades terrestres. Neste contexto, os resultados de qualquer análise deste tipo alterar-se-ão se a localização ou a extensão do referencial se alterar ou ainda se os objectos forem reposicionados dentro do mesmo. Caso nada disto se verifique, a localização torna-se irrelevante, e será mais apropriado recorrer a técnicas convencionais, no sentido de serem não-espaciais (de Smith et al., 2015).

Os estudos de Epidemiologia Espacial podem recorrer a um vasto conjunto de técnicas, sendo por vezes difícil escolher o caminho a seguir. Nesse sentido, Pfeiffer et al. (2008) fornecem como orientação um quadro conceptual para a realização de análise espacial de dados em Epidemiologia (figura 42).

Desta forma, a análise epidemiológica espacial tem como objectivos descrever padrões espaciais, identificar clusters de doença e explicar ou prever o risco de doença.

Nada disto seria possível sem a disponibilidade de dados, sendo neste caso necessário ter acesso a uma base de dados capaz de combinar dados convencionais, contendo os atributos da(s) entidade(s) em estudo, com dados georreferenciados, contendo a representação espacial das entidades, seja qual for o modelo do espaço e o tipo de representação adoptado para as entidades (estas questões serão particularizadas no ponto 1.7.2).

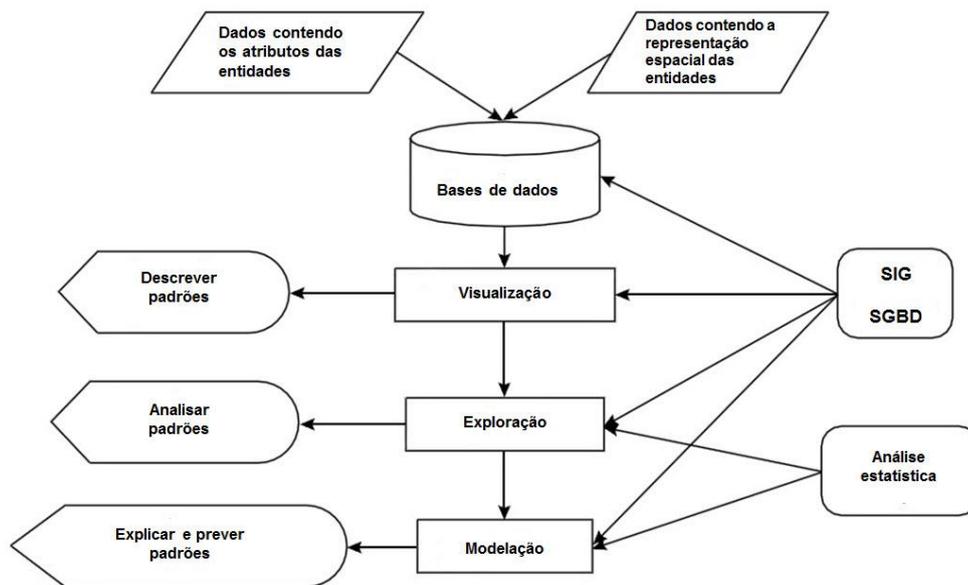


Figura 42 – Quadro conceptual da análise espacial de dados epidemiológicos.
 Fonte: Adaptado de Pfeiffer et al. (2008).

As tarefas de gestão e análise de dados (espaciais ou não) são efectuadas por Sistemas de Informação Geográfica (SIG) e Sistemas de Gestão de Bases de Dados (SGBD) e exigem geralmente ao epidemiologista o conhecimento de uma variedade de conceitos de base cuja relevância para a questão analítica de partida pode não ser imediatamente aparente.

A resposta à questão de partida desdobra-se geralmente em objectivos analíticos específicos, os quais levam por sua vez a três tipos de métodos analíticos, nomeadamente a visualização, a exploração e a modelação. Os dois primeiros tipos envolvem o uso de técnicas que se focam apenas em examinar a dimensão espacial dos dados. Neste ponto, o método utilizado com mais frequência é a análise visual, a qual implica a preparação de mapas ilustrativos dos padrões espaciais dos fenómenos em estudo, e cuja observação permite não apenas estimular a realização de análises mais complexas, mas também comunicar os resultados dessas mesmas análises. No que diz respeito à exploração, esta implica já o recurso a métodos estatísticos, no sentido de se percepcionar se os padrões espaciais observados são ou não aleatórios. Finalmente, a modelação prende-se com a investigação de relações de causa-efeito, utilizando-se tanto dados espaciais como não-espaciais com o intuito de explicar ou prever padrões espaciais.

Existe também alguma sobreposição entre tipos de métodos, sobretudo entre visualização e exploração, na medida em que a exibição visual pode necessitar de recorrer a métodos analíticos quantitativos (por exemplo na aplicação de escalas de representação para a produção de um mapa temático). É importante ter-se também em conta que, embora os grupos incluídos neste quadro conceptual permitam organizar a execução das técnicas de análise espacial num processo lógico e sequencial, este pode não ser sempre um processo linear, bastando ter-se em conta que a apresentação dos resultados da exploração e modelação implica um regresso à visualização.

A Epidemiologia Espacial constitui actualmente um campo de estudo de relevo, sendo que ao longo dos últimos 20 anos, a aplicação de técnicas de análise espacial em situações de vigilância epidemiológica ou de pesquisa tem sofrido um crescimento exponencial (Pfeiffer et al., 2008). Mas talvez a referência mais emblemática e o primeiro caso de estudo que se aproxima das aplicações actuais tenha sido posto em prática em Londres por John Snow, médico da época, durante a epidemia de cólera de 1854. Neste estudo, a avaliação do padrão espacial (de concentração) dos pontos representativos dos casos de cólera face à localização das bombas de água para consumo público que representavam um factor de risco potencial (figura 43) veio a desempenhar um papel importante na identificação do foco de infecção, embora se reconheça nos nossos dias que não constituiu a única contribuição, do ponto de vista epidemiológico.



Figura 43 – Mapa de John Snow do surto de cólera de Londres de 1854. As localizações dos óbitos são mostradas como pontos e as fontes de água como cruzes. Fonte: Pfeiffer et al. (2008).

Desde os tempos de Snow até aos nossos dias, a Epidemiologia Espacial sofreu enormes avanços, grande parte relacionados com o advento das ciências da computação, tendo-se notado um crescimento acentuado dos estudos nesta área de aplicação da década de 1980 em diante. Desde então que se tem verificado um crescente aumento da produção científica nesta área, existindo já uma variedade de textos, desde os focados em aspectos particulares da análise espacial, como Lawson e Williams (2001), que abordam a Cartografia das doenças, e Lawson (2001), que se centra especificamente na análise estatística em Epidemiologia Espacial, até aos textos que abrangem toda a área de estudo, como Elliott et al. (2001). Neste domínio, têm vindo também a surgir várias publicações científicas com revisão por pares, entre as quais se destaca o *International Journal of Health Geographics* (IJHG, 2015), por ter sido a primeira publicação periódica especializada em Epidemiologia Espacial a surgir, e conferências especializadas neste domínio, tais como a Conferência de Epidemiologia Espacial, realizada em Londres em 2006 e, mais recentemente e remetendo para o contexto Português, o GeoSaúde 2014 - I Congresso de Geografia da Saúde dos Países de Língua Portuguesa.

Tal como afluído anteriormente, as ciências da computação têm tido aqui um papel impulsionador. De facto, a disponibilidade crescente de aplicações de *software* capazes de lidar com Informação Geográfica “amigáveis para o utilizador” tornou a análise espacial mais acessível a epidemiologistas e outros investigadores. É curioso constatar que a maior parte dos avanços tiveram mais a ver com a implementação de funcionalidades e com a crescente variedade deste tipo de software, enquanto que, no domínio da análise estatística espacial as ferramentas ainda se encontram algo dispersas por diversas aplicações, algumas das quais exigem inclusive conhecimentos de programação (Pfeiffer et al., 2008). Neste ponto, é também útil distinguir entre o software de visualização de mapas, que permite apenas produzi-los, possuindo eventualmente capacidades limitadas de análise e introdução de dados, sendo exemplos as aplicações *ESRI ArcGIS Explorer* (ArcGIS Explorer, 2015) e *Google Earth* (Google Earth, 2014) e o software capaz de suportar um sistema de informação geográfica completo (esta questão será mais desenvolvida no subcapítulo 1.7.2). Entre os exemplos representativos do segundo tipo encontram-se o software comercial *ESRI ArcGIS* (ESRI, 2014) e a sua contrapartida livre e de código-fonte aberto *QGIS* (QGIS, 2014), que engloba já um grande leque de ferramentas que permitem a introdução, gestão e análise de dados espaciais e a apresentação de resultados (Pfeiffer e Hugh-Jones, 2002). Existem ainda produtos mais virados para a análise estatística espacial, tais como o software *GeoDa* (GeoDa, 2014), de livre acesso e que disponibiliza um número razoável de ferramentas exploratórias, e o software comercial *Biomedware ClusterSeer* (BioMedware, 2014) dedicado à realização de análise espacial e espaço-temporal de clusters. A este nível, embora os softwares SIG incorporem um pequeno número de ferramentas de análise estatística espacial, são ainda os pacotes estatísticos, como a linguagem de programação *R* (R, 2014a) para computação estatística e gráficos, que incorporam um maior número destas ferramentas e oferecem mais opções de configuração.

Como anteriormente indicado, um factor crucial na realização destes tipos de análises é a disponibilidade de dados. Assim, também a disponibilização de dados georreferenciados tem vindo a aumentar, potenciando o avanço da análise espacial em Epidemiologia. Aqui, um requisito fundamental consiste em obter os dados relativos tanto ao numerador como ao denominador geograficamente referenciados, e a uma

resolução espacial suficientemente elevada para permitir a obtenção de inferências significativas (Shekhar e Xiong, 2008). Embora tenha sido sempre possível obter estes dados no âmbito de estudos específicos (tais como os que recorrem a inquéritos, ou aqueles que se apoiam na vigilância epidemiológica de rotina), grande parte das vezes estes dados não existiam ou não se encontravam disponíveis. Todavia, avanços recentes no desenvolvimento de *hardware* e software permitem já o processamento regular de dados de alta resolução para uso na gestão e análises simples levadas a cabo pelas autoridades administrativas locais.

Actualmente, muitas fontes de dados são hoje geridas por organizações governamentais ou intimamente ligadas à actividade governativa, tais como institutos produtores de cartografia e cadastro, institutos nacionais de estatística ou ainda ligados à distribuição de correio. E se muitas destas organizações cobram um valor pela cedência dos dados, por outro lado tentam em geral manter e melhorar níveis elevados de qualidade nos dados que produzem. Outra questão relevante associada ao custo dos dados prende-se com o seu grau de actualização. Também a este nível tem havido progressos, sendo que, por exemplo, certos tipos de dados relativos a variáveis ambientais e obtidos através de detecção remota (Longley et al., 2005) podem ser actualizados quase em tempo real e com custos reduzidos. Também a disponibilidade geral de sistemas de posicionamento global (GPS) de baixo custo permite actualmente a recolha de dados no terreno que podem ser prontamente georreferenciados.

E assim como aumentaram os tipos de dados disponibilizados e a tecnologia para os disponibilizar, o mesmo se verificou ao nível das fontes de dados, sendo que a qualidade pode variar muito em todas estas frentes. Esta situação tem motivado esforços para padronizar formatos e qualidade e para facilitar o acesso *online* através de portais de informação geográfica a dados ou metadados, tais como o *GeoNetwork* (GeoNetwork, 2015), a plataforma *Earth Resources Observation and Science* (EROS) *Center* (EROS Center, 2015) para disponibilização de imagens de observação da Terra suportada pelo governo dos Estados Unidos da América, o Sistema Nacional de Informação Geográfica (SNIG) (SNIG, 2015) e a iniciativa iGEO – Informação Geográfica, ambas no contexto Português (iGEO, 2015). Ao nível dos esforços para padronizar formatos de dados espaciais, formas de acesso e qualidade, destacam-se a iniciativa europeia *INSPIRE* (Inspire, 2015), o *OGC* (Open Geospatial Consortium)

(OGC, 2015) e o contínuo desenvolvimento das normas *ISO* que incidem sobre este domínio, tais como a norma 19115 (ISO, 2015).

1.7.2 Considerações essenciais em Epidemiologia Espacial

Como já referido, a disponibilidade de dados ocupa um lugar de relevo entre os temas essenciais a ter em conta na análise epidemiológica espacial. Tipicamente, os dados recolhidos com o propósito de serem utilizados na investigação epidemiológica encontram-se mais focados nos atributos das observações, tal como o diagnóstico de uma determinada doença nas pessoas observadas. Adicionalmente, a recolha da localização absoluta (através de coordenadas) dos locais das observações permite estender, de várias formas, o leque de possibilidades face às análises mais convencionais dos factores de risco. Num primeiro momento, a inclusão de uma referência geográfica para cada observação vai permitir efectuar análises que incorporam as relações espaciais entre observações e seus atributos. Adicionalmente, vai também permitir a obtenção de atributos adicionais através da ligação a outros dados georreferenciados com base na localização. Tudo isto é feito no seio de sistemas digitais de informação cujo componente chave reside na representação da dimensão espacial, e desta forma, a investigação ligada à descrição e compreensão dos mecanismos que influenciam a ocorrência das doenças pode beneficiar muito destes sistemas, na medida em que representam o ambiente nos quais estes mecanismos operam. Outra questão relevante reside na representação da dimensão temporal no seio destes sistemas, sendo esta geralmente incluída sob a forma de mais um atributo das entidades espaciais (Pfeiffer et al., 2008).

É importante referir que a disponibilidade de dados espaciais e a modelação da dimensão temporal são duas questões de tal relevância que alguns autores não hesitam em referir que a aplicação de Sistemas de Informação Geográfica na Geografia da Doença só concretizará todo o seu potencial após o desenvolvimento de sistemas de vigilância do ambiente e das doenças capazes de fornecer dados espaciais acerca das condições ambientais, dos agentes causais das doenças e dos efeitos na saúde das pessoas ao longo do tempo, com base em interrogações definidas pelos utilizadores sobre áreas geográficas delineadas pelos mesmos (Cromley, 2003). Embora esta

situação tenha sofrido desenvolvimentos recentes, por exemplo na aplicação de estatísticas espaçotemporais para detecção precoce de surtos (Takahashi et al., 2008), talvez o estudo de Tsai e Perng (2011), analisando a distribuição espacial da ocorrência dos 13 neoplasmas mais relevantes em Taiwan nos períodos de 1995-1998 e 2005-2008, ilustre melhor as aplicações mais frequentes que fazem uso da dimensão temporal.

Convém sempre ter em conta que, graças à enorme complexidade do mundo real, estas representações digitais das entidades correspondem sempre a abstrações, o que exige um elevado grau de generalização e simplificação (Haining, 2003).

Outra questão relevante tem a ver com os sistemas digitais de informação onde os dados são incluídos, o que no caso dos dados espaciais equivale, na maioria dos casos, a falar de SIG. Todavia, uma das formas mais simples de incluir a dimensão espacial pode ser encontrada em dados relativos a localizações pontuais (como a representação dos centros geográficos das áreas dos municípios) e consiste apenas em adicionar duas colunas de dados contendo as coordenadas X e Y do ponto pretendido, algo que pode ser feito em qualquer sistema de gestão de bases de dados. A partir daqui, é possível produzir um “mapa” rudimentar usando a função de diagrama de dispersão numa aplicação de folha de cálculo. Mas basta necessitarmos de representar também os limites administrativos dos mesmos municípios para precisarmos de recorrer a software mais especializado, tal como um SIG, ou uma aplicação de produção/visualização de mapas. Estas aplicações fazem então toda a diferença, pois permitem integrar de forma precisa vários tipos de dados espaciais, tornando possível relacioná-los espacialmente. Segundo Burrough e McDonnell (1998), a tecnologia SIG resulta da integração de hardware, software e da componente organizacional, sendo incontornável ter-se em conta os requisitos de cada um. Em termos organizacionais, esta situação prende-se com a disponibilidade de pessoas com as competências apropriadas, uma realidade que, parecendo óbvia, muitas vezes não se concretiza. No caso do hardware, é necessário que os computadores, incluindo os dispositivos de entrada e saída de dados, estejam capacitados para lidar com os volumes de dados (por vezes maciços) e com os requisitos computacionais destas aplicações (entre outros componentes, são geralmente necessárias placas gráficas de grande capacidade). No que diz respeito ao software, um SIG pode ser entendido como um sistema que contém no seu núcleo uma base de dados capaz de lidar com informação georreferenciada, à qual se juntam um conjunto de

ferramentas de software que permitem a introdução, a gestão e a análise dos dados e ainda a geração de mapas e produtos associados (Pfeiffer et al., 2008).

Acerca dos dados espaciais propriamente ditos, existem diversos modelos conceptuais que podem ser utilizados para representar o espaço geográfico. Os mais divergentes consistem em racionalizar o espaço como entidades ou, ao invés, como campos. No primeiro caso, o espaço é entendido como estando ocupado por entidades discretas com atributos específicos, cuja posição pode ser mapeada através de coordenadas geográficas. As entidades são tipicamente representadas em ambiente SIG através de pontos, linhas ou áreas, cujo significado pode ser por exemplo uma árvore, uma estrada ou o perímetro de uma barragem, respectivamente. Todas as entidades podem possuir vários atributos associados, como por exemplo o registo de uma pessoa se encontrar afectada ou não por uma determinada doença e o seu telefone de contacto, no caso de pontos que assinalam a residência das pessoas participantes num estudo de caso-controlo. O segundo modelo já conceptualiza o espaço como estando coberto por um determinado atributo que varia continuamente, de acordo com uma determinada função matemática ou campo, sendo exemplo a representação de uma nuvem tóxica com o objectivo de modelar a sua dispersão. Os campos contínuos, como aqueles que representam os padrões espaciais da temperatura do ar ou da altitude, contém geralmente um só atributo, cujo valor varia ao longo da área representada. A escolha de cada modelo dependerá dos dados e do uso a dar-lhes, sendo as entidades tendencialmente utilizadas em processos administrativos e os campos empregues na investigação de processos espaciais. Na prática, a representação destes modelos conceptuais em ambiente SIG pode ser feita nos formatos vectorial, *raster* ou de rede irregular triangulada (TIN), como pode ser observado na figura 44. O formato vectorial representa as formas das entidades espaciais através de um conjunto ordenado de coordenadas ao qual são ligados por sua vez os atributos das entidades, sendo particularmente adequado para descrever linhas, pontos e polígonos representando entidades. O armazenamento dos dados espaciais neste formato permite realizar todo o tipo de cálculos geométricos, tal como determinar comprimentos e áreas e realizar uma multiplicidade de operações de análise espacial específicas deste formato, tais como as que se baseiam em funções de proximidade.

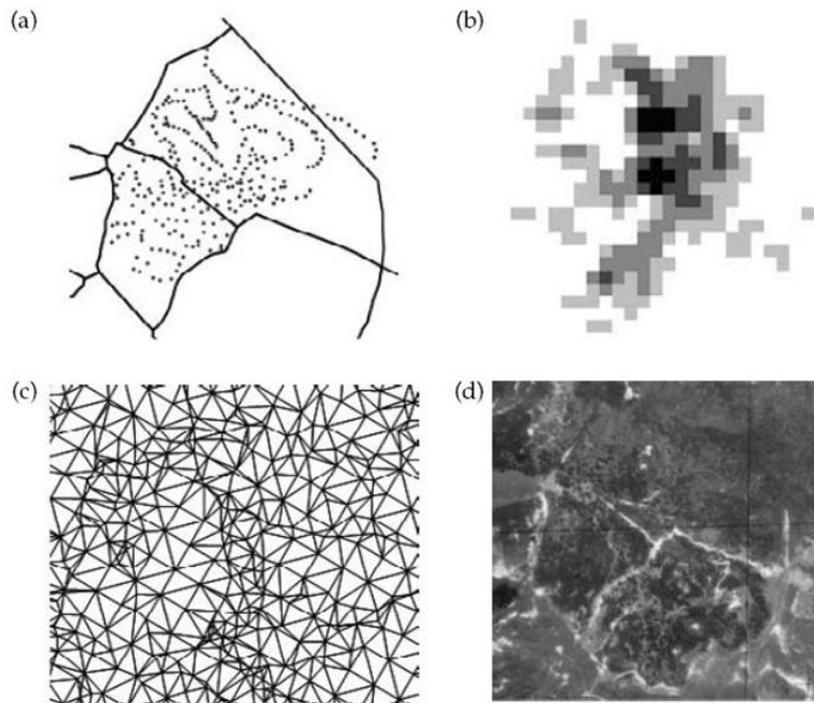


Figura 44 – Exemplos de representações de dados em SIG: a) mapa vectorial, b) mapa raster, c) estrutura TIN, d) fotografia aérea. Fonte: Pfeiffer et al. (2008).

No que respeita ao formato raster, este é particularmente adequado à representação de campos contínuos. Para representar os dados espaciais, recorre a uma grelha bidimensional que segmenta o espaço em células, num procedimento semelhante ao acto de cobrir uma determinada área com uma rede de pesca. Cada célula possui um só atributo que corresponde ao valor do fenómeno espacial descrito, por exemplo a altitude em metros acima do nível do mar. Este valor representa uma sumarização do atributo dentro da área abrangida por cada célula, sendo esta área entendida como a resolução espacial de um conjunto de dados espaciais deste tipo, na medida em que não é possível obter valores para fenómenos que ocorram no terreno em áreas mais pequenas do que as abrangidas pela célula. Assim, a escolha do tamanho da célula é uma decisão que contém as suas ambiguidades, pois se por um lado a escolha de células mais pequenas vai permitir descrever melhor a variação espacial dos valores do atributo, por outro lado vai aumentar as necessidades em termos de espaço de armazenamento e volume de processamento.

Quanto ao formato TIN, este destina-se à representação de superfícies tridimensionais, sendo formado por um conjunto de nós integrados que registam os valores (por exemplo

de altitude) e formam áreas triangulares entre si, ficando o conjunto com a aparência de uma malha de triângulos. Estas estruturas têm como objectivo a realização de análises tais como a identificação de bacias hidrográficas e o cálculo de valores de altitude através de interpolação em qualquer ponto dentro da área abrangida pela rede, entre outras.

Uma possibilidade relevante e que deriva dos modelos de dados já apresentados reside na possibilidade da informação contida em certos atributos ser gerada com base em relações definidas no modelo de dados SIG. Estas relações podem ser classificadas como topológicas, espaciais ou gerais. As relações topológicas têm a ver com a análise das relações espaciais dos objectos entre si, sendo exemplo a identificação de áreas adjacentes, tais como lotes de terreno. As relações espaciais envolvem geralmente a realização de operações entre diferentes temas de dados espaciais, como por exemplo quantificar a área de um parque natural que se encontra abrangida pelos limites administrativos de um município. Nas relações gerais incluem-se aquelas que não podem ser inferidas a partir da posição espacial das entidades relevantes, tornando-se necessário defini-las explicitamente, incluindo-se aqui as ligações a tabelas de dados internas ou externas ao SIG (Pfeiffer et al., 2008).

Existem ainda outras questões de carácter técnico acerca dos dados espaciais que não serão abordadas aqui. Entre estas, incluem-se as estruturas de dados que suportam o armazenamento e partilha destes dados, e ainda os processos relativos à sua recolha, gestão e qualidade. Uma discussão mais aprofundada e actualizada das matérias relacionadas com a tecnologia SIG pode ser encontrada em Longley et al. (2010).

Para além das questões relativas aos dados espaciais, existe uma outra discussão incontornável e com relevância para esta tese, que tem a ver com os efeitos espaciais, os quais decorrem directamente da expressão geográfica dos fenómenos.

Entre estes efeitos a ter em conta encontram-se três conceitos relacionados, a autocorrelação espacial, a dependência espacial e a heterogeneidade espacial. A dependência espacial diz respeito ao relacionamento entre dados espacialmente referenciados, situação que pode ter várias origens, tais como a natureza da(s) variável(s) em estudo e o tamanho, forma e configuração das unidades espaciais. Assim, quanto mais pequenas estas unidades forem, maior a probabilidade das unidades próximas serem espacialmente dependentes, e as unidades espaciais que apresentam

uma configuração longa e estreita têm maior probabilidade de relações de dependência espacial com unidades próximas do que aquelas que apresentam formas mais compactas (Anselin et al., 2010). Por outras palavras, a dependência espacial assume como princípio base que os valores dos atributos medidos em localizações próximas entre si apresentam mais similaridade do que aqueles medidos em localizações mais distantes (Pfeiffer et al., 2008).

Quanto à autocorrelação espacial, esta constitui a forma de medir a dependência espacial. Uma variável pode designar-se autocorrelacionada quando apresenta um padrão sistemático na sua distribuição espacial. A presença e força desta interdependência entre os valores de uma determinada variável, tendo em conta a localização, pode ser quantificada como autocorrelação positiva, inexistente (zero) ou negativa. A presença de autocorrelação espacial positiva indica que valores semelhantes terão tendência para se agregarem, enquanto no caso da autocorrelação espacial negativa serão os valores dissemelhantes que vão apresentar tendência para a agregação. Na maior parte dos estudos, o enfoque é colocado na autocorrelação espacial positiva, chegando Tobler (1970) a considerá-la como endémica nos dados geográficos, ao enunciar este efeito como a Primeira Lei da Geografia. Os dados podem também exibir um padrão espacial aleatório, o qual é indicado por um valor de autocorrelação igual a zero, situação em que os dados (ou seja, os valores dos atributos) não variam em função da sua localização espacial, sinalizando a presença de dados aleatórios, independentes entre si e identicamente distribuídos (Shekhar e Xiong, 2008).

Se a estrutura da dependência espacial não variar, ou seja, for igual para todas as localizações numa determinada área geográfica, o processo espacial designa-se estacionário. Por outro lado, se a estrutura de dependência variar ao longo da área, o processo é entendido como não-estacionário ou heterogéneo. Se a dependência num processo estacionário for afectada apenas pela distância, o processo considera-se isotrópico, mas se a dependência divergir em diferentes direcções, designar-se-á anisotrópico (Pfeiffer et al., 2008).

Assim, a heterogeneidade espacial ocorre quando se verifica a ausência de uniformidade espacial nos efeitos da dependência espacial e/ou dos relacionamentos entre as variáveis em estudo. Uma estrutura de dependência que é inconsistente ao longo da área de estudo denota a ausência de homogeneidade, podendo desta forma a heterogeneidade

espacial ser entendida como um caso particular de dependência espacial, no sentido em que representa uma realização complexa da natureza da variável em estudo e dos efeitos resultantes do tamanho, forma e configuração das unidades espaciais (Anselin et al., 2010). Dito de uma maneira mais relacionada com a forma do que com o processo, a heterogeneidade espacial implica a variabilidade das condições humanas e ambientais ao longo do espaço. Ao nível local, isto significa a variação dos fenómenos, tais como factores socioeconómicos ou o uso do solo, ao longo de uma determinada área geográfica. Estes fenómenos não se distribuem uniformemente no espaço e vão formar padrões diversos, assimétricos, a diferentes escalas e em múltiplas direcções (Shekhar e Xiong, 2008).

A variação total entre os valores dos atributos num processo espacial resulta geralmente de grandes e médias/pequenas variações de escala, ou seja, ao nível da macro-escala ou mesoescala/microescala, respectivamente. Estas variações são também por vezes designadas como efeitos espaciais de primeira e segunda ordem, sendo geralmente medidas numa escala contínua.

A variação ao nível da macroescala manifesta-se através de uma tendência ao longo da área geográfica em estudo, como por exemplo o risco de uma determinada doença diminuir de norte para sul como resultado de diferenças de temperatura que afectam a sobrevivência de um determinado organismo causador de uma infecção. Sob esta perspectiva e a esta escala de análise, este constitui um processo de primeira ordem, pois aqui a disposição espacial dos casos de infecção não decorre da interacção entre indivíduos. Por outro lado, a variação ao nível da mesoescala descreve a variação na estrutura de dependência local de um determinado processo, e configura também heterogeneidade espacial, sendo exemplo o aparecimento de um cluster de pessoas afectadas por uma determinada doença infecciosa em localidades específicas (e não noutras). Aqui já se configura geralmente um processo de segunda ordem, na medida em que o aparecimento de um cluster poderá depender da interacção entre indivíduos infectados, ocorrendo transmissão da doença por contágio (Longley et al., 2010).

Em geral, verifica-se um dos dois tipos, dependente da escala e da extensão a que as observações são efectuadas, sendo que a maior parte dos métodos de estatística espacial apenas permite modelar um destes efeitos, podendo ser obtidos resultados enviesados se

ambos estiverem presentes e forem empregues métodos de modelação convencionais que consideram estes efeitos como fixos.

Outra classe de efeitos espaciais a ter em conta designa-se efeitos de fronteira (*edge effects*), sendo que as fronteiras ou limites de uma área podem consistir em barreiras físicas tais como o mar ou ter a ver com limites impostos como no caso dos municípios ou países ou áreas de estudo. A existência destas fronteiras pode trazer problemas, uma vez que os dados para lá dos limites estão com frequência incompletos, indisponíveis (por exemplo quando já é outro país) ou inexistentes (por exemplo quando o mar constitui a fronteira). Desta forma, é provável que as localizações pontuais ou áreas que se encontram próximo destes limites possuam menos vizinhos do que aqueles que se encontram no centro da área de estudo. Isto coloca um problema quando são efectuados cálculos cuja robustez depende das áreas vizinhas ou quando se efectuam análises para detectar clusters, pois a presença de um número menor de vizinhos pode distorcer as estimativas para pontos ou áreas posicionados perto dos limites, sendo estas distorções designadas efeitos de fronteira. Embora estes efeitos possam ser negligenciáveis quando se trabalha com grandes escalas, podem já ter um impacto considerável na estimação de efeitos em pequenas escalas próximo dos limites. A forma de lidar com estes efeitos consiste em recorrer a um sistema de ponderação que atribui menos peso às observações próximo dos limites, ou através de áreas de protecção.

Uma outra questão que se relaciona também com as fronteiras prende-se com a forma de modelar as relações de vizinhança. A contiguidade e a conectividade são características típicas de um conjunto de processos espaciais conhecido como Topologia. No caso dos dados em formato raster, a topologia encontra-se implicitamente definida pela própria estrutura dos dados, devido à forma regular como as células se dispõem entre si na grelha. Mas no caso dos dados vectoriais, a situação é mais complexa, podendo a topologia ser representada através de vários métodos. A forma mais simples consiste em guardar apenas as coordenadas, sendo as relações de vizinhança derivadas no decurso de uma interrogação à base de dados ou como parte de uma análise estatística. No caso de dados representando polígonos, é possível guardar a informação topológica directamente nos dados, por exemplo as fronteiras partilhadas entre polígonos. Aliás, uma das particularidades que definem o software SIG é a sua capacidade para gerar novos dados com base em transformações e interrogações sobre

os dados disponíveis tendo em conta as relações topológicas e espaciais. Entre os exemplos mais simples, os cálculos de distância e área podem ser facilmente efectuados tanto em dados raster como vectoriais, e o declive pode ser extraído com base em modelos digitais de terreno em formato raster ou TIN. Áreas de *buffer* (zonas-tampão) à volta de entidades espaciais podem ser definidas, por exemplo para circunscrever um cordão sanitário à volta de um local contaminado. As operações de sobreposição tiram partido das relações espaciais entre diferentes temas de dados espaciais, podendo também envolver operações matemáticas simples baseadas em operadores booleanos ou outras operações mais complexas. Um exemplo seria contar os hospitais (representados sob a forma de pontos) que se encontram dentro de uma determinada unidade administrativa, fazendo esta parte de um tema vectorial contendo polígonos representativos de todas as unidades, indicando-se durante o processo que a análise será efectuada apenas para essa unidade administrativa. No caso dos métodos estatísticos que contam com a dependência espacial, é necessário gerar uma matriz de ponderações (ou pesos) espaciais que descreve como as observações num conjunto de dados se relacionam entre si espacialmente. Existe uma diversidade de opções para o cálculo de matrizes deste tipo, sendo o caso mais simples uma matriz binária que regista se os objectos espaciais (por exemplo municípios) são ou não vizinhos contíguos, sendo possível ampliar esta vizinhança (dita de primeira ordem) para adjacências de múltiplas ordens, entre outras opções (Pfeiffer et al., 2008).

Um desafio particularmente relevante tem a ver com a realização de testes de significância envolvendo dados espaciais. Grande parte das tarefas realizadas em estatística são do tipo inferencial, ou seja, recorrem a informação obtida a partir de amostras para tecer conclusões gerais acerca de uma população maior, partindo do pressuposto que a amostra foi retirada dessa população. Dois conceitos constituem a base da estatística inferencial, sendo um deles o uso de intervalos de confiança, situação em que os valores calculados a partir de uma amostra (por exemplo 45% das pessoas, numa amostra de 1000 pessoas retirada da população geral, votam num candidato a um cargo político) são assumidos como representando os valores da população geral, sendo todavia calculada uma margem de erro associada a uma probabilidade, que constitui o intervalo de confiança (por exemplo, 42% a 48% das pessoas vão votar nesse candidato com um nível de confiança superior a 95%, ou seja, a probabilidade de erro nesta

estimativa seria inferior a 0.05). O segundo conceito são os testes inferenciais (mais frequentemente designados como testes de significância ou de hipóteses), situação do tipo “tudo ou nada” em que se coloca uma hipótese (dita hipótese nula) que será rejeitada ou não a favor de uma hipótese alternativa, com base na probabilidade de se cometer um erro de inferência. Por exemplo, a hipótese de 50% dos votantes apoiarem um candidato colocando a eleição num impasse poderia ser rejeitada com um nível de confiança superior a 99%, ou seja, a probabilidade de erro nesta estimativa seria inferior a 0.01 (Longley et al., 2010), ou por outras palavras, teria associado um *p-value* inferior a 0.01.

Todavia, embora a utilização de testes de hipóteses seja prática comum na maior parte da actividade científica, a sua aplicação em estudos envolvendo dados espaciais acarreta muitos problemas, os quais decorrem das propriedades fundamentais dos dados geográficos.

Como já dito atrás, grande parte dos testes inferenciais pressupõe a existência de uma população, da qual é extraída uma amostra através de um processo geralmente bem estabelecido. Para começar, é muitas vezes difícil conceber um determinado conjunto de dados geográficos como uma amostra representativa de todos os conjuntos passíveis de serem extraídos da mesma área coberta, não obstante a existência de todo um conjunto de técnicas de amostragem espacial. Por outro lado, quando se pretende que um conjunto de dados deste tipo seja representativo de uma área mais vasta da superfície terrestre, surgem igualmente problemas, quiçá mais difíceis de contornar. Em primeiro lugar, na maioria dos procedimentos clássicos de estatística que recorrem ao teste de hipóteses, um dos pressupostos fundamentais é a independência das observações, ou seja, assume-se que as amostras foram obtidas independentemente (Longley et al., 2010); (Pfeiffer et al., 2008). Todavia, em muitas situações, um conjunto de dados geográfico inclui todas as ocorrências de um fenómeno existentes na área de estudo, constituindo, no sentido estatístico, a própria população e nesse caso a recolha de uma amostra deixa de fazer sentido. Mesmo considerar um conjunto de dados geográfico representativo de uma área maior trará o problema da amostra não ter sido obtida aleatoriamente, na medida em que teria sido na verdade recolhida por selecção sistemática de todas as observações dentro da área de estudo. As próprias observações poderiam não ser independentes, na medida em que seria muito provável que

observações vizinhas apresentassem valores semelhantes, devido à existência de dependência espacial, a qual constitui, como já debatido, um dos efeitos que mais frequentemente ocorre em dados espaciais. As consequências potenciais de não se contar com o efeito da dependência espacial numa análise estatística são, por um lado, a subestimação dos erros e por outro lado, a sobrestimação dos níveis de significância estatística (Pfeiffer et al., 2008).

Entretanto, a presença do efeito espacial oposto (a heterogeneidade espacial) pode também comprometer a realização de testes de hipóteses, pois a variabilidade da superfície terrestre leva a que seja muito difícil determinar um “local médio”. Por exemplo, a comparação entre folhas de uma carta topográfica evidenciará características muito diferentes, mesmo se forem contíguas. Também as secções estatísticas de uma cidade não podem ser consideradas como uma amostra independente e aleatória de todas estas unidades existentes, na medida em que não são independentes entre si nem são definidas aleatoriamente. Desta forma, seria arriscado generalizar, tentando-se inferir as características de todas as secções a partir das características das secções existentes dentro de uma determinada subárea.

Assim, é conveniente que, antes da realização de testes de hipóteses sobre dados geográficos se tente perceber, por um lado, se existe uma população maior acerca da qual se deseja afinal inferir e, por outro lado, se os dados que se pretende utilizar constituem uma amostra aleatória e independente dessa população. Se a resposta for negativa em qualquer dos casos, não será apropriado aplicar testes de hipóteses. Entretanto, várias formas de lidar com este problema têm sido propostas, com diferentes graus de complexidade.

Uma das mais minimalistas (e provavelmente a menos satisfatória) consiste em reduzir as observações até se verificar o pressuposto de independência, o que, no caso dos dados pontuais, equivale a atingir uma distribuição espacial em que as observações se encontram suficientemente afastadas para poderem ser consideradas independentes. O outro extremo reside em renunciar aos testes de inferência, passando a utilizar-se estatísticas locais para observar as diferenças nos resultados da análise ao longo do espaço. Não obstante, é possível generalizar, através de um desenho experimental apropriado que permita conferir confiança ao processo de generalização, através da replicação do estudo num número suficiente de áreas distintas (Longley et al., 2010).

Uma forma do teste de hipóteses não ser adversamente afectado pela dependência espacial consiste em recorrer ao Método de Monte Carlo, sendo produzidas distribuições sucessivas da hipótese nula com base em aleatorizações sucessivas dos dados usados na análise. Os valores individuais da estatística de teste calculados em cada aleatorização são então utilizados em conjunto para representar a distribuição da hipótese nula, com a qual o valor observado na estatística de teste é comparado, sendo então calculado um valor de significância. Este método necessita geralmente de um grande número de aleatorizações, podendo ser computacionalmente intensivo quando aplicado sobre conjuntos de dados espaciais muito grandes ou em processos espaciais complexos. Alguns procedimentos estatísticos envolvem múltiplos testes com base no mesmo procedimento sobre os mesmos dados, por exemplo em técnicas de análise de clusters, tais como o *índice I de Moran* (Pfeiffer et al., 2008), que será introduzido no próximo subcapítulo.

1.7.3 Aplicações de Epidemiologia Espacial e Doenças Cerebrovasculares

Embora o esquema da figura 42, anteriormente apresentado, constitua uma forma intuitiva de introduzir a análise espacial de dados em Epidemiologia, pretende-se nesta parte, sempre que possível, fornecer exemplos de casos de estudo no âmbito da Epidemiologia do AVC que recorreram a técnicas representativas dos três grupos de métodos analíticos referidos na figura 42, a qual fornece aqui a estrutura para apresentação destes casos.

Desta forma, os métodos de visualização das características espaciais dos dados constituem geralmente o início de qualquer análise epidemiológica, permitindo a avaliação dos padrões de distribuição espacial do(s) fenómeno(s) em estudo, a detecção de erros evidentes nos dados, a formulação de hipóteses acerca dos agentes que possam estar a influenciar os padrões observados e ainda a apresentação dos resultados da análise, por exemplo através de mapas com a distribuição da doença.

Os métodos de visualização disponíveis separam-se entre técnicas para representar dados discretos (pontos e áreas, quando temos dados agregados) e dados contínuos (superfícies). Por simplicidade, serão aqui abordados apenas os métodos para dados

discretos, aconselhando-se a consulta de Pfeiffer et al. (2008) para a discussão relativa aos dados contínuos.

No que diz respeito à visualização de dados pontuais, o método mais simples consiste em cartografar estes dados usando as suas coordenadas cartesianas, num processo análogo a colocar pinos num mapa para referenciar localizações. Hu et al. (2008) relacionam a mortalidade por AVC com a poluição atmosférica, o rendimento e a distribuição dos espaços verdes, e fazem uso desta técnica para representar as fontes de emissões poluentes com e sem dados disponíveis (figura 45 esquerda e direita, respectivamente).

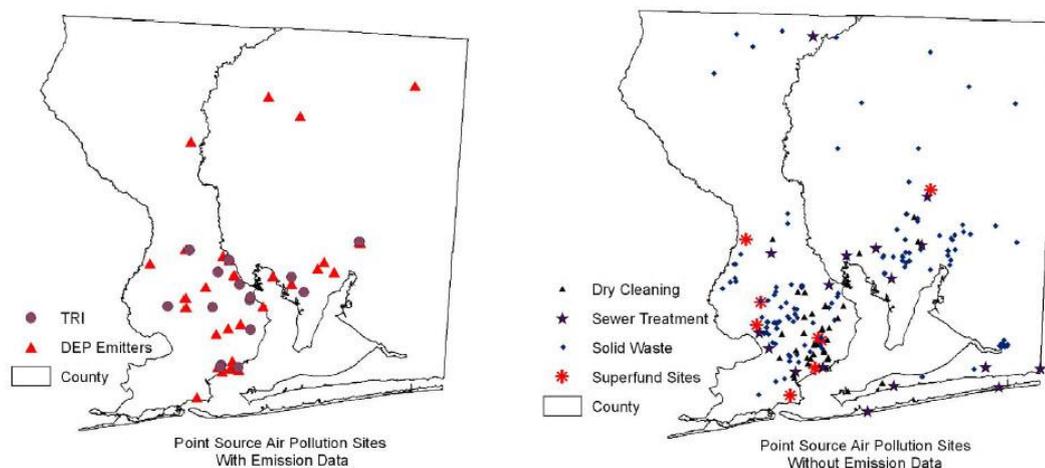


Figura 45 – Localizações pontuais de fontes de emissão de poluentes. Fonte: Hu et al. (2008).

Não obstante a simplicidade deste método, seja ele empregue para indicar fontes emissoras de poluentes ou a residência de pessoas afectadas por AVC numa determinada área, a sua utilização torna-se por vezes menos eficaz quando ocorre um grande número de fenómenos pontuais ao longo da área de estudo ou, inversamente, vários eventos na mesma localização.

Existem, não obstante, algumas alternativas que permitem lidar com estes problemas. Quando o objectivo é avaliar a densidade de pontos numa área, a presença de um grande número de pontos na mesma localização torna difícil esta avaliação. Uma alternativa consiste na utilização de métodos de *Kernel Smoothing* (Pfeiffer et al., 2008), que facilitam a avaliação de padrões, pois permitem visualizar simultaneamente a distribuição espacial e a densidade dos eventos. Um caso particular de kernel smoothing

consiste em recorrer à técnica de *Kernel Density*, a qual produz uma superfície de densidade através do cálculo da densidade de pontos numa vizinhança (geralmente designada largura de banda ou raio de busca) em redor de um ponto. Assim, conceptualmente, uma superfície curva suave (uma função de kernel) é ajustada sobre cada ponto, sendo cada ponto substituído pela sua respectiva função de kernel na qual a soma dos valores da superfície corresponde a uma unidade, e o valor da superfície é o mais alto na localização do ponto e vai diminuindo com a distância até zero, no limite do raio de busca (figura 46a). O cálculo finaliza-se com a soma das várias funções de kernel, obtendo-se uma superfície agregada, ou campo contínuo de densidade (Longley et al., 2005), como ilustrado na (figura 46b).

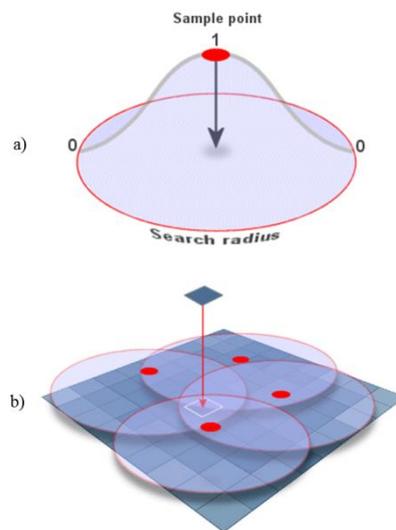


Figura 46 – Modelação de superfícies através de *Kernel Smoothing*. Fonte: Geography Hunter (2015).

Hu et al. (2008) recorreram a esta técnica para calcular superfícies da densidade da média anual das contagens diárias de tráfego ao longo da rede viária da área de estudo, como parte do processo de estudar a associação entre mortalidade por AVC e poluição atmosférica (figura 47).

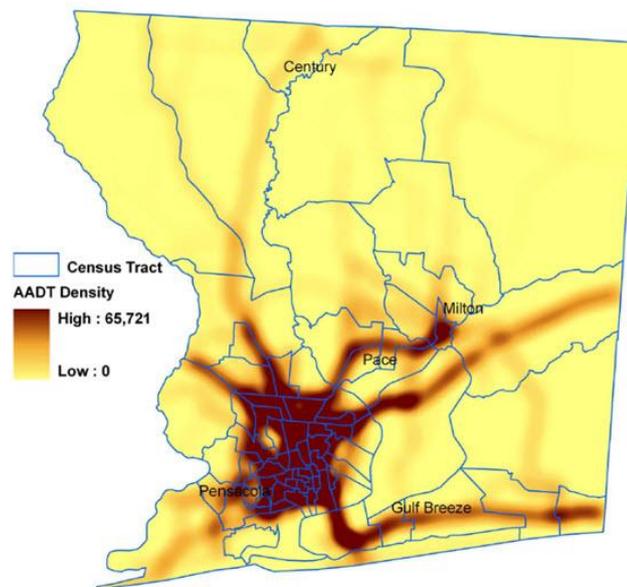


Figura 47 – Superfície raster da densidade da média anual das contagens diárias de tráfego obtida através de *Kernel Density*. Fonte: Hu et al. (2008).

Freqüentemente, ao invés de pontuais, os dados disponíveis encontram-se agregados (por exemplo, por questões que se prendem com o sigilo estatístico), no contexto de um desenho epidemiológico do tipo ecológico. O processo de agregação consiste em sumarizar um grupo de pontos representativos de dados individuais num só valor, que pode representar o seu total ou média, entre outros. O valor resultante deste processo é geralmente associado a uma localização espacial, em geral uma área discreta como os limites de um município ou outra unidade administrativa. Nos estudos de Epidemiologia Espacial, a forma de agregação mais comum consiste em contar ocorrências (tais como os casos de uma doença ou o valor de um agente ambiental associado a uma doença) para obtenção do total atribuível a cada área, podendo este ser depois apresentado em função de um determinado corte populacional, obtendo-se assim, por exemplo, a prevalência ou a incidência da doença por área administrativa. A forma mais comum de apresentar estes resultados consiste em recorrer a mapas *coropletos* (designação originada nos termos gregos *khoros*, significando “lugar” e *plethein*, que significa “preencher”). Assim, estes mapas apresentam a informação preenchendo cada área neles incluída com uma tonalidade de cor (ou através de tramas), representativa da magnitude da variável numa área (Pfeiffer et al., 2008). Num estudo ecológico relacionando a poluição atmosférica com a incidência e severidade de AVC isquêmico no sul da cidade

de Londres, Maheswaran et al. (2014) recorreram a mapas coropletos para representar a distribuição espacial de concentrações de poluentes previamente modeladas por área censitária (*census output area*) (ONS, 2014), como se pode observar na figura 48.

Apesar da popularidade destes mapas para ilustrar a distribuição espacial dos dados epidemiológicos, impõem-se algumas cautelas no seu uso, na medida em que podem ser afectados por três tipos de problemas que podem surgir, consoante o contexto em que o estudo se desenrola.

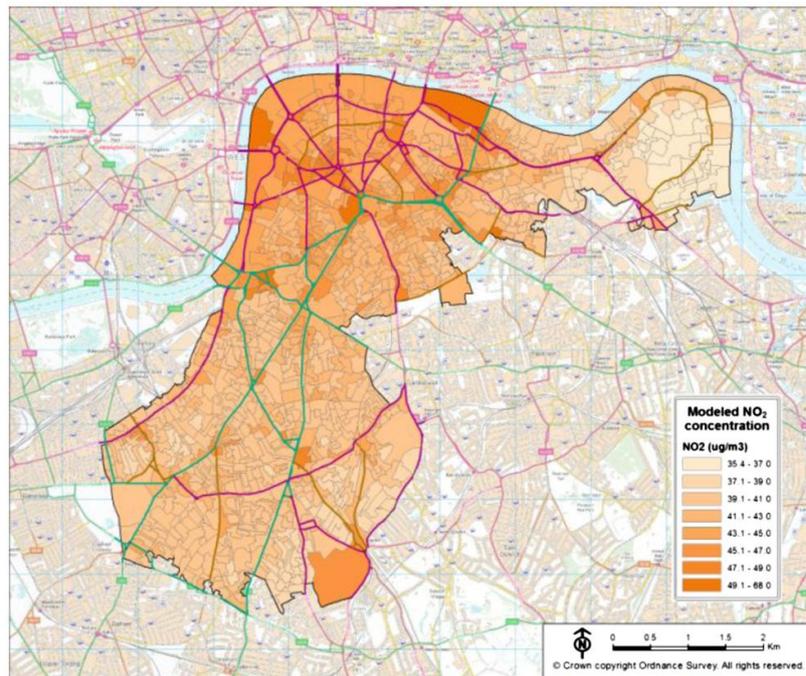


Figura 48 – Concentrações de dióxido de azoto, por áreas censitárias no sul de Londres.
Fonte: Maheswaran et al. (2014).

O primeiro problema tem a ver com a dimensão dos polígonos que compõem a área de estudo, pois nas situações em que as suas dimensões variam muito, aqueles com maior área têm tendência a dominar a visualização, podendo introduzir viés na interpretação do mapa (Monmonier, 1996). Um segundo problema, talvez mais importante, relaciona-se com a possibilidade dos padrões observados nas unidades estatísticas espaciais ao longo da área de estudo poderem resultar da distribuição espacial do atributo de interesse ou, por outro lado, serem induzidos pelas fronteiras dos polígonos (por exemplo quando representam os limites de áreas administrativas, como os municípios), situação conhecida como o Problema da Unidade de Área Modificável, ou MAUP

(Openshaw, 1984). Um terceiro problema, mais subtil, surge quando a distribuição da variável a ilustrar no mapa é altamente assimétrica, o que dificulta a sua representação através de escalas de cor com um número finito de tons (Pfeiffer et al., 2008).

Existem várias soluções alternativas para mitigar estes problemas. No caso da dimensão dos polígonos, uma das soluções prevê o recurso a cartogramas (Dorling, 1995), sendo que nestes mapas a área de representação é também distorcida proporcionalmente ao valor da variável a representar.

No caso da MAUP, uma das soluções consiste em reanalisar os dados a diferentes desagregações territoriais, embora na prática nem sempre seja possível obter dados muito desagregados (Arlinghaus et al., 1996); (Lawson e Williams, 2001). Finalmente, no caso de distribuições muito assimétricas, é possível transformar os dados, mas correndo o risco de comprometer a sua interpretação. Uma maneira de analisar mapas coropleto sem transformar os dados é recorrer a técnicas dinâmicas de análise exploratória de dados espaciais (ESDA), as quais conjugam formas tradicionais de visualização gráfica de dados (histogramas, caixas de bigodes ou outras) com o mapa coropleto. Entre as várias opções, a técnica de *brushing* permite seleccionar valores no gráfico que são interactivamente destacados no mapa (Haslett et al., 1991), permitindo uma análise menos enviesada da distribuição espacial do fenómeno. Várias ferramentas deste tipo podem ser encontradas no software GeoDa (Anselin, 2005), entre outros.

No que diz respeito à exploração, como indicado anteriormente, esta difere da mera visualização, na medida em que recorre a métodos estatísticos para tentar perceber se os padrões espaciais observados são ou não aleatórios, sendo que esta fase consiste, no essencial, em realizar análise de clusters espaciais.

Antes de mais, é importante indicar, quer nos estejamos a referir a técnicas de análise de clusters espaciais ou não espaciais, que as técnicas de análise de clusters consistem, de um modo geral, em ferramentas de análise exploratória de dados com o propósito universal de resolver problemas de classificação. Pretende-se assim ordenar dados representativos de algo que se pretende estudar (fenómenos, pessoas, objectos...) em grupos, designados clusters, sendo que estes grupos devem apresentar, por um lado, o máximo de homogeneidade interna (dentro do cluster) e por outro lado o máximo de heterogeneidade externa (entre clusters) (Salah et al., 2012); (Ahmad et al., 2013).

De certa forma, esta é uma extensão do processo iniciado na fase anterior de visualização, visto que permite já a identificação rápida de alguns padrões espaciais mais óbvios, e esses padrões podem ser classificados, de forma abrangente, como regulares, aleatórios ou clusterizados. Porém, algumas cautelas se impõem neste ponto, pois embora o termo *clustering* seja empregue para indicar a agregação espacial invulgar dos fenómenos epidemiológicos, existe o risco do padrão espacial percebido ser apenas resultante da distribuição da população em risco ou de outros factores de risco. Neste sentido, Elliott et al. (2001) propõem que se considere a doença espacialmente agregada quando se verifica que existe variação espacial residual no risco após as influências conhecidas terem sido tidas em conta.

Outra consideração importante consiste em classificar as técnicas de análise de clustering como específicas ou não-específicas, embora entre epidemiologistas estas técnicas sejam mais frequentes designadas como locais ou globais, respectivamente. Assim, os métodos globais são usados para avaliar se existe tendência para se verificarem clusters ao longo da área de estudo, mas sem os localizar. Por outro lado, os métodos locais destinam-se a detectar a localização e extensão dos clusters. Ambos pressupõem o recurso a testes de significância estatística para tentar obter alguma confirmação em ambas as situações. Por um lado, os métodos globais fornecem em geral uma só estatística, que mede o grau de clustering espacial, sobre a qual incide um teste de significância estatística, e neste caso a hipótese nula indica que o fenómeno não se verifica, admitindo-se antes a existência de aleatoriedade espacial. Por outro lado, nos métodos locais, o teste de significância incide sobre cada potencial cluster detectado (Pfeiffer et al., 2008).

Existe uma quantidade razoável de técnicas disponíveis para avaliação de clustering ou para a detecção de clusters espaciais, podendo ser encontrada uma discussão compreensiva em Rogerson e Yamada (2009).

Abordar-se-ão aqui as duas mais utilizadas, uma delas global, nomeadamente o índice I de Moran (também designado coeficiente I de autocorrelação de Moran) e as estatísticas locais de Moran, também designadas índices *LISA* (indicadores locais de autocorrelação espacial) (Anselin, 2005); (Longley et al., 2010).

O coeficiente I de autocorrelação de Moran é semelhante ao coeficiente de correlação de Pearson, e oferece uma forma de quantificar a similaridade entre os valores de uma

variável em estudo entre áreas nas quais esta variável foi medida e que se encontram relacionadas no espaço (Moran, 1950). A estatística I de Moran é dada pela seguinte fórmula:

$$I = \frac{n \sum_i \sum_j W_{ij} (Z_i - \bar{Z})(Z_j - \bar{Z})}{(\sum_i \sum_j W_{ij}) \sum_k (Z_k - \bar{Z})^2}$$

Nesta, Z_i pode ser o valor dos resíduos ($O_i - E_i$), por exemplo da taxa de mortalidade padronizada por idades ou do Rácio Padronizado de Mortalidade (SMR), associado a uma área, e W_{ij} representa a medida da proximidade das áreas i e j . Para definir as relações espaciais entre regiões é utilizada uma matriz de pesos espaciais, com o propósito de fazer com que, no cálculo do coeficiente, sejam atribuídos pesos mais elevados às regiões que se encontram mais próximas no espaço e inferiores às que se encontram mais distantes (Moran, 1950). O coeficiente I apresenta uma distribuição próxima da normal e um valor esperado de $-1/(N-1)$ quando não existe correlação entre valores vizinhos. Sendo N correspondente ao número de unidades espaciais areais na região de estudo, o valor esperado de I irá assim aproximar-se de zero à medida que N aumenta. Embora o valor do coeficiente I de Moran varie entre $+1$ e -1 , não está necessariamente balizado por estes limites, ao contrário do coeficiente de correlação de Pearson (Waller e Gotway, 2004). Assim, um valor de zero sinaliza a não existência de clustering (hipótese nula), um valor positivo indica a presença de autocorrelação positiva (agregação de áreas com valores similares) e um valor negativo sinaliza autocorrelação espacial negativa, ou seja, que as áreas vizinhas têm tendência para apresentar valores dissimilares da variável em estudo (Pfeiffer et al., 2008).

A figura 49 ilustra o índice I de Moran calculado em ambiente GeoDa e apresentado juntamente com o gráfico de dispersão de Moran, para o valor médio dos alojamentos em 51 estados americanos, obtendo-se um valor de 0.4011 que sinaliza a presença de autocorrelação espacial positiva, por outras palavras a tendência para a agregação espacial de valores altos ou baixos, respectivamente.

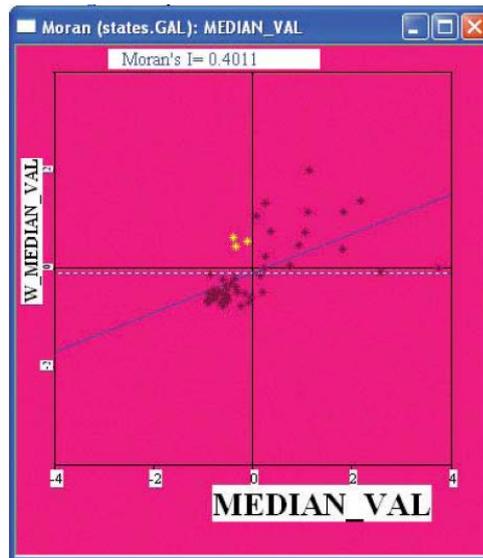


Figura 49 – Resultado do cálculo do Índice I de Moran. Fonte: Longley et al. (2005).

Coloca-se aqui a necessidade de confirmar que este valor não foi obra do acaso. Ou seja, se a distribuição espacial dos valores fosse aleatória, continuaria o índice a atingir um valor de 0,4011 ou próximo?

A solução proposta neste caso consiste em estimar os valores da média e desvio-padrão da estatística I de Moran através do rearranjo aleatório dos valores espacializados das variáveis, ilustrando a figura 50 o resultado obtido após a simulação de 999 rearranjos do valor médio dos alojamentos observado nos 51 estados. Do ponto de vista estatístico, assume-se aqui como hipótese nula que a distribuição dos valores pelos 51 estados ocorra de forma aleatória, sendo o valor de cada estado independente dos valores dos estados vizinhos. Assume-se também que a população corresponde ao conjunto de todos os arranjos possíveis, sendo o arranjo real dos valores uma amostra dessa população.

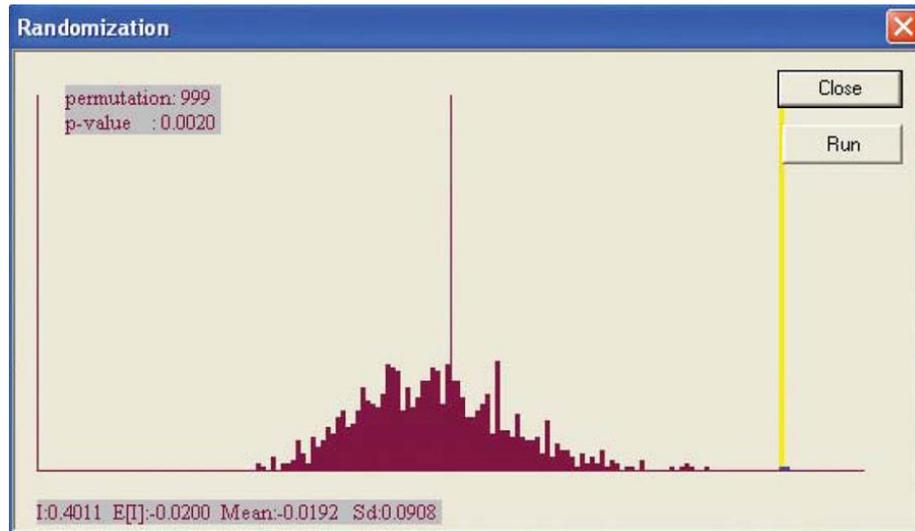


Figura 50 – Processo de aleatorização no cálculo do Índice I de Moran. Fonte: Longley et al. (2005).

O teste irá assim comparar o valor do índice (a estatística de teste obtida com a “amostra”) com a distribuição dos valores obtidos na hipótese nula. A linha amarela na figura representa o valor real, o qual teria muito pouca probabilidade de surgir num arranjo aleatório, e reforça a tese da presença de autocorrelação espacial positiva nos dados. Assim, após a realização do teste, observa-se que o valor obtido (a linha amarela na figura 50) se encontra muito distante daqueles obtidos a partir dos arranjos espaciais produzidos aleatoriamente, o que parece reforçar a presença de dependência espacial, rejeitando-se a hipótese nula com um nível de significância inferior a 0.01 (ou seja, 99% de certeza). Desta forma, torna-se possível adaptar uma técnica convencional de teste de hipóteses à natureza particular dos dados espaciais (Longley et al., 2010).

Quanto aos índices LISA, este método detecta as ocorrências localizadas de autocorrelação espacial nos dados agregados recorrendo à decomposição do índice I de Moran em valores para cada unidade espacial areal englobada pela região de estudo. As estatísticas LISA para cada área são calculadas através da seguinte fórmula:

$$I_i = Z_i \sum_{j, j \neq i}^n w_{ij} Z_j$$

Onde Z_i e Z_j correspondem aos valores observados (estandardizados) e W_{ij} representa a matriz de pesos espaciais padronizados por linha (*row standardized*). Desta forma, estes indicadores permitem detectar clusters de valores semelhantes ou opostos à volta de uma determinada observação (Pfeiffer et al., 2008).

A este respeito, (Lanska e Peterson, 1995) recorreram a índices LISA na análise da variação espaçotemporal da mortalidade por AVC nos Estados Unidos entre 1970 e 1978 (figura 51), tendo reportado uma tendência para a agregação espacial dos estados com declínio mais rápido nas taxas médias de variação anual da mortalidade por AVC em homens de raça caucasiana, numa área que reúne cerca de 11 estados da região sudeste dos Estados Unidos da América e que é frequentemente designada como “Stroke Belt” (Howard et al., 1995); (Howard et al., 1997); (Palermo, 2013).

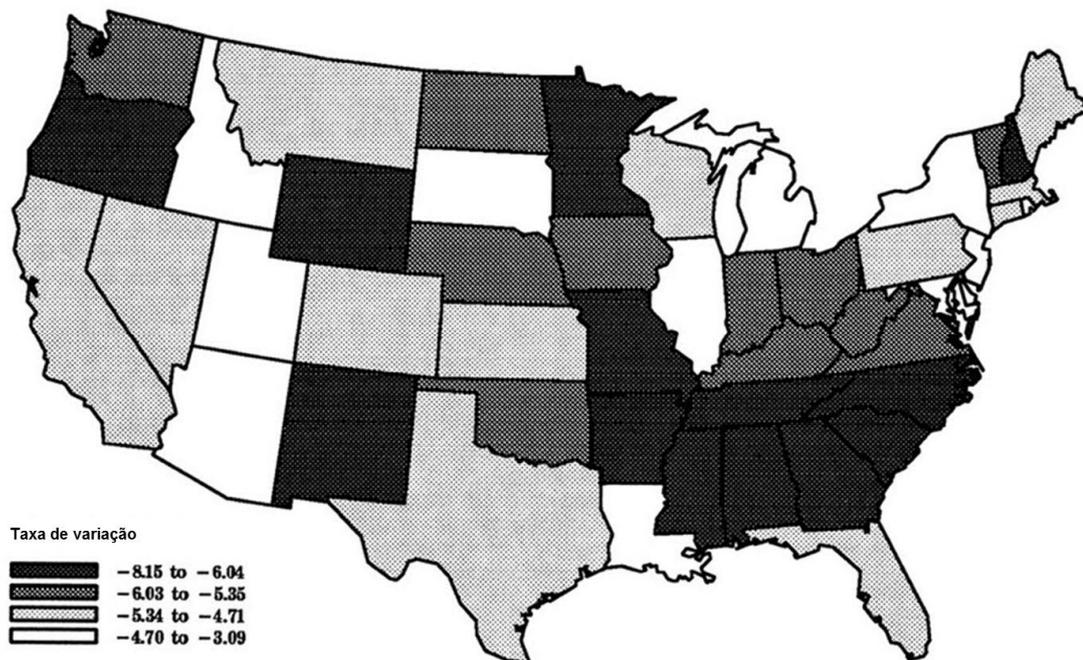


Figura 51 – Mapa da taxa média anual de variação nas taxas de mortalidade em homens caucasianos entre 1970 e 1978 nos Estados Unidos da América. Os estados com declínio mais rápido encontra-se agregados de forma significativa nos estados do sudoeste ($I=0.266$, $p<0.01$). Fonte: Adaptado de Lanska e Peterson (1995)

Num exemplo muito mais recente, e também no contexto dos Estados Unidos da América, Schieb et al. (2013) utilizaram também índices LISA para detectar clusters de municípios (*counties*) com valores altos ou baixos de taxas de hospitalização por AVC, após o que calcularam várias estatísticas (e.g. a média e mediana) dos valores de variáveis relativas à caracterização socioeconómica dos municípios incorporados em cada tipo de cluster e à sua rede de cuidados de saúde. Desta forma, este estudo teve como objectivo explorar a relação entre o risco de hospitalização devido a AVC e vários factores socioeconómicos e relativos à rede de cuidados de saúde, medidos ao nível do

município (e.g. o rendimento médio do agregado familiar ou a população abaixo do limiar de pobreza) e avaliar a evolução desta relação entre dois períodos temporais, na medida em que efectuou esta análise em 1995–1996 e em 2005–2006. A figura 52 ilustra a cartografia dos clusters num dos períodos do estudo.

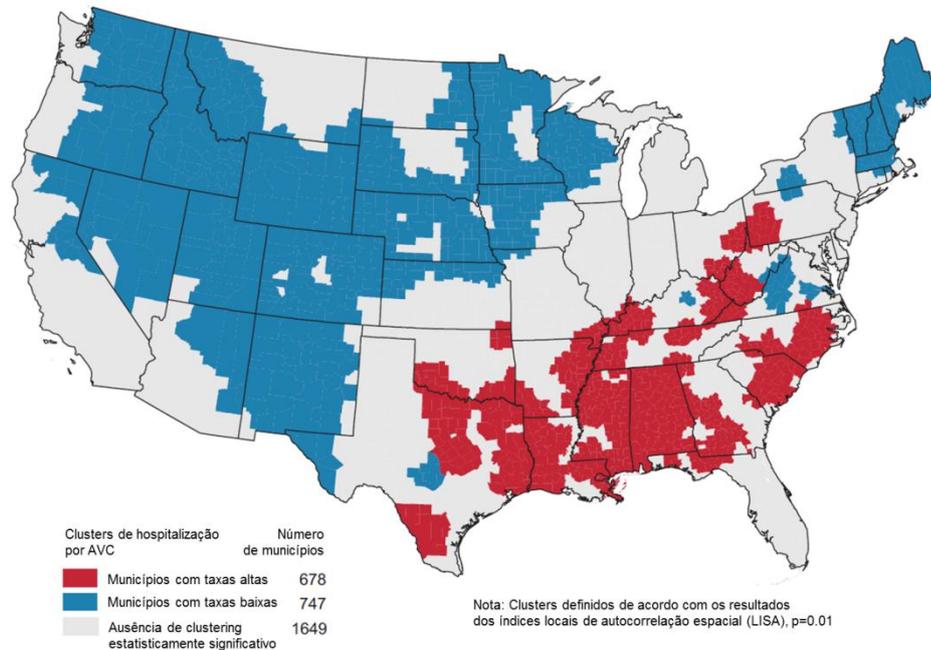


Figura 52 – Clusters das taxas de hospitalização por AVC em pessoas com 65 ou mais anos de idade no período 2005-2006. Fonte: Adaptado de Schieb et al. (2013)

É importante referir que, para além dos métodos de clustering espacial até aqui referidos e em geral empregues de forma univariada, existem também técnicas de análise de clusters que não lidam (pelo menos não directamente) com a componente espacial, e são geralmente empregues em situações de análise multivariada.

As formas de análise de clusters mais utilizadas dividem-se em técnicas não-hierárquicas (ou de optimização) e em técnicas hierárquicas, estas últimas mais frequentemente utilizadas.

Quanto às técnicas não-hierárquicas, estas apoiam-se simultaneamente num determinado critério de agrupamento de casos e na predefinição de um número de grupos (clusters). Desta maneira, a optimização (quer por minimização ou maximização do critério) levará a que cada caso venha a ser incluído num dos grupos predefinidos.

Entre estas técnicas, a de *K-means* (Hartigan e Wong, 1979); (Everitt et al., 2011) é a mais frequentemente utilizada.

As técnicas hierárquicas não exigem a predefinição de um número de grupos, embora seja necessário, tal como nas anteriores, a escolha de um determinado critério de agrupamento. As técnicas hierárquicas separam-se em divisivas e aglomerativas.

As técnicas divisivas começam por um único grupo de partida e dividem-no sucessivamente em subgrupos até chegarem ao ponto em que o número de grupos iguala o número de indivíduos, sendo geralmente menos utilizadas por serem mais intensivas em termos computacionais.

As técnicas aglomerativas funcionam de forma inversa, tendo como ponto de partida a definição de uma matriz de distâncias entre indivíduos, correspondendo a n grupos de partida (cada indivíduo é considerado nesta fase um grupo) e vão alocando sucessivamente os indivíduos a grupos (através do sucessivo relaxamento da proximidade absoluta, aquela que existe entre cada indivíduo e si próprio), até ser atingido um único grande grupo que aglomera todos os n elementos. Surge assim uma hierarquia, constituindo cada nível de agregação, um nível hierárquico (Mendes e Bação, 1998). Uma questão que se coloca nestas técnicas é o facto de não estarem, à partida, dotadas de mecanismos para lidar com a estrutura espacial dos dados, quando se pretende que os clusters obtidos tenham em conta a distribuição espacial do fenómeno em estudos. Uma forma de induzir este efeito é incorporar variáveis adicionais contendo a localização das observações. As questões relativas às técnicas hierárquicas e incorporação de efeitos espaciais serão desenvolvidas e operacionalizadas no capítulo 3. Um outro tipo de técnica mais recente baseia-se numa aplicação de redes neuronais designada *Self-Organizing Maps* ou *SOM* (Kohonen, 1990); (Hulle, 2012), tendo inclusive sido já proposta como a alternativa mais conveniente à técnica de *K-means*, desde que devidamente parametrizada (Bação et al., 2005). Esta técnica foi mais recentemente adaptada para lidar explicitamente com o carácter espacial dos dados geográficos, numa variante designada *GeoSOM*. Esta adaptação consiste em modificar a forma com as *Best Matching Units* - *BMU* (que constituem aqui essencialmente os clusters) são escolhidas. Desta forma, ao invés de ser atribuída uma ponderação igual a todas as variáveis, é reforçado o peso da localização espacial associada a cada observação (por exemplo com recurso à inclusão de coordenadas geográficas X e Y).

Isto tem como efeito que as unidades estatísticas espaciais a agregar (aqui designadas neurónios) que possuem coordenadas semelhantes terão maior tendência a ser consideradas BMUs (GeoSOM, 2015); (Henriques et al., 2012).

No que diz respeito à modelação, esta constitui a fase de investigação em Epidemiologia Espacial que se dedica à investigação de relações de causa-efeito, utilizando para o efeito todas as “peças” que se revelem necessárias à construção de modelos com o intuito de explicar ou prever padrões espaciais das doenças. Este é um processo pragmático, no sentido em que incorpora frequentemente tanto dados não-espaciais como espaciais. Por seu lado, a incorporação de dados espaciais torna necessária a utilização de modelos especificamente desenvolvidos para o efeito, ou obriga a que seja por vezes necessário operar adaptações nos métodos tradicionais de modelação para os tornar capazes de lidar com este tipo de dados.

Entre os métodos mais frequentemente empregues nesta fase encontram-se os modelos de regressão, os quais são aqui introduzidos enquanto meios de quantificar o efeito de um conjunto de variáveis explanatórias na distribuição espacial de um determinado evento de saúde. Serão assim introduzidas aqui as técnicas mais relevantes para os casos de aplicação desenvolvidos nos capítulos 2 e 3, bem como algumas questões que devem ser tidas em conta nestes estudos.

Um aspecto que deve ser tido em conta antes de mais tem a ver com o tipo de variável de resposta que se pretende investigar, a qual pode ser por exemplo uma contagem de casos de uma doença por áreas administrativas (que são entendidas aqui como unidades estatísticas espaciais) dentro da área de estudo, uma variável binária indicando apenas a presença ou ausência de doença num conjunto de localizações, ou ainda uma variável medida numa escala contínua. Uma discussão mais aprofundada sobre tipos de variáveis e escalas de medição em estatística pode ser encontrada em Cunha et al. (2007) e em Gravetter e Wallnau (2013).

O conhecimento da natureza da variável de resposta é fundamental pois condiciona o tipo de técnica de regressão que deve ser utilizada (em geral) e as opções técnicas disponíveis para lidar com a presença de efeitos espaciais (em particular).

Passando agora às técnicas de regressão apropriadas para cada tipo de variável de resposta, quando a variável de resposta (ou dependente) é medida numa escala contínua,

a técnica mais frequentemente utilizada é a Regressão Linear. Esta técnica obedece à seguinte formulação:

$$\mu_i = \beta_0 + \beta_1 X_{1i} + \dots + \beta_m X_{mi} + \varepsilon_i$$

Desta forma, a regressão linear permite que o valor médio de uma variável de resposta μ_i seja função de m variáveis explicativas (também designadas como predictoras, covariáveis ou independentes). Assim, o termo X representa a matriz ($m \times i$) de variáveis explicativas, o termo β_0 uma constante que assinala o valor de μ_i quando $X = 0$ (também designada termo de intercepção). As constantes β_1, \dots, β_m (geralmente designadas coeficientes de regressão) determinam a quantidade de mudança na variável explicativa em resposta a cada mudança unitária em cada uma das variáveis explicativas (Pfeiffer et al., 2008). Finalmente, o termo ε_i é geralmente designado termo de erro (ou componente aleatório ou estocástico), na medida em que se assume, no modelo clássico de regressão linear, que este termo não depende das covariáveis, ou seja, não pode ser explicado por estas. Assim, após a estimação, este termo corresponderá aos *resíduos* da regressão, podendo estes, por sua vez ser entendidos como estimativas dos erros ε_i (Fahrmeir et al., 2013). Este modelo de regressão assume cinco pressupostos essenciais, nomeadamente que: 1) para todos os valores de μ_i deve existir um valor correspondente de X (pressuposto de existência); 2) o valor de μ_i num dado ponto não é afectado pelo valor de μ_i em qualquer outro ponto (pressuposto de independência); 3) a relação entre μ_i e X deve ser aproximadamente linear (pressuposto de linearidade); 4) A variância de μ_i ao longo da linha de regressão estimada é igual para todos os valores de X (pressuposto de homocedasticidade) e 5) os resíduos ε obedecem a uma distribuição normal com média 0 (pressuposto de normalidade) (Pfeiffer et al., 2008).

E é justamente ao nível destes pressupostos que se colocam problemas quando os fenómenos em estudo têm expressão espacial. De facto, na presença de dados espacializados, o pressuposto de homocedasticidade pode ser posto em causa pela presença de não-estacionaridade espacial, uma vez que estrutura do processo pode não ser constante ao longo do espaço da área de estudo, sendo possível observar, representando o processo num mapa, que a estrutura da variância deriva (*drift*) ao longo da área de estudo. Uma forma mais elaborada de identificar a presença de não-

estacionaridade espacial nas variáveis explicativas consiste em recorrer a um modelo de regressão geograficamente ponderada (*GWR*) (Longley e Tobón, 2003). Este modelo será referido mais à frente neste texto.

Também o pressuposto de independência das observações pode ser posto em causa se se verificar a presença de dependência espacial, pois a similaridade (ou diferença) entre observações espacialmente próximas estará funcionalmente relacionada com a sua proximidade através de um processo activo. A presença de dependência espacial pode ser detectada através de análise de clusters, recorrendo-se ao cálculo dos índices LISA sobre os resíduos de uma regressão linear. Caso seja detectada a presença de autocorrelação espacial, existem duas formas possíveis de incorporar o efeito da dependência espacial no modelo clássico de regressão linear, passando este a designar-se um modelo de regressão espacial. Estas duas técnicas são o modelo espacial autoregressivo (*SAR*), mais conhecido como modelo *Spatial Lag* e o modelo *Spatial Error* (*SEM*), e a diferença fundamental entre os dois reside em partirem de diferentes pressupostos, no que respeita ao processo espacial que exerce influência sobre a estrutura espacial dos dados.

Quanto ao modelo spatial lag, este assume que se encontra presente um processo de difusão ou escolha através do qual a variável dependente é activamente influenciada pelos seus vizinhos. Desta forma, a dependência espacial pode ser formalmente modelada como um modelo autoregressivo espacial, sendo incorporada no modelo uma variável correspondente que induz um desvio (*lag*) espacial sistemático, como se observa na respectiva formulação:

$$\mu_i = \beta_0 + \rho W y + \beta_1 X_{1i} + \dots + \beta_m X_{mi} + \varepsilon_i$$

Como se pode observar, a única diferença entre este e o modelo de regressão linear tradicional reside no termo $\rho W y$, no qual ρ representa a medida de autocorrelação espacial, funcionando como o coeficiente espacial da regressão, e $W y$ representa uma matriz de pesos espaciais análoga à que é empregue nos métodos de clustering espacial de Moran previamente mencionados neste texto.

No que respeita ao modelo spatial error, este assume que existe estrutura espacial no termo de erro, examinando a autocorrelação espacial entre os resíduos das áreas

adjacentes. Essa autocorrelação é assumida neste modelo como resultante de algo negativo, como por exemplo uma possível variável omitida que apresenta clustering espacial. Desta forma, a única diferença em termos de formulação entre este e o modelo de regressão linear reside na estrutura do termo de erro, o qual irá incorporar os efeitos espaciais. Assim,

$$\varepsilon = \lambda W\varepsilon + \xi$$

onde ε constitui um vector de termos de erro, espacialmente ponderados através da matriz de pesos espaciais, λ representa o coeficiente de erro espacial da regressão e ξ consiste num vector de termos de erro não correlacionados. Desta forma, na eventualidade de não existir autocorrelação espacial entre os erros, λ será igual a zero (Anselin, 2001); (LeSage e Pace, 2009). Em termos de aplicação destas técnicas, um exemplo com interesse pode ser encontrado no estudo de Sparks e Sparks (2010), focado na utilização de modelos espaciais autoregressivos no estudo da associação entre a mortalidade ao nível dos municípios nos Estados Unidos da América e vários determinantes socioeconómicos.

Retomando a questão das técnicas de regressão apropriadas para cada tipo de variável de resposta, quando esta consiste numa contagem do número de eventos que ocorrem numa população de uma determinada dimensão, ou numa contagem do número de eventos em relação ao número de pessoas-ano em risco, é razoável assumir que estas contagens seguem uma distribuição de Poisson, particularmente quando se trata de doenças não-contagiosas ou raras (Pfeiffer et al., 2008). Como tal, neste caso uma técnica apropriada será o recurso a uma Regressão de Poisson ou, em alternativa, a uma regressão baseada na distribuição Binomial Negativa. O modelo de Regressão de Poisson pode ser descrito pelas seguintes equações:

$$\ln(\lambda_i) = \sum_{k=0}^K \beta_k x_{ik}$$

(1)

$$P(Y_i = y_i) = \frac{e^{-\lambda_i} \lambda_i^{y_i}}{y_i!}$$

(2)

A equação 1 ilustra uma regressão que relaciona o logaritmo natural da média ou número esperado de eventos para cada caso i , $\ln(\lambda_i)$, com a soma dos produtos para cada

covariável, x_{ik} , multiplicada pelo coeficiente de regressão β_k (sendo β_0 uma constante multiplicada por 1 para cada caso). Com a equação 2, pretende-se indicar que a probabilidade de Y_i (resultado esperado no caso i) segue uma distribuição de *Poisson* (correspondente ao lado direito da equação) para a média das contagens da Equação 1, λ_i . Desta forma, a distribuição esperada das contagens de casos, bem como a correspondente distribuição dos resíduos da regressão, dependem da média das contagens ajustada, λ_i . O papel do logaritmo natural na equação 1 é comparável à transformação logarítmica da variável dependente que acontece frequentemente na análise de taxas obtidas a partir da agregação de casos. Em ambos os casos, os coeficientes de regressão provocam diferenças proporcionais nas taxas (Osgood, 2000). Embora tanto os modelos de regressão de Poisson como Binomial negativo sejam adequados à análise de dados de contagem, estes dois modelos diferem no que respeita aos seus pressupostos em relação à média e à variância da variável dependente. Desta forma, enquanto os modelos de Poisson assumem que a média e a variância são iguais (equidispersão), os modelos binomiais negativos não assumem esta situação e tentam assim corrigir para a presença de sobredispersão nos dados, entendendo-se que existe uma situação de sobredispersão quando a variância é superior à média. Este fenómeno pode constituir um problema na medida em que pode fazer com que os erros padrões das estimativas sejam subestimados, por outras palavras, que a variável explicativa pareça ser significativa quando na verdade não o é (Piza, 2012); (Hilbe, 2007).

Para tornar possível a capacidade de lidar com a sobredispersão, a distribuição binomial baseia-se em dois parâmetros, na medida em que possui um parâmetro adicional em relação à de Poisson que permite ajustar a variância de forma independente da média. Desta forma, ao contrário da distribuição de Poisson, que é totalmente caracterizada pela sua média μ , a binomial negativa é simultaneamente função de μ e do parâmetro auxiliar α . Consequentemente, a sua média continua a ser μ , mas a sua variância passa a ser $\mu(1 + \mu\alpha)$. Obviamente, se $\alpha \rightarrow 0$, esta distribuição transforma-se numa distribuição de Poisson, razão porque alguns autores afirmam que a distribuição de Poisson constitui um caso particular da distribuição binomial negativa (Hilbe, 2007).

Existem vários exemplos de aplicação na área de estudo do AVC. Maheswaran e Elliott (2003) recorreram ao modelo de regressão de Poisson para investigar a associação entre

a mortalidade por AVC e a residência na proximidade de estradas principais no Reino Unido. Mais recentemente, Maheswaran et al. (2014) utilizam também esta técnica para investigar a associação entre os efeitos da poluição atmosférica e a severidade dos casos de AVC isquémico.

Embora permitam lidar com tipos de dados de contagem ou contínuos (não serão aqui abordados modelos para variáveis binárias, isto é modelos de Regressão Logística, acerca dos quais se sugere a consulta de Kleinbaum e Klein (2010), os modelos de regressão abordados até agora constituem modelos globais, nos quais se assume que a variação é igual em todas as localizações. Todavia, pode dar-se o caso de um modelo global não conseguir representar correctamente a variação numa dada localização. Outra característica dos modelos globais é o facto de fazerem uso de todo o conjunto de dados, enquanto os modelos locais são geralmente definidos por recorrerem apenas a um determinado subconjunto dos dados. A variação local de certas propriedades, algumas já abordadas neste texto, faz surgir a necessidade de modelos locais. Um exemplo reside na maior parte dos dados espaciais apresentarem dependência espacial. Outra situação prende-se com a maior parte destas análises (espaciais) estarem sujeitas ao problema da MAUP, sendo que os resultados dependem da divisão do espaço, como já referido anteriormente neste texto. Outra questão ainda tem a ver com o facto de ser difícil assumir estacionaridade em qualquer processo que se desenrole ao longo do espaço geográfico, sendo mais frequentes as situações de não-estacionaridade, nas quais a médias e a variância variam marcadamente entre localizações, sendo o processo claramente não-estacionário (Lloyd, 2010).

As formas locais de análise espacial têm, por estes motivos, vindo a ganhar importância, e entre estas encontra-se a Regressão Geograficamente Ponderada (GWR). Esta consiste numa técnica local de estatística espacial particularmente apropriada à análise da não-estacionaridade espacial, entendida como a situação na qual as medições das associações entre variáveis variam entre localizações (Fotheringham et al., 2002). Assim, ao contrário da regressão convencional, que produz uma única equação de regressão que sumariza as relações globais entre variáveis explicativas e dependentes, esta técnica gera dados espaciais que expressam a variação espacial dos relacionamentos entre variáveis. Desta forma, os mapas gerados a partir destes dados

assumem um papel central na exploração e interpretação da não-estacionaridade espacial.

Em termos de processo, o modelo GWR, em vez de calibrar uma única equação de regressão (tal como o modelo de regressão linear convencional formulado anteriormente neste subcapítulo), gera ao invés uma regressão em separado para cada observação. Desta forma, cada equação é calibrada através de um conjunto distinto de ponderações, obtidas a partir de observações do mesmo conjunto de dados. Assim, cada equação GWR obedece à seguinte formulação:

$$\hat{y}_i = \beta_0(u_i, v_i) + \sum_k \beta_k(u_i, v_i) x_{ik} + \varepsilon_i$$

Na qual o \hat{Y}_i estimado representa o valor estimado para a variável independente na observação i , β_0 representa o termo de intercepção e β_k o parâmetro estimado para a variável K , X_{ik} é o valor na observação i da última variável K , ε_i representa o termo de erro. A diferença para a equação de regressão linear reside, realmente, em (u_i, v_i) , que regista as coordenadas de localização da observação i (Fotheringham et al., 2002). O registo da localização de cada observação é o ponto de partida para o pressuposto fundamental nesta técnica de que cada observação mais próxima de uma outra tem uma maior influência nas estimativas dos parâmetros dessa outra que as observações mais distantes. O peso atribuído a cada observação baseia-se numa função de diminuição da distância centrada na observação i , correspondendo esta função a um kernel espacial, tal como ilustrado graficamente na figura 53. Fazendo aqui um aparte, o termo kernel refere-se a uma função de padronização da ponderação (ou peso), nomeadamente uma função de ponderação que vai considerar $h = 1$. O kernel determina assim a forma da função de ponderação, sendo o parâmetro h denominado largura de banda ou, em alternativa, constante de suavização (Zucchini, 2003).

No caso específico dos dados que representam áreas, a distância entre observações corresponde à distância entre os centróides dos polígonos que representam essas áreas. Quanto à função de diminuição da distância, esta pode assumir uma variedade de formas, sendo modificada para incluir uma regulação da largura de banda (*bandwidth*) correspondente à distância a que o peso rapidamente se aproximará de zero.

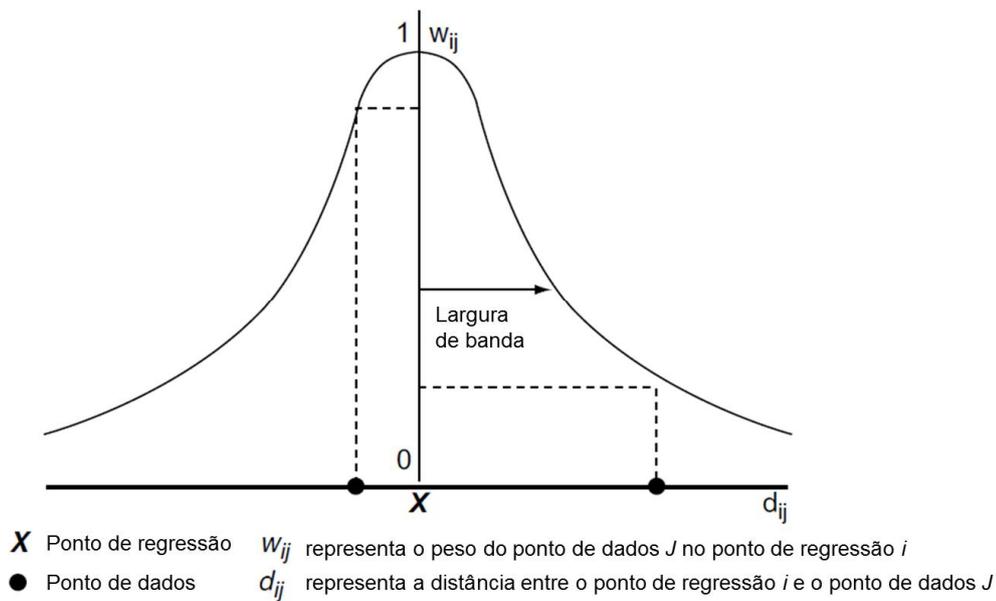


Figura 53 – Mecanismo de um kernel espacial. Fonte: Adaptado de Fotheringham et al. (2002).

Tal como ilustrado na figura 53, a largura de banda constitui a medida que controla o ritmo de diminuição do peso em função da distância (*distance-decay weighting function*), indicando a intensidade da suavização dos resultados locais de calibração. Desta forma, *kernels* espaciais com uma largura de banda mais pequena darão origem a uma diminuição mais abrupta do peso em função da distância, produzindo superfícies mais grosseiras do que quando são empregues larguras de banda maiores.

Por exemplo, num kernel Gaussiano, a função de ponderação corresponde à seguinte formulação:

$$w_{ij} = \exp \left[-\frac{1}{2} \left(\frac{d_{ij}}{b} \right)^2 \right]$$

Na qual b representa a largura de banda, w_{ij} diz respeito à ponderação e d_{ij} corresponde à distância entre as observações i e j .

A largura de banda, por sua vez, pode ser escolhida manualmente pelo operador ou otimizada através de uma variedade de algoritmos, sendo muito frequente o recurso à minimização do valor do Critério de Informação de Akaike (AIC), método que apresenta a vantagem de tomar em conta que os graus de liberdade podem variar entre modelos centrados em diferentes observações. O operador pode ainda escolher uma largura de banda fixa que é usada para cada observação ou uma largura de banda variável que se expande em áreas em que as observações se encontram menos

concentradas e encolhe em áreas de maior densidade de observações (Fotheringham et al., 2002). A figura 54 ilustra graficamente aplicações de GWR com o ajustamento de um kernel espacial aos dados utilizando uma largura de banda fixa (figura 54a) ou adaptativa (figura 54b).

Porque a regressão é calibrada independentemente para cada observação, torna-se possível obter uma estimativa separada dos seus parâmetros, bem como um valor de t e da qualidade do ajuste para cada observação. Desta forma, estes valores podem ser mapeados, permitindo ao analista interpretar visualmente a distribuição espacial da natureza e da força dos relacionamentos entre variáveis explicativas e dependentes (Mennis, 2006). É importante referir que a GWR pode obedecer a vários modelos de regressão. Por exemplo, a versão 4 da aplicação dedicada GWR (GeoDa Center, 2015) permite, consoante a natureza dos dados, adoptar como modelo as distribuições de Poisson, Gaussiana ou Logística. Mais recentemente, alguns autores referem já soluções baseadas na distribuição binomial negativa (Silva e Rodrigues, 2013), embora esta opção não esteja ainda implementada no software dedicado atrás referido.

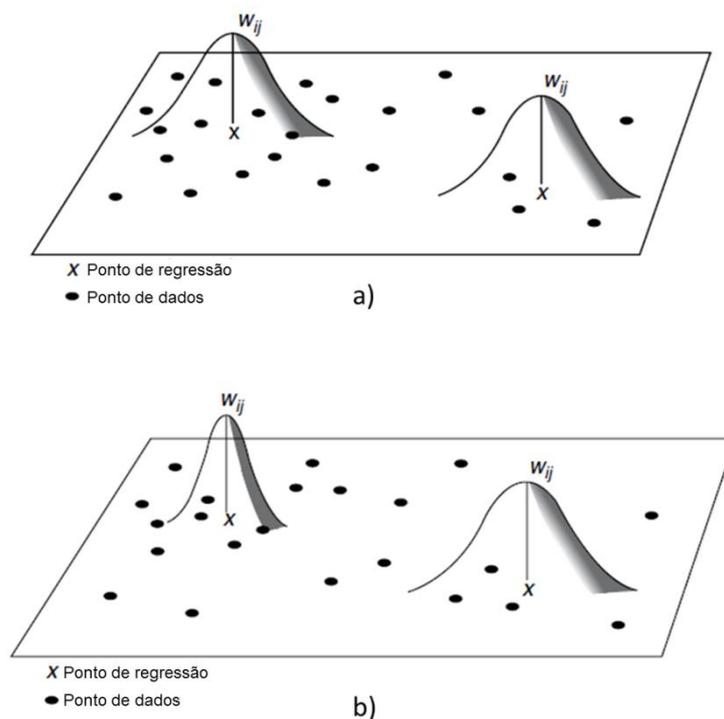


Figura 54 – *Kernels* espaciais com larguras de banda fixa (a) e adaptativa (b).
 Fonte: Adaptado de Fotheringham et al. (2002).

Embora exista já um número razoável de exemplos de estudos focados noutras doenças (St-Hilaire et al., 2010); (Cheng et al., 2011); (Helbich et al., 2012); (Weisent et al., 2012), na aplicação desta técnica na área de estudo do AVC registam-se ainda relativamente poucos estudos. Balamurugan et al. (2013) utilizaram uma metodologia baseada na GWR para investigar a associação entre factores de risco socioeconómico e mortalidade por AVC, tendo concluído que factores como a pobreza e o nível de escolaridade actuam como impulsionadores das disparidades de mortalidade por AVC entre áreas censitárias no estado norte-americano do Arkansas. Soljak et al. (2011) recorreram também a esta técnica para estudar a variação espacial de situações de subdiagnóstico de vários tipos de doenças cardiovasculares (entre as quais AVC) ao longo do território do Reino Unido. Nakaya et al. (2005) demonstram o uso desta técnica no estudo da associação entre mortalidade geral em pessoas em idades laboral, tipo de ocupação profissional e desemprego, tendo os resultados sinalizado disparidades geográficas estatisticamente significativas nestas associações ao longo da área metropolitana de Tóquio. Este autor afirma, a este propósito, que a aplicação de modelos de regressão globais neste caso poderia levar a resultados enganadores.

Finalmente, é importante referir que os modelos de regressão aqui referidos representam apenas uma parte das opções disponíveis, existindo outras alternativas que não serão aqui abordadas, tais como os modelos baseados em estatística bayesiana, entre os quais os modelos hierárquicos bayesianos, tendo já surgido na área do AVC alguns estudos empregando estes modelos. A título de exemplo, Hu et al. (2008) recorrem a um modelo deste tipo para investigar associações entre poluição atmosférica, nível de rendimentos e proximidade residencial a espaços verdes no estado norte-americano da Florida. Mais recentemente, Odoi e Busingye (2014) utilizaram também este tipo de modelos para investigar a associação entre disparidades geográficas ao nível dos factores socioeconómicos entre pequenas áreas censitárias (*census tracts*) e a mortalidade por AVC.

2. Análise espaçotemporal da relação entre factores socioeconómicos e acidente vascular cerebral na população de Portugal Continental abaixo de 65 anos de idade

2.1 Introdução

O Acidente Vascular Cerebral (AVC) representa, globalmente, um problema de saúde pública de grande magnitude, quer em termos de morbilidade como de mortalidade. Englobando o conjunto das doenças cerebrovasculares entre os códigos I60 a I69 da Classificação Internacional das Doenças e Problemas relacionados, 10^a revisão (ICD-10, 2015), esta doença foi em 2010 classificada como a segunda maior causa de morte, bem como a terceira causa de DALYs em todo o Mundo (Krishnamurthi et al., 2013), ocorrendo com mais frequência em adultos de meia-idade ou em idosos (WHO, 2014a). Nesse mesmo ano de 2010, identificaram-se 16,9 milhões de pessoas tendo sofrido um primeiro episódio de AVC e 5,9 milhões de mortes relacionadas com a doença. Esta foi também responsável por 102 milhões de anos de vida perdidos ajustados pela incapacidade (DALYs) e 33 milhões de sobreviventes a um episódio. Mais ainda, verificou-se um aumento estatisticamente significativo destes indicadores entre 1990 e 2010 (Krishnamurthi et al., 2013).

O risco de AVC, definido como a probabilidade de se sofrer um AVC (em termos de morbilidade) ou morrer em consequência de um (mortalidade), tem apresentado padrões diversos no que respeita à sua distribuição geográfica, não só entre países mas também entre regiões do mesmo país (Cox et al., 2006); (Addo et al., 2012); (Krishnamurthi et al., 2013). Em Portugal, por exemplo, o risco de morte em consequência de um AVC era, na população total em 2011, de 61,9 óbitos por 100.000 habitantes (DGS, 2013). Todavia, em 2012 os valores por NUTs II (Eurostat, 2014) variavam entre 53 óbitos na NUT Algarve e 66,7 óbitos na NUT Norte, por 100.000 habitantes (INE, 2014). As razões subjacentes a estas variações regionais não são ainda conhecidas. Consequentemente, existe interesse em identificar os determinantes destas disparidades geográficas, quer para o risco de morte quer para o risco de ocorrência da doença, com o propósito de desenvolver esforços de prevenção e mitigação.

O AVC é uma doença multifactorial e os seus determinantes podem ser discutidos tanto a nível individual como ecológico. Existe actualmente muito conhecimento sobre os determinantes individuais das doenças cardiovasculares em geral, bem como sobre o AVC em particular, tanto no que respeita aos modificáveis como não-modificáveis (Sacco et al., 1997). Entre os não-modificáveis, os mais relevantes são a idade, o sexo e os de origem genética. A hipertensão é classificada como a situação de maior risco entre os determinantes modificáveis, embora seguida de perto pelo fumo de tabaco, fibrilação atrial e inactividade física (Brainin e Heiss, 2009). No entanto, numa perspectiva abrangente de Saúde Pública, estes determinantes individuais constituem geralmente a fase final de uma cadeia de eventos (Labarthe, 2011). Logo, existe uma necessidade de enfoque na “causa das causas” (Marmot e Wilkinson, 2005), ou seja, de se investigar e actuar no início de uma cadeia de eventos que pode ser longa e complexa. No que respeita às doenças cardiovasculares em geral (e aplicável também ao AVC), esta cadeia tem normalmente início numa conjugação de situações desfavoráveis de carácter ambiental e social, que contribuem para o aumento de comportamentos nocivos na população. Estes comportamentos, com destaque para as dietas desequilibradas e a inactividade física conduzem, por seu lado, ao desenvolvimento de factores de risco individuais como a hipertensão e obesidade, entre outros (Labarthe, 2011). Colectivamente, esta conjugação de situações ambientais e sociais desfavoráveis pode ser entendida como o estatuto socioeconómico de um indivíduo e compreende uma classe diferente de determinantes (Cox et al., 2006), que incluem o nível de escolaridade (Avendaño et al., 2004); (Lambert et al., 2013); (Wu et al., 2013), ocupação profissional (Mendes, 1988); (Franks et al., 1991); (Hayashi et al., 1996); (Kunst et al., 1998); (Cesana et al., 2001), rendimento e posse de bens materiais (Engström et al., 2001); (Li et al., 2008). Os determinantes ambientais incluem, adicionalmente, num sentido mais restrito, a exposição ao ar poluído e a distância a uma unidade de cuidados de saúde como um hospital (Busingye et al., 2011); (Labarthe, 2011). Apesar dos elos de ligação entre estatuto socioeconómico e ocorrência de AVC terem vindo a ser identificados, a medida exacta em que este estatuto contribui para esta doença permanece ainda controversa (Cox et al., 2006); (Addo et al., 2012). Consequentemente, compreender a associação causal entre estatuto socioeconómico e AVC contribui para a correcta

aplicação e avaliação das intervenções levadas a cabo por profissionais de Saúde Pública e decisores políticos.

A relação entre o estatuto socioeconómico e o risco de AVC é certamente complexa, e um dos aspectos desta complexidade relaciona-se com a sua variabilidade geográfica. Em 2010, a incidência por país (em casos por 100000 pessoas/ano) variou amplamente entre países e seus respectivos níveis de rendimento (por exemplo 60 casos no Kuwait e 504 na Lituânia). Também em 2010, quando se considera o valor absoluto de novos casos em todo o Mundo, a maioria (69%) dos 16,9 milhões ocorreu em países de baixo e médio rendimento. Adicionalmente, entre 1990 e 2010 ocorreu um aumento global e significativo de 25% na incidência em pessoas com idades compreendidas entre os 20 e os 64 anos, tendo os países de baixo e médio rendimento sido responsáveis por uma parte substancial deste aumento. Todavia, a mortalidade apresenta um cenário mais complexo, uma vez que as taxas padronizadas apresentam um decréscimo global e significativo de 25%, tanto em países de alto como baixo rendimento. Inversamente, o número absoluto de óbitos aumentou em 26% no mesmo período (Krishnamurthi et al., 2013).

A partir do momento em que se reconhece a existência de disparidades geográficas nas associações entre os determinantes relacionados com o estatuto socioeconómico e a ocorrência de AVC, os modelos de regressão utilizados para medir estas associações precisam de contar com o facto dos coeficientes de regressão poderem também variar espacialmente. As abordagens locais de modelação permitem aos investigadores estimar com maior rigor as relações entre os determinantes e o risco da doença, dado permitirem a estimação de coeficientes de regressão para cada localização na área de estudo. Entre estas abordagens locais, as técnicas de modelação baseadas na Regressão Geograficamente Ponderada (GWR) calculam coeficientes de regressão locais, permitindo que as estimativas da associação entre doença e variáveis explicativas variem espacialmente, e disponibilizando assim uma estratégia de modelação mais flexível que os métodos globais (Weisent et al., 2012). Com isto em mente, vários estudos implementaram já este método, com o objectivo de melhorar a nossa compreensão dos determinantes das disparidades geográficas na saúde (St-Hilaire et al., 2010); (Cheng et al., 2011); (Helbich et al., 2012); (Weisent et al., 2012). A técnica de GWR tem suscitado alguma controvérsia, pois enquanto alguns autores a consideram

mais apropriada para a análise explanatória, outros defendem a flexibilidade deste tipo de modelo na investigação das relações que variam espacialmente (Helbich et al., 2012), no que constitui um dos principais objectivos da presente investigação.

Assim, o presente estudo teve como objectivos: (i) identificar os determinantes socioeconómicos das disparidades geográficas do risco de AVC ao nível do município em Portugal Continental; (ii) investigar se os coeficientes de regressão para a associação entre factores socioeconómicos e o risco de AVC demonstram variabilidade espacial; (iii) comparar o desempenho de vários métodos de regressão na modelação destas associações, nomeadamente as regressões binomial negativa e global, e ainda modelos locais de Poisson; e (iv) avaliar a variação espaçotemporal das associações entre determinantes e risco de AVC.

2.2 Materiais e métodos

2.2.1 Área de estudo e fontes de dados

Este estudo foi implementado no território continental português, englobando 275 municípios em 1991 e 278 em 2001, com um total populacional de 8,1 milhões em 1991 e 9,9 milhões em 2001, aproximadamente (INE, 2014). De acordo com a Comissão Europeia, os municípios são divisões administrativas locais, podendo também ser consideradas como unidades territoriais estatísticas (NUTs), anteriormente correspondentes ao nível 5 do sistema europeu (Eurostat, 2014a); (INE, 2014). Os municípios são relativamente heterogéneos quanto a população, variando em 2001 entre 1924 habitantes em Barrancos e mais de meio milhão em Lisboa (564657 habitantes). Apesar disto, estas unidades territoriais revelam alguma homogeneidade no que respeita às características da sua população, tais como o estatuto socioeconómico (INE, 2014).

Os dados cartográficos relativos a limites administrativos, mortalidade por AVC, população residente e factores socioeconómicos, todos ao nível municipal, foram obtidos no Instituto Nacional de Estatística (INE). As distâncias rodoviárias foram calculadas através da aplicação Google Earth (Google Earth, 2014). As áreas

predominantemente rurais no interior dos municípios foram obtidas a partir do Instituto de Apoio às Pequenas e Médias Empresas e à Inovação (IAPMEI, 2014).

Os dados relativos à mortalidade cobrem 2 períodos de 5 anos, o primeiro de 1 de Janeiro de 1992 a 31 de Dezembro de 1996, e o segundo de 1 de Janeiro de 2002 a 31 de Dezembro de 2006. Durante o primeiro período registaram-se 9601 óbitos por AVC, e 5852 óbitos no segundo.

2.2.2 Cálculo do risco de AVC, mapeamento e análise espacial inicial

O risco de morte por AVC foi inicialmente usado para avaliar a frequência e a distribuição espacial desta doença entre municípios. O risco de morte municipal foi calculado como o número de óbitos por AVC registados no município em pessoas com menos de 65 anos de idade, homens ou mulheres (todas as análises foram desenvolvidas para cada sexo em separado), ao longo de cada período de 5 anos, dividido pelas estimativas da população deste coorte residente no município, homens ou mulheres com menos de 65 anos de idade, nos anos intermédios desses períodos, 1994 e 2004. Devido ao potencial confundimento resultante da variação de estruturas etárias entre municípios, o risco de morte por AVC foi padronizado pela idade, recorrendo ao método de padronização indirecto (Siegel, 2012) e usando a população total de Portugal Continental como população padrão. Este procedimento assegurou que as diferenças na distribuição geográfica do risco de morte por AVC não seriam afectadas pelas diferenças geográficas na distribuição da população por idades. O valor padronizado resultante para cada um dos 2 períodos de 5 anos foi assim apresentado como o número de óbitos por AVC por 100000 habitantes.

Foram produzidos mapas coropletos para as variáveis socioeconómicas e para o risco de morte por AVC padronizado por idade, para uma avaliação visual comparativa da distribuição espacial destas variáveis. Todos os mapas foram produzidos em ambiente ESRI ArcMap (ESRI, 2014), utilizando o método de optimização de classificação de Jenks para determinar os intervalos mais apropriados para a apresentação espacial das variáveis.

O risco de morte por AVC e as variáveis socioeconómicas foram de seguida avaliados em termos de clustering espacial ao nível do município usando os índices Local e

Global de Moran I (Lloyd, 2010), em ambiente GeoDa (GeoDa, 2014), permitindo uma primeira análise explanatória da presença de autocorrelação espacial.

Foram calculados coeficientes de correlação linear de Pearson para identificar pares de variáveis explicativas com elevada correlação entre si. Vários critérios foram definidos para guiar esta fase de modelação. O primeiro consistiu em reter no modelo, tanto quanto possível, apenas as variáveis com coeficientes de correlação inferiores a 0,6 entre si. Para lá deste limiar, apenas uma das variáveis de um par altamente correlacionado (i.e. com $r \geq 0,6$) foi retida no modelo para posterior análise, reduzindo-se assim o número de variáveis independentes a incluir no subsequente processamento da modelação. Em segundo lugar, cada modelo deveria reter, se a correlação permitisse, pelo menos uma variável representando cada categoria. Com base nestes 2 critérios, foi filtrado um subconjunto de variáveis socioeconómicas a partir do conjunto original de todas as variáveis disponíveis, em ambos os modelos para homens e mulheres, antes de se dar início à análise da regressão.

2.2.3 Análises de regressão univariada de Poisson

As variáveis socioeconómicas foram investigadas para a sua associação potencial com o risco de morte por AVC. Estas variáveis incluem-se em várias categorias. A percentagem de pessoas que trabalham mais de 45 horas por semana no total da população com actividade laboral no município (calculada separadamente para homens e mulheres) e a percentagem de trabalhadores manuais mais trabalhadores agrícolas no total da população com actividade laboral (sem separação de sexos) em 1991 e 2001, foram incluídas como variáveis representativas do estatuto social. Incluiu-se uma variável relacionada com o emprego, nomeadamente a percentagem de população desempregada no município (calculada separadamente para homens e mulheres) em 1991 e 2001. Foi incluída uma variável respeitante ao nível de escolaridade atingido, especificamente a percentagem de população tendo atingido o ensino básico (calculado separadamente para homens e mulheres) em 1991 e 2001. Entre as variáveis representativas do rendimento, incluíram-se o índice municipal de poder de compra *per capita*, a percentagem municipal de poder de compra em todos os municípios em 1993 e 2000, e o valor médio municipal das rendas de casa (disponível apenas para 2000). O

número de infra-estruturas de saúde sediadas no município (disponível apenas para 2001), a distância em minutos entre a câmara municipal e o hospital mais próximo (baseado na rede rodoviária de 2011 e portanto utilizada somente no período 2002-2006) e a percentagem de pessoas que vivem em áreas predominantemente rurais (calculada separadamente para homens e mulheres) na área do município, foram incluídas como variáveis ambientais (com base nos dados de 2008 e assim testadas apenas no período 2002-2006). Deste conjunto de variáveis, apenas as que permaneceram após a avaliação de multicolinearidade, como discutido no passo anterior, foram avaliadas nesta fase.

As variáveis dependentes especificadas nos modelos de regressão univariada de Poisson foram o número de óbitos por AVC registados em cada município, em homens e mulheres com menos de 65 anos de idade em cada um dos 2 períodos da investigação. Especificou-se um ponderador (offset), correspondente à estimativa da população municipal segmentada por sexos e com menos de 65 anos de idade para os anos de 1994 e 2004, anos intermédios dos referidos períodos de 1992/96 e 2002/06, respectivamente. Quando se lida com os casos de doença (dados de contagem) e particularmente para doenças com um reduzido número de ocorrências, o modelo de Poisson pode revelar-se um modelo de regressão adequado. Outros estudos epidemiológicos, incidindo sobre áreas de pequena dimensão, aplicaram também o modelo de Poisson para descrever a distribuição de uma doença (Cheng et al., 2011). Assim, foram ajustados modelos de regressão univariada (simples) aos dados utilizando a função apropriada para modelos lineares generalizados (Dobson, 2002) com o software de estatística R (R, 2014b). Uma avaliação da presença de sobredispersão (R, 2014c) detectou a presença de sobredispersão significativa nos modelos de regressão univariada de Poisson, revelando-se estes últimos inapropriados para os dados em presença. Consequentemente, recorreu-se a modelos de regressão binomial negativa em todos os processos subsequentes envolvendo modelos de regressão não-espaciais multivariados e para as comparações finais destes com as outras abordagens de modelação (espaciais) testadas neste estudo.

2.2.4 Modelos multivariados de regressão binomial negativa

Tal como nas análises prévias de regressão univariada, o número de óbitos por AVC registados em cada município, para sexos e para cada um dos 2 períodos, constituiu a variável dependente especificada nos modelos de regressão multivariada binomial negativa, tendo-se especificado também como ponderador (offset) as estimativas de população residente no município em 1994 e 2004, homens e mulheres com idade inferior a 65 anos. Assim, a equação da regressão para o modelo binomial negativo é:

$$\ln(\lambda) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k$$

onde λ representa $E(Y)$, Y a variável dependente, β_k são parâmetros estimados (coeficientes de regressão) e X_k pertencem ao vector de K variáveis socioeconómicas em estudo (Weisent et al., 2012). Previamente à análise da regressão binomial negativa, aplicou-se o método de selecção de variáveis de McHenry (all possible variables), no software NCSS (NCSS, 2014) aos conjuntos de variáveis não colineares, tentando-se sinalizar o número adequado de variáveis a incluir nas análises de regressão multivariada e obter assim os modelos de regressão multivariada mais parcimoniosos. Paralelamente a este método, realizaram-se também experiências empíricas com diferentes modelos de variáveis, no sentido de se obter informação adicional sobre o grau de parcimónia dos modelos.

Com o objectivo de determinar todos os modelos possíveis que incluíssem combinações parcimoniosas das variáveis indicadas no passo anterior, utilizou-se uma função apropriada em ambiente R (R, 2014d) a partir do conjunto de variáveis independentes não colineares previamente definidas. De seguida desenvolveu-se em (R, 2014e) o processo de modelação baseado na regressão binomial negativa múltipla. Todas as combinações possíveis (e parcimoniosas) de variáveis foram testadas, sendo escolhidos os melhores modelos, primeiramente com base no melhor (mais baixo) valor do critério de informação de Akaike (AIC), (R, 2014f) e em segundo lugar, com base na significância estatística (pelo menos $p < 0.05$) dos coeficientes de regressão das variáveis

2.2.5 Regressão geograficamente ponderada de Poisson, global e local

Os modelos globais estimam um coeficiente para cada variável explicativa, o qual consiste numa média de todas as localizações, enquanto os modelos de regressão geograficamente ponderada estimam um conjunto de coeficientes para cada localização, ou seja, neste estudo, para cada município. A equação para o modelo GWR local é:

$$Y_i(\mu) = \beta_{0i}(\mu) + \beta_{1i}(\mu) X_{1i} + \beta_{2i}(\mu) X_{2i} + \dots + \beta_{ki}(\mu) X_{ki}$$

$\beta_{ki}(\mu)$ denotam coeficientes de regressão para a relação entre uma variável explanatória e a variável dependente em redor de uma localização μ e são assim específicos para essa localização enquanto X_{ki} representam as variáveis independentes incluídas no modelo (Fotheringham et al., 2002). O modelo GWR local permite assim ao investigador calcular um coeficiente de regressão específico para cada localização quando é avaliada a relação entre as variáveis dependente e independente. Por conseguinte, ao assumir uma relação causal, o investigador fica apto a avaliar como o impacto de um factor de risco específico varia com a localização geográfica. A utilização do modelo de distribuição de Poisson como distribuição de referência na GWR é presentemente a estratégia disponível mais adequada para analisar dados de contagem de casos de doença por áreas geográficas, particularmente quando estamos em presença de números pequenos (Weisent et al., 2012). Tal como nos modelos binomiais negativos, o número de óbitos por AVC registados a nível municipal em homens ou mulheres com idade inferior a 65 anos, para cada período considerado, constituiu a variável dependente especificada no modelo, e a variável ponderadora (offset) foram as estimativas de população residente no município de homens e mulheres com idade inferior a 65 anos, em 1994 e 2004. Ambos os modelos foram ajustados no software estatístico GWR, desenvolvido especificamente para os propósitos deste tipo de análise (GeoDa, 2014). Para o modelo de Poisson GWR local, escolheu-se um método de kernel adaptativo, mais apropriado para lidar com as diferenças de densidade e diversidade dos municípios ao longo do território de Portugal Continental. A capacidade de lidar com as configurações irregulares das áreas dos municípios é particularmente importante, pois as suas formas, dimensão e densidade variam substancialmente entre áreas metropolitanas

e áreas rurais. O método de kernel adaptativo possibilita a variação do tamanho da janela de análise, permitindo assim incorporar o mesmo número de municípios em cada estimativa local, sendo aplicada, em cada análise de regressão local, uma ponderação igual a zero a todos os municípios fora da janela de análise. Para cada modelo, recorreu-se a uma abordagem manual iterativa, baseada no AIC, para identificar o número de vizinhos (municípios) mais próximos a utilizar como largura de banda óptima do modelo. Para analisar a variabilidade espacial na associação entre o risco de AVC e as variáveis independentes, os coeficientes de regressão estimados a partir do modelo GWR local foram apresentados sob a forma de mapas coropletos, usando o esquema de classificação de Jenks. A avaliação do melhor ajuste dos modelos de regressão binomial negativa e GWR Poisson global e local foi feita através do AIC.

2.3 Resultados

2.3.1 Distribuição especial do risco de AVC e os factores socioeconómicos

As estimativas do risco de AVC padronizado por idades a nível municipal apresentaram um elevado grau de variabilidade, em ambos os sexos, com uma tendência geral para a diminuição entre os períodos de 1992/96 e 2002/06, e para valores consistentemente mais baixos nas mulheres. Assim, no período de 1992/96, e para os homens, os valores de risco oscilaram entre 0 (em 2 municípios) e 558 óbitos por 100000 habitantes, com uma mediana de 160. Em 2002/06, estes valores oscilavam entre 0 (em 7 municípios) e 353 óbitos por 100000 habitantes, com uma mediana de 93. Nas mulheres, o risco de AVC apresentou valores proporcionalmente mais baixos, variando entre 0 (em 7 municípios) e 255 por 100000 habitantes, com uma mediana de 91 no período de 1992/96. Todavia, em 2002/2006, o valor mínimo foi de 0 (em 29 municípios) e 283 óbitos por 100000 habitantes, com uma mediana de 52, apresentando assim um cenário diferente, com o aumento do valor superior do risco.

Quanto à distribuição espacial, o mapeamento do risco de AVC evidência a existência de disparidades geográficas ao longo da área de estudo, entre sexos e períodos temporais definidos. Em 1992/96, o mapa do risco de AVC nos homens mostra que os

Análise espaçotemporal da relação entre factores socioeconómicos e acidente vascular cerebral na população de Portugal Continental abaixo de 65 anos de idade

municípios com maior risco de AVC pareciam localizar-se mais longe do litoral nacional, com maior agregação nos municípios no nordeste português, enquanto o restante território de Portugal Continental evidenciava valores mais baixos mas tendencialmente com um aumento para leste das regiões litorais (Figura 55a). No que respeita à distribuição espacial do risco de AVC nas mulheres, o padrão de disparidades geográficas é ligeiramente distinto, com os valores mais elevados a registarem-se em municípios na região norte do país, com alguns valores elevados no sul, também ocorrendo maioritariamente no interior (Figura 55b).

No entanto, dez anos depois, o padrão parece quase sofrer uma inversão entre homens e mulheres. Embora o mapa correspondente ao período de 2002/06 nos homens mostre que os municípios com valores mais elevados de risco de AVC tendem a localizar-se longe do litoral em direcção ao Norte e Centro do país, surge já no sul (mais concretamente em municípios do Alentejo) um padrão diferente, que aparenta uma cintura desde o litoral até à fronteira sudeste alentejana (Figura 56a). No que respeita à distribuição espacial do risco para as mulheres, o padrão de disparidades geográficas apresenta-se muito mais suave, com valores mais elevados nos municípios localizados num eixo entre o centro norte e a metade sudeste de Portugal Continental (Figura 56b).

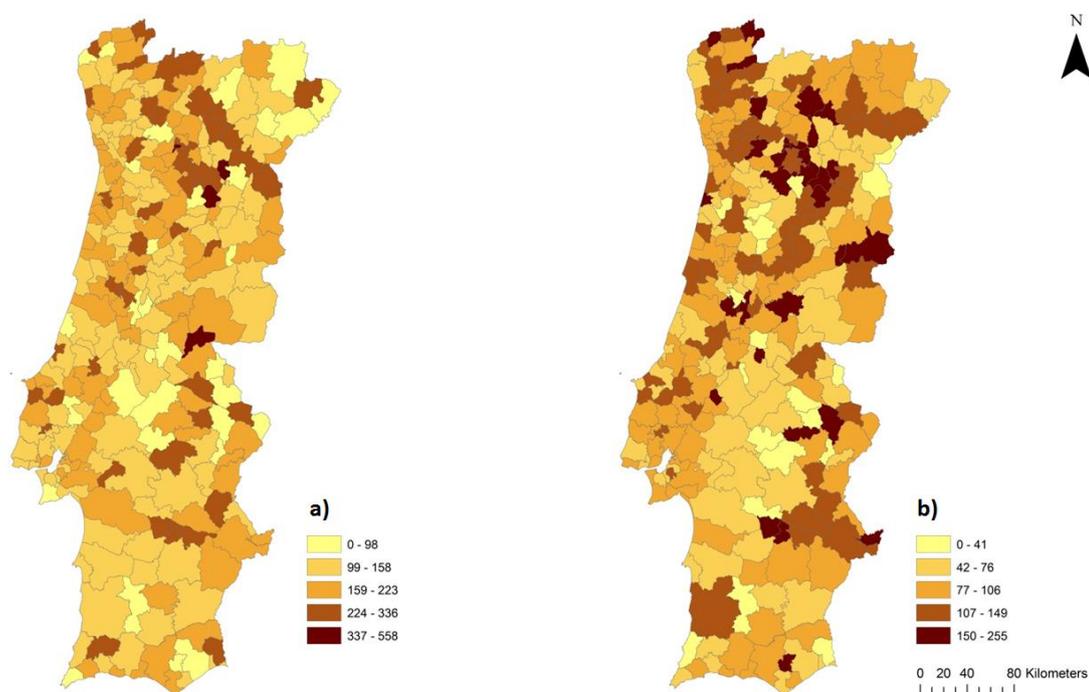


Figura 55 – Taxa de mortalidade padronizada por AVC em municípios de Portugal Continental, 1992-1996, por 100 000 habitantes em a) Homens b) Mulheres. Dados: INE.

Análise espacotemporal da relação entre factores socioeconómicos e acidente vascular cerebral na população de Portugal Continental abaixo de 65 anos de idade

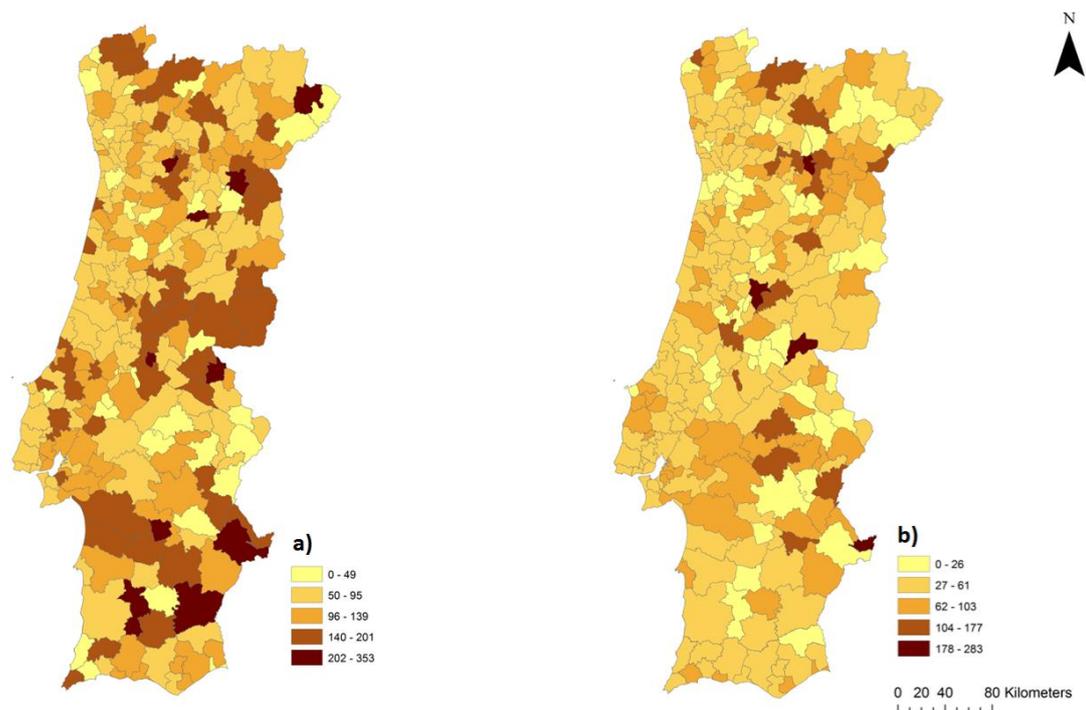


Figura 56 – Taxa de mortalidade padronizada por AVC em municípios de Portugal Continental, 2002-2006, por 100 000 habitantes em a) Homens b) Mulheres). Dados: INE.

As estatísticas descritivas dos factores socioeconómicos também evidenciavam importantes disparidades, entre municípios, entre os 2 períodos temporais e entre sexos (Tabela 3).

Tabela 3 – Estatísticas descritivas dos factores socioeconómicos investigados para potenciais associações com o risco de morte por AVC

Categoria	Variável	Md ¹ (1991)	Md ¹ (2001)	Dvp ² (1991)	Dvp ² (2001)	Med ³ (1991)	Med ³ (2001)	Min ⁴ (1991)	Min ⁴ (2001)	Max ⁵ (1991)	Max ⁵ (2001)
Estatuto Social	Homens que trabalhavam mais de 45 horas por semana (%)	34,8	24,7	9,6	5,5	34,0	24,5	14,7	11,6	67,9	46,4
	Mulheres que trabalhavam mais de 45 horas por semana (%)	21,9	11,5	8,1	3,6	20,7	10,9	5,8	5	56,0	29,1
	Trabalhadores manuais e agrícolas (%)	67,9	57,9	10,7	9,7	70,3	60,2	27,6	23,1	87,5	75,5
Emprego	Taxa de desemprego - Homens (%)	3,8	4,6	1,8	1,6	3,4	4,5	0,9	1,6	14,7	13,6
	Taxa de desemprego - Mulheres (%)	11,2	10,9	7,1	4,7	9,4	9,6	2,4	3,4	45	33,1
Nível de escolaridade	Homens que atingiram o ensino básico (%)	14	13	3,6	1,8	13,2	13	6,2	7,8	24,3	17
	Mulheres que atingiram o ensino básico (%)	11,7	10	3,2	1,7	11,4	10	5,2	5,3	21,1	14,4
Rendimento	Índice municipal de poder de compra per capita	61,5	66,6	33,9	28,8	52,0	58,6	18,9	33,7	291	305,2
	Poder de compra municipal (%)	0,4	0,4	1,4	1,2	0,1	0,1	0,01	0,01	20	16,4
	Valor médio mensal das rendas dos alojamentos (euros)		117,5		31,8		116		48,0		256
Ambiente	Infraestruturas de saúde municipais (nº)		8,6		8,4		6		1		93
	Distância ao hospital mais próximo (minutos)		26,3		18,4		25		1		89
	Homens a viver em áreas predominantemente rurais (%)		42,7		60		43		0		85,3
	Mulheres a viver em áreas predominantemente rurais (%)		31,6		25		29		0		84,8

¹Média; ²Desvio-padrão; ³Mediana; ⁴Mínimo; ⁵Máximo

Assim, na metade superior destes territórios administrativos em 1991, entre 34 a 68% dos homens integrados na população com actividade laboral trabalhavam mais de 45 horas por semana, enquanto em 2001 a situação melhorou, diminuindo para 25 a 46%. Nas mulheres, pode ser observada semelhante evolução, embora os valores absolutos

Análise espaçotemporal da relação entre factores socioeconómicos e acidente vascular cerebral na população de Portugal Continental abaixo de 65 anos de idade

mostrem uma melhor perspectiva global. Nos municípios da metade superior, em 1991 de 21 a 56% das mulheres com actividade laboral trabalhavam mais de 45 horas por semana, enquanto em 2001 esta situação apenas envolvia entre 10 a 29% das mulheres trabalhadoras. Em mais de metade dos municípios do território Continental em 1991, uma elevada percentagem de população (entre 70 e 88%) estava empregada em profissões manuais, diminuindo para 60 a 76% em 2001. Quanto ao nível de escolaridade atingido, em 1991 entre 6 e 13 % da população masculina tinha atingido a escolaridade básica na metade inferior dos municípios, com uma situação ligeiramente pior para as mulheres (entre 5 e 10%). Em 2001 esta situação não terá melhorado significativamente, registando-se entre 8 e 13% nos homens e 5 a 10% nas mulheres. A análise de clustering efectuada com base no Índice I de Moran evidenciou a presença de clustering espacial estatisticamente significativo ($p < 0.001$) na distribuição espacial dos factores socioeconómicos investigados (Figuras 57-60). Pode ser observada uma concentração de valores mais altos nas percentagens das populações de ambos os sexos que trabalhavam mais de 45 horas por semana, em 1991, nos municípios do interior norte e ao longo da fronteira nordeste (figuras 57a e 57b).

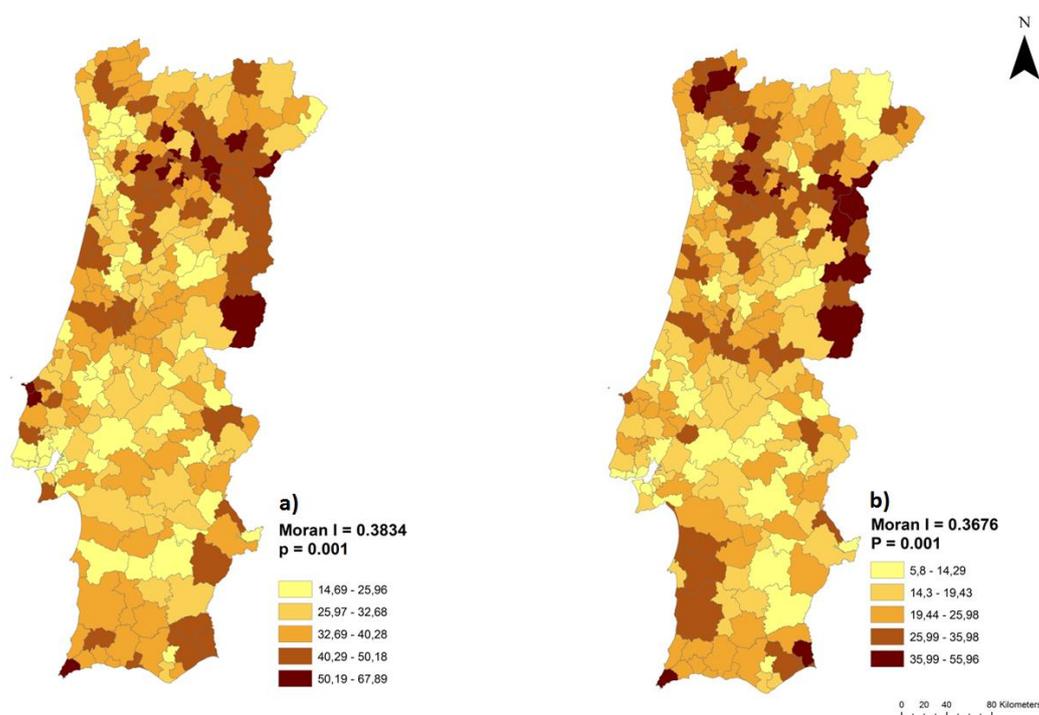


Figura 57 – Homens (a) e Mulheres (b) que trabalhavam mais de 45 horas por semana (%), em 1991, em municípios de Portugal Continental. Dados: INE.

Análise espaçotemporal da relação entre factores socioeconómicos e acidente vascular cerebral na população de Portugal Continental abaixo de 65 anos de idade

Em 2001, este padrão parece tornar-se mais suave, mas permaneceu uma concentração de maiores valores na região interior mais a norte (figuras 58a e 58b).

No que respeita à distribuição espacial da percentagem de população, por sexos, que atingiu o ensino básico, em 1991 verifica-se uma consistente concentração de valores elevados (para homens e mulheres) ao longo do litoral e em direcção ao sul do país, com alguns valores mais elevados a registarem-se em cidades do interior como Évora e Beja (figuras 59a e 59b). Este padrão manteve-se em 2001, mas com reforço da concentração nas regiões litorais (figuras 60a e 60b).

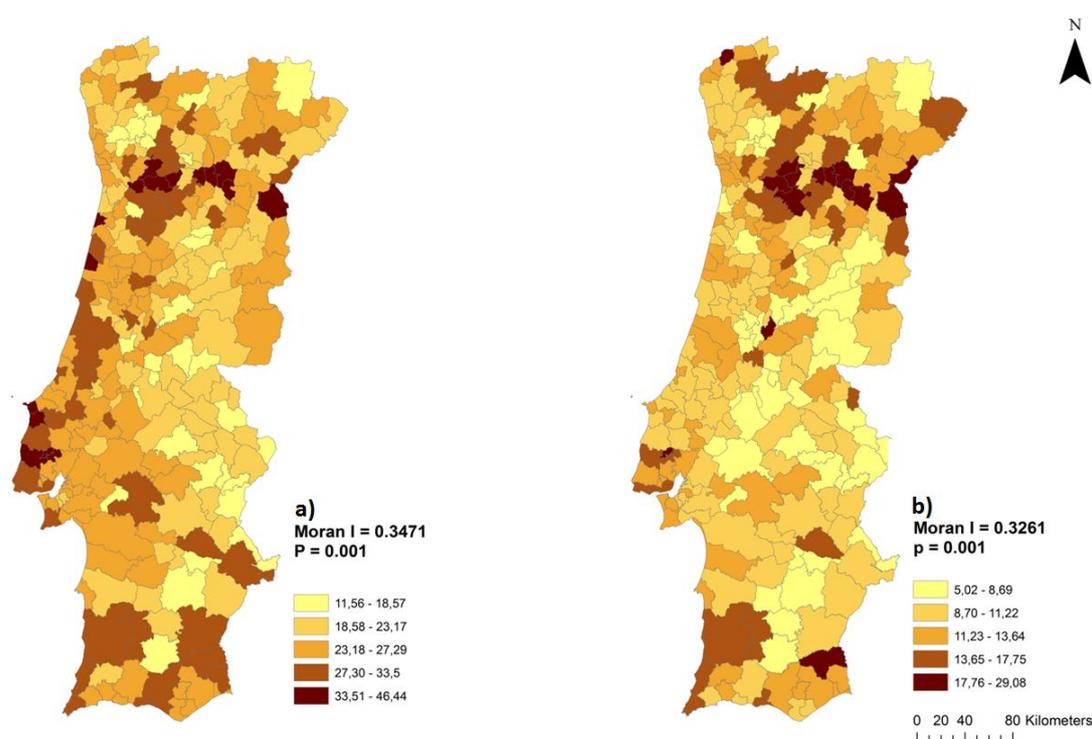


Figura 58 – Homens (a) e Mulheres (b) que trabalhavam mais de 45 horas por semana (%), em 2001, em municípios de Portugal Continental. Dados: INE.

A busca de multicolinearidade entre variáveis socioeconómicas determinou a presença de correlações positivas fortes, como expectável, entre o poder de compra municipal per capita e a percentagem municipal de poder de compra, tanto para 1992/96 como para 2002/06 e para ambos os sexos (Anexos 2 a 5).

Análise espaçotemporal da relação entre factores socioeconómicos e acidente vascular cerebral na população de Portugal Continental abaixo de 65 anos de idade

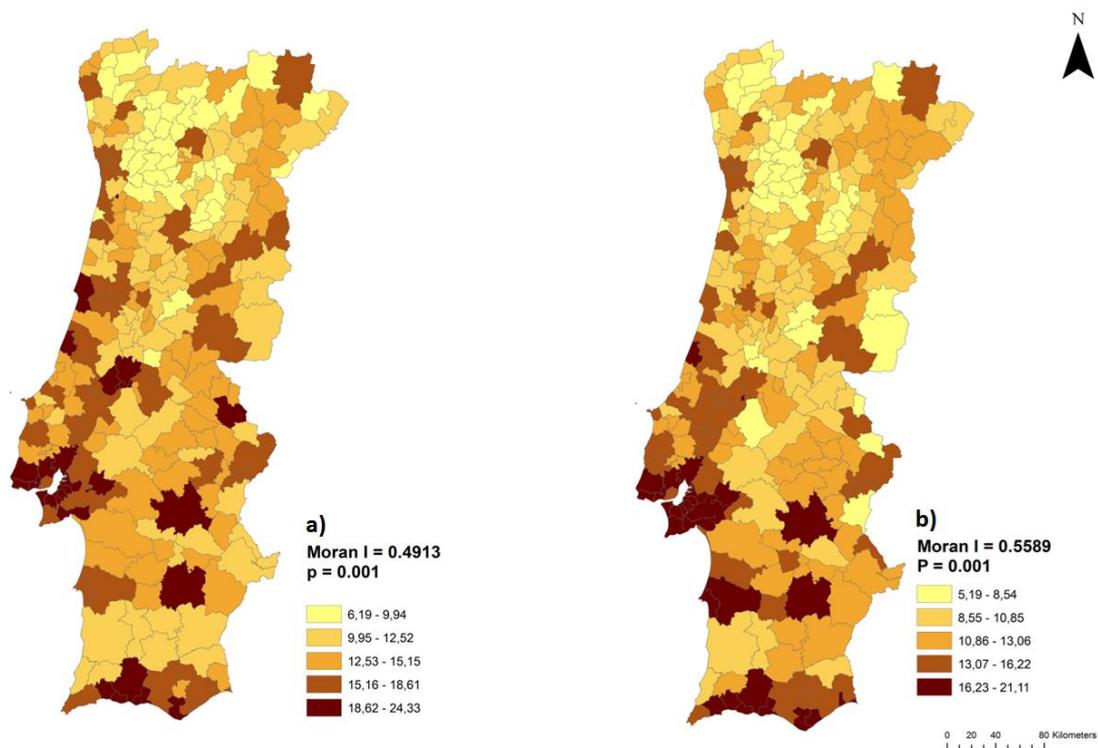


Figura 59 – Homens (a) e Mulheres (b) que atingiram o ensino básico (%) em 1991, em municípios de Portugal Continental. Dados: INE.

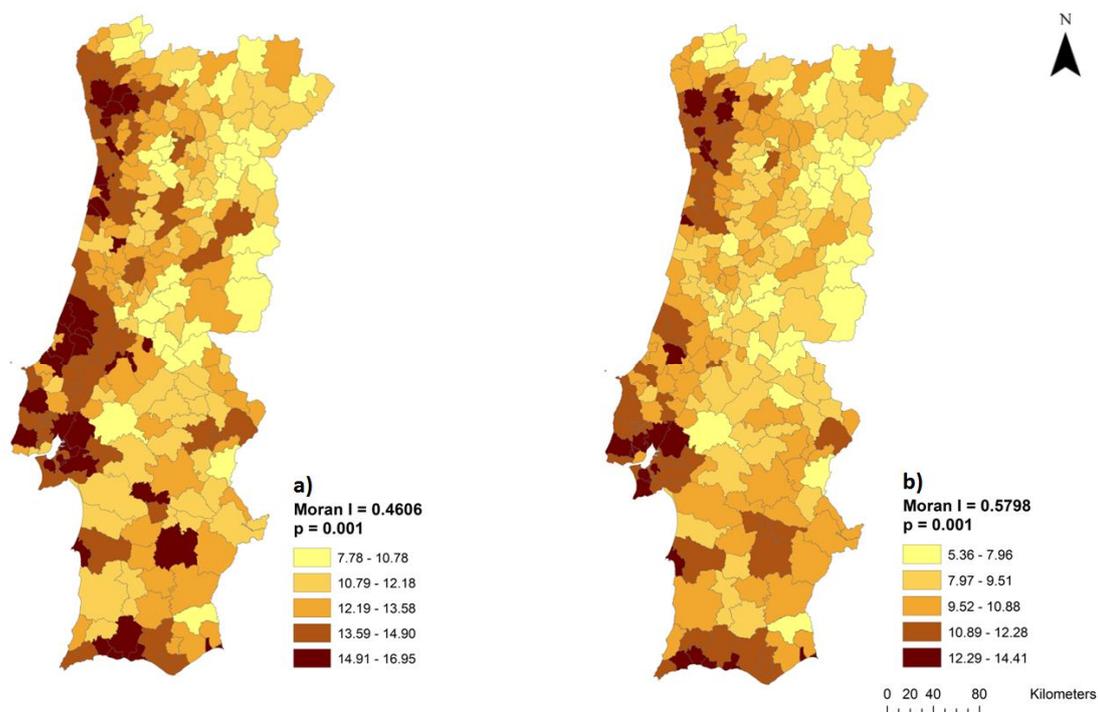


Figura 60 – Homens (a) e Mulheres (b) que atingiram o ensino básico (%) em 2001, em municípios de Portugal Continental. Dados: INE.

Em ambos os modelos do período 1992/96, ocorreu uma situação similar entre o poder de compra municipal *per capita* e a percentagem de população, masculina ou feminina, que atingiu o ensino básico (Anexos 4 e 5). Foram encontradas também correlações positivas fortes e significativas nos modelos para 2002/06 entre as 2 variáveis representativas do rendimento e o número de infra-estruturas municipais de saúde (Anexos 2 e 3). Inversamente, registaram-se correlações negativas fortes e significativas ($p < 0.001$) entre o poder de compra municipal e a percentagem de trabalhadores manuais e agrícolas no total de população activa, em 1992/96, para ambos os sexos. Verificou-se a mesma situação entre a percentagem de população que atingiu o ensino básico e a percentagem de trabalhadores manuais e agrícolas no total da população com actividade laboral (Anexos 4 e 5). Em ambos os modelos 2002/06, verificaram-se correlações negativas fortes e significativas entre o poder de compra municipal *per capita* e a percentagem de trabalhadores manuais e agrícolas no total da população com actividade laboral. Finalmente, notou-se uma correlação semelhante neste segundo período temporal, para homens e mulheres, entre a percentagem de população que vive em áreas predominantemente rurais e a percentagem de população que atingiu o ensino básico (Anexos 2 e 3).

2.3.2 Determinantes socioeconómicos da distribuição geográfica do risco de AVC

Depois de eliminadas as variáveis multicolineares no passo anterior, os factores socioeconómicos restantes foram investigados no que respeita à sua associação univariada com o risco de AVC, e a maioria revelou associações estatisticamente significativas ($p < 0.001$) com o risco (Anexo 6).

No entanto, foi detectada sobredispersão significativa nestas associações, situação que torna os modelos comuns de Poisson inapropriados para lidar com associações entre as distribuições destas variáveis. Uma vez que os modelos binomiais negativos se ajustam melhor a este tipo de dados que os modelos ordinários de Poisson (por acomodarem melhor a sobredispersão), a análise prosseguiu, a partir deste ponto, com recurso a modelos binomiais negativos.

Com base no método de selecção de McHenry e também nas experiências empíricas com vários modelos diferentes, foi estabelecido um conjunto de variáveis para os

modelos mais parcimoniosos. Especificamente, os modelos finais óptimos, para o período 2002/06, deveriam conter 5 variáveis explanatórias cada. Já para o período de 1992/96, os modelos deviam conter 2 ou 3 variáveis. Neste ponto, foram determinadas todas as combinações possíveis com o número de variáveis estabelecido, partindo do conjunto de variáveis não colineares previamente definido para cada modelo. Em seguida, todas as possíveis combinações para cada modelo foram testadas através do modelo de regressão binomial negativa, tendo sido escolhidos os melhores modelos masculinos e femininos para cada período temporal, com base no menor AIC. Finalmente, foram sequencialmente excluídas as variáveis cujos coeficientes não atingiam o critério de significância $p < 0.05$ e os modelos corridos iterativamente até que restassem apenas as variáveis que atingiam pelo menos um grau de significância aceitável ($p < 0.05$). A partir deste ponto, os 4 modelos finais foram usados para comparar o desempenho das 3 abordagens de modelação (binomial negativa, GWR global de Poisson e GWR local de Poisson) testadas neste estudo.

Os modelos de regressão binomial negativa para o período de 2002/06 evidenciaram uma tendência para o risco de AVC ser significativamente mais baixo em municípios com maiores percentagens de homens ou mulheres que atingiram o ensino básico, uma tendência já presente no período de 1992/96 para o sexo feminino (Tabela 4). Na tendência oposta, e para o período de 2002/06, o risco de AVC parece ser significativamente mais alto nos municípios com taxas de desemprego mais elevadas nas mulheres. Tal como para o período de 1992/96, o risco de AVC parece ser significativamente mais alto em municípios com maiores percentagens de homens que trabalham mais de 45 horas por semana (Tabela 4).

A avaliação dos resíduos padronizados de Pearson resultantes dos modelos binomiais negativos evidenciou a presença de autocorrelação espacial positiva nos resíduos de todos os modelos (Moran I: 0.13, $p=0.002$ no modelo masculino 2002-2006; I: 0.07, $p=0.03$ no modelo das mulheres 2002-2006; I: 0.14, $p=0.001$ no modelo das mulheres 1992-1996; I: 0.09, $p=0.01$ no modelo dos homens 1992-1996). Isto implica que, mesmo tendo o modelo binomial negativo apresentado melhor desempenho do que o modelo de Poisson comum, os seus resíduos contêm ainda assim autocorrelação espacial, não tendo este modelo de regressão sido capaz de eliminar totalmente a autocorrelação espacial nos modelos de dados testados. Consequentemente, a presença

Análise espaçotemporal da relação entre factores socioeconómicos e acidente vascular cerebral na população de Portugal Continental abaixo de 65 anos de idade

de autocorrelação espacial significativa nos resíduos de todos os modelos de dados requer o uso de um modelo espacial mais apropriado.

2.3.3 Avaliação comparativa das abordagens de modelação e análise espacial dos coeficientes locais

Usando o valor do critério de informação de Akaike (AIC) para comparar e avaliar o desempenho dos modelos, aquele que apresentar o menor valor de AIC será o modelo com melhor ajuste. Este critério foi escolhido para comparar as 3 abordagens de modelação, testadas nos 4 modelos finais. Para todos os modelos de dados, o modelo de regressão GWR local de Poisson evidenciou o melhor ajuste, seguido pelo modelo GWR Poisson global e por fim, pelo modelo binomial negativo (Tabela 4).

Tabela 4 – Avaliação comparativa de desempenho dos modelos testados

	Tipo de modelo e estimativas dos coeficientes (p-values)						
	Modelo Binomial Negativo		Modelo Global Poisson GWR ¹		Modelo Local Poisson GWR ¹		
Modelo 2002-2006 (Homens)	Coeficiente (p-value)	AIC ²	Coeficiente (p-value)	AIC ²	Mínimo	Máximo	AIC ²
Termo de intercepção	-5.7782 (0.0001)	1470	-5.9075 (0.0001)	520	-7.4228	-3.7536	431
Homens que atingiram o ensino básico	-0.0853 (0.0001)		-0.0783 (0.0001)		-0.2616	0.0376	
Modelo 2002-2006 (Mulheres)							
Termo de intercepção	-7.2131 (0.0001)	1215	-7.2182 (0.0001)	467	-7.2760	-6.5239	354
Mulheres que atingiram o ensino básico	-0.0570 (0.0015)		-0.0578 (0.0002)		-0.1229	-0.0384	
Taxa de desemprego (Mulheres)	0.0262 (0.0005)		0.0264 (0.0001)		0.0164	0.0274	
Modelo 1992-1996 (Homens)							
Termo de intercepção	-6.7290 (0.0001)	1657	-6.6724 (0.0001)	616	-7.4739	-5.6543	534
Homens que trabalham mais de 45 horas por semana	0.0089 (0.0002)		0.0061 (0.0001)		-0.0327	0.0354	
Modelo 1992-1996 (Mulheres)							
Termo de intercepção	-6.6254 (0.0001)	1378	-6.6884 (0.0001)	397	-7.4714	-5.3142	322
Mulheres que atingiram o ensino básico	-0.0307 (0.0001)		-0.0269 (0.0001)		-0.1324	0.0290	

¹Regressão Geograficamente Ponderada; ²Critério de Informação de Akaike

É relevante assinalar que todos os 4 modelos de Poisson GWR apresentaram evidência de não estacionaridade nos coeficientes de regressão, com a única excepção da taxa de desemprego nas mulheres para o período de 2002/06. Esta situação é evidenciada pela

Análise espaçotemporal da relação entre factores socioeconómicos e acidente vascular cerebral na população de Portugal Continental abaixo de 65 anos de idade

comparação entre as amplitudes interquartil dos coeficientes locais de regressão, cujos valores se apresentam superiores ao dobro do erro padrão dos coeficientes de regressão do modelo GWR de Poisson global, excepto para o período de 2002/06 na taxa de desemprego nas mulheres (Tabela 5).

Isto indica que os coeficientes de regressão para a maior parte das variáveis incluídas nos modelos GWR locais não se revelaram constantes, variando ao longo dos municípios de Portugal Continental. Isto significa que a força das associações entre o risco de AVC e cada uma das variáveis explanatórias finais varia em função da localização espacial, excepto para o período de 2002/06 na taxa de desemprego nas mulheres.

Tabela 5 – Avaliação da estacionaridade dos coeficientes nos modelos GWR de Poisson locais

Modelo	Variável	Modelo Global Poisson GWR ¹		Modelo Poisson GWR ¹ Local	Não-estacionaridade do coeficiente de regressão
		Erro Padrão (SE)	2 X Erro Padrão (2SE)	Distância Interquartilica Range (IQR)	IQR > 2SE
Mulheres com menos de 65 anos de idade (2002-2006)	Mulheres que atingiram o ensino básico	0,015	0,031	0,048	Sim
	Taxa de desemprego feminina	0,007	0,014	0,003	Não
Mulheres com menos de 65 anos de idade (1992-1996)	Mulheres que atingiram o ensino básico	0,004	0,009	0,058	Sim
Homens com menos de 65 anos de idade (2002-2006)	Homens que atingiram o ensino básico	0,01	0,021	0,073	Sim
Homens com menos de 65 anos de idade (1992-1996)	Homens que trabalham mais de 45 horas por semana	0,002	0,003	0,013	Sim

¹Regressão Geograficamente Ponderada

Assumindo uma relação causal, esta situação implica que os efeitos dos determinantes não são estáticos ao longo da área em estudo mas, pelo contrário, dependem em larga medida da localização espacial.

As figuras 61 a 64 apresentam os padrões espaciais dos coeficientes municipais da regressão GWR local para as variáveis explanatórias que evidenciaram não-estacionaridade espacial nos modelos finais, e a evidência estatística para a sua não-estacionaridade pode ser observada na tabela 5. De modo a analisar a deriva espacial dos coeficientes da regressão da GWR local, e como cada modelo só tinha uma variável

Análise espaçotemporal da relação entre factores socioeconómicos e acidente vascular cerebral na população de Portugal Continental abaixo de 65 anos de idade

com alta significância final e não estacionária, algumas das variáveis que foram descartadas no refinamento final nos modelos de regressão binomial negativa (mas apresentavam pelo menos um nível de confiança de 90%) foram usadas como variáveis auxiliares para providenciar contexto espacial e temporal, tendo sido também usada com esta finalidade a variável mortalidade por AVC padronizada pela idade.

Assim, no modelo GWR de Poisson local, nos homens e para o período de 1992/96, o risco de AVC tendeu a ser maior nos municípios com percentagens relativamente mais elevadas de pessoas a trabalhar mais de 45 horas por semana (33,4%). Esses municípios estavam maioritariamente localizados na área mais remota da região nordeste transmontana, ao longo de uma faixa que ocupava grande parte do Alentejo e um conjunto de municípios do distrito de Leiria e Coimbra (Figura 61).

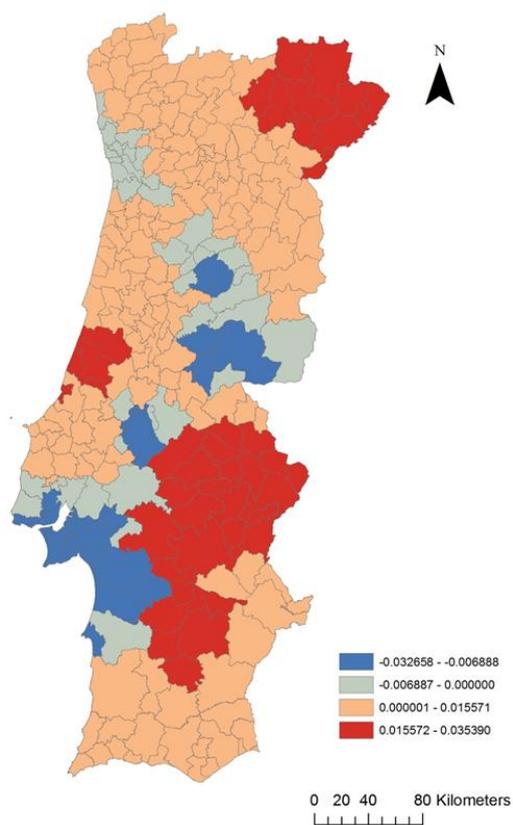


Figura 61 – Distribuição geográfica dos coeficientes do modelo GWR de Poisson local - Homens que trabalhavam mais de 45 horas por semana (%) no período 1992-1996, em municípios de Portugal Continental.

Estas áreas tenderam também a registar valores mais elevados de mortalidade padronizada pela idade (153,73 óbitos por 100000 indivíduos) e valores mais baixos nas

taxas de desemprego (4,46%). Todavia, áreas com percentagens comparativamente baixas de pessoas a trabalhar mais de 45 horas por semana (28,54%) tiveram a mais forte associação negativa com o risco de AVC. Estas áreas apresentaram, em média, valores mais reduzidos de mortalidade padronizada pela idade (150,06 óbitos por 100000 indivíduos), valores ligeiramente mais elevados nas taxas de desemprego (4,76%) e a mais larga estendia-se entre a costa a norte de Lisboa e o meio da costa alentejana. Assumindo uma relação causal, o risco pareceu diminuir entre algumas das mais remotas regiões de Portugal e o litoral, onde as mais importantes cidades estão localizadas, embora algumas áreas intermédias de baixo risco constituam um desafio.

No modelo masculino relativo ao período 2002/06, o risco de AVC evoluiu para associações significativas com os homens que atingiram o ensino básico, sendo estas associações negativas na maior parte do território, com excepção de uma faixa que se estende entre os municípios de Rio Maior e Serpa e ainda no município de Niza (isolado). É interessante registar que áreas com associações mais negativas podem ser encontradas em municípios ao longo da fronteira norte e centro com Espanha, a par de um cluster praticamente no centro de Portugal, centrado no município de Góis (Figura 62).

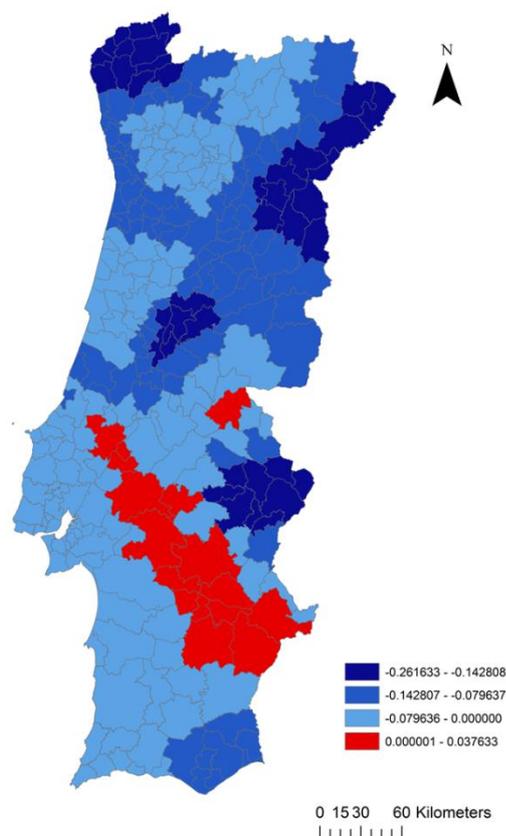


Figura 62 – Distribuição geográfica dos coeficientes do modelo GWR de Poisson local - Homens que atingiram o ensino básico (%), no período 2002-2006, em municípios de Portugal Continental.

Estes municípios estavam, em média, mais distantes de hospitais (44 minutos versus 27 em municípios com associações positivas) e registavam maiores valores de mortalidade padronizada pela idade do que aqueles com associações positivas (111,93 óbitos por 100000 indivíduos versus 105,63, respectivamente).

No que respeita ao modelo feminino para 1992/96, as associações entre as mulheres que atingiram o ensino básico e o risco de AVC são negativas na maior parte do território, mas passam a positivas numa área genericamente centrada na Área Metropolitana do Porto, a segunda região mais densamente povoada do país, após a região de Lisboa, e ainda numa área próxima, mas muito mais pequena, centrada no município de Santa Marta de Penaguião.

Pelo contrário, o município de Lisboa e a sua área metropolitana estão entre os que registam a mais forte associação negativa, um claro padrão de divergência entre as 2 grandes áreas urbanas do país (Figura 63).

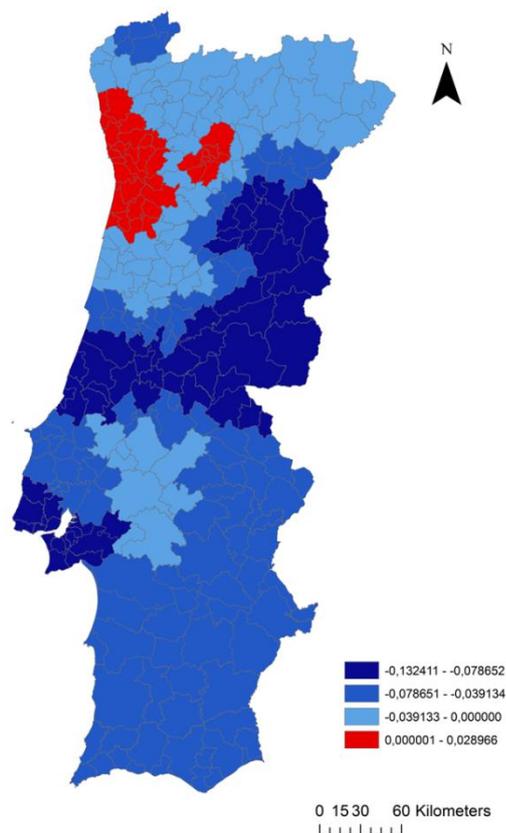


Figura 63 – Distribuição geográfica dos coeficientes do modelo GWR de Poisson local – Mulheres que atingiram o ensino básico (%), no período 1992-1996, em municípios de Portugal Continental.

Os municípios com associações positivas tiveram, em média, menos mulheres a atingir a escolaridade básica (10,54%) do que as áreas com fortes associações negativas (12,53%) e maior mortalidade padronizada pela idade. No entanto, municípios com forte associação negativa apresentavam uma média relativamente mais elevada de mulheres que trabalham mais de 45 horas por semana (21,45% versus 20,88) e uma maior taxa de desemprego (9,17 versus 7,71), quando comparados com municípios com associação positiva.

Finalmente, no modelo feminino para 2002/06, registaram-se associações negativas entre mulheres que atingiram o ensino básico e o risco de AVC em todos os municípios de Portugal Continental, com os valores mais negativos a registarem-se na região nordeste interior, e com tendência para aumento em direcção ao litoral e sul do país (Figura 64).

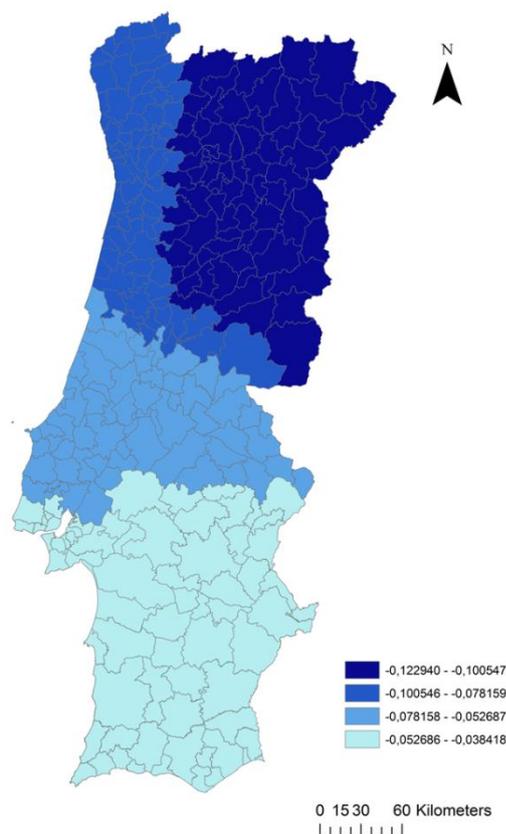


Figura 64 – Distribuição geográfica dos coeficientes do modelo GWR de Poisson local – Mulheres que atingiram o ensino básico (%), no período 2002-2006, em municípios de Portugal Continental.

É também interessante assinalar que os indicadores auxiliares parecem piorar na ordem inversa, dado que, em média, os municípios com as associações negativas mais elevadas se encontravam mais distantes dos hospitais (33 minutos versus 28), tinham uma maior percentagem de mulheres a trabalhar mais de 45 horas por semana (13,47% versus 10,96%), uma maior taxa de desemprego (12,85% versus 12,48%) e também uma mais elevada taxa de mortalidade padronizada pela idade, com 62,61 versus 59,05 óbitos por 100000 indivíduos.

2.4 Discussão

Que seja do nosso conhecimento, nenhum estudo anterior investigou a variação geográfica da associação entre o risco de AVC e factores socioeconómicos na população abaixo dos 65 anos com recurso a modelos GWR locais de Poisson, em combinação com uma análise multitemporal referida a 2 períodos distintos, tal como foi

feito no presente estudo. Por conseguinte, esta investigação pretende contribuir para o avanço do conhecimento nesta área, tanto no contexto de Portugal como a nível internacional.

Os modelos GWR locais existem já há algum tempo e podem ser uma ferramenta válida para a exploração de relações complexas entre efeitos na saúde e factores socioeconómicos potenciadores do risco de doença, com alguns autores a pugnar claramente pela necessidade de serem realizadas investigações sobre as variações no espaço geográfico dos coeficientes de regressão de modo a assegurar que os programas de controlo de doenças são postos em prática de forma eficaz, independentemente do tipo de doença e da área geográfica sob análise (Weisent et al., 2012). Isto é particularmente verdadeiro no caso de uma doença multifactorial como o AVC, cujas associações com factores socioeconómicos têm vindo a ser mencionadas na literatura desde há já algum tempo (Cox et al., 2006); (Addo et al., 2012), não obstante a exacta medida destas associações permaneça até ao presente algo difusa.

Observa-se neste estudo uma marcada diferença entre os coeficientes dos modelos multivariados globais e os dos modelos locais. De facto, os modelos globais produzem um valor único (um coeficiente apenas) que quantifica a força da associação entre os factores socioeconómicos e o risco, ou por outras palavras, assumem implicitamente que a medida da associação é a mesma em todos os municípios.

Por exemplo, as abordagens globais sinalizaram que os municípios com elevadas percentagens de pessoas tendo atingido o ensino básico tenderam para um risco de AVC consistentemente baixo, como se observa nos modelos para ambos os sexos, no período de 2002/06, e no modelo para mulheres no período temporal de 1992/96, e os municípios com elevadas percentagens de homens a trabalhar mais de 45 horas por semana, em 1992/96 tenderam de forma consistente para um maior risco de AVC. Reciprocamente, os modelos GWR locais de Poisson desdobraram-se numa diversidade de coeficientes para a associação entre risco de AVC e os seus potenciais determinantes ao longo do território de Portugal Continental. Isto sugere que os modelos globais poderão ser menos exactos e fiáveis, dado que os coeficientes de regressão dos determinantes variaram entre o negativo e o positivo ao longo de todo o território. Assim, os modelos locais parecem apresentar uma melhor abordagem na avaliação da natureza das relações entre risco de AVC e determinantes, ao longo da área de estudo.

Estes resultados transmitem a ideia de que a força de associação entre AVC e os seus determinantes varia em função da localização, algo que necessita ser levado em conta nos programas de Saúde Pública delineados com o objectivo de prevenir ou mitigar esta doença, dado que um factor socioeconómico pode ser mais relevante enquanto determinante desta doença em alguns municípios (ou grupo de municípios), e menos em outros. Neste estudo, os resultados da GWR local de Poisson revelaram, de forma clara, a existência de padrões espaciais na variação geográfica dos coeficientes, como aliás já terá sido observado noutros estudos que seguiram a mesma metodologia, embora para outras doenças (St-Hilaire et al., 2010); (Cheng et al., 2011); (Helbich et al., 2012); (Weisent et al., 2012) e inclusive num estudo muito recente e focado no AVC, especificamente na influência das disparidades geográficas de vizinhança na doença coronária e AVC no Estado norte-americano do Tennessee (Odoi e Busingye, 2014). Por exemplo, nos modelos para cada um dos sexos no período de 2002/06, e também para mulheres no período de 1992/96, que concluíram o ensino básico, parece existir um maior impacto na diminuição do risco de AVC nos municípios mais remotos do país. Isto é especialmente visível no modelo para mulheres no período temporal de 2002/06. Coloca-se assim a hipótese de atingir este nível de ensino ser mais relevante para a diminuição do risco de AVC nestas áreas mais remotas, com uma população mais escassa e tipicamente mais envelhecida, enquanto este determinante inverte a sua influência em direcção ao litoral, onde há mais população e com maiores níveis de escolaridade. Assumindo que os factores socioeconómicos auxiliares, usados para contextualização, contribuem para agravar o risco de AVC, a distribuição espacial destes também parece reforçar esta hipótese, dado que as áreas mais remotas estão, em geral, mais afastadas das infra-estruturas de saúde e têm uma percentagem relativamente maior de mulheres a trabalhar mais de 45 horas por semana, bem como maiores taxas de desemprego e de mortalidade padronizada por idade.

Paradoxalmente, o uso de uma metodologia incluindo uma análise GWR de Poisson local também contribui para confirmar a tendência geral das variáveis finais para baixarem ou aumentarem o risco de morte por AVC, transmitida pelos modelos globais. Assim, mesmo no modelo dos homens para o período de 1992/96 (o que registou maior heterogeneidade no coeficiente), os coeficientes para a percentagem de homens a trabalhar mais de 45 horas por semana foram positivos na maior parte dos municípios

de Portugal Continental (77%). Isto reforça os resultados alcançados noutros estudos focados na influência deste factor no risco de AVC nos homens em idade laboral (Hayashi et al., 1996). Do mesmo modo, nos 3 modelos restantes, a percentagem de pessoas que atingiram o ensino básico apresenta um coeficiente negativo na maioria dos municípios, mais concretamente 94% (modelo de 2002/06 para os homens), 87% (modelo de 1992-1996 para as mulheres) e mesmo 100% (modelo 2002/06 para as mulheres). Isto reforça também a ideia geral de que níveis de escolaridade mais elevados atingidos conduzem a melhores resultados de saúde (Avendaño et al., 2004); (Henriques et al., 2009); (Wu et al., 2013).

No que concerne à variação espaçotemporal dos modelos, é relevante notar que, embora se verifiquem mudanças nos modelos masculinos, não apenas nos factores significativos em jogo mas também nos seus padrões de distribuição espacial, alguns factores permanecem consistentemente significativos. É o caso de ambos os modelos nas mulheres. Assim, assumindo uma relação causal, a percentagem de mulheres que atingiram o ensino básico permaneceu um factor de protecção entre a última década do século XX e a primeira do século XXI na maior parte dos municípios do território continental português. Acrescente-se que alguns municípios do sul do país exibem coeficientes positivos e elevados nos modelos dos homens nos 2 períodos temporais, sugerindo que estes territórios devam ser investigados.

Alguns padrões mantêm-se, todavia, difíceis de interpretar, como seja o cluster isolado de municípios centrados na área do Porto no modelo das mulheres em 1992/06, e o grupo de municípios distribuídos num padrão com orientação noroeste-sudeste para os homens em 1992/96, todos evidenciando coeficientes positivos para associação entre risco de AVC e pessoas que atingiram o ensino básico. Levanta-se aqui a hipótese de poderem existir outros factores determinantes em jogo nestas áreas, para além dos incluídos no presente estudo, que podem eventualmente possuir força suficiente para confundir ou ofuscar a tendência geral de menor risco de AVC quando aumenta a percentagem de população que atinge o ensino básico.

Vários rumos de acção permanecem em aberto para exploração sobre o tema do presente estudo, alguns decorrentes de limitações próprias, outros mais relacionados com o contexto português. Por exemplo, a ausência de uma medição directa e sistemática da pobreza e do rendimento em Portugal (e da sua distribuição espacial a

uma escala de análise como o município) leva a que os investigadores tenham de socorrer-se de variáveis alternativas, com foi feito neste estudo. Coloca-se assim a hipótese de que o uso de variáveis mais explícitas na medição dos níveis de pobreza e de rendimento da população portuguesa possa fortalecer as conclusões retiradas dos modelos aqui utilizados e destas variáveis poderem apresentar níveis mais elevados de significância. Algumas variáveis poderiam beneficiar com um melhor alinhamento temporal em relação aos períodos de referência da mortalidade. Ou seja, os modelos da distância ao hospital mais próximo, em 2002/06, foram calculados a partir da rede de estradas existente em 2011. Todavia, a escolha deste ano foi a única opção disponível, não tendo sido possível obter uma rede rodoviária mais próxima dos períodos temporais em análise. Outra limitação prende-se com a avaliação da mortalidade por AVC. O facto do número absoluto de óbitos por AVC ter apresentado tendência para a diminuição no território continental português e na maioria dos municípios entre os 2 períodos temporais escolhidos, levanta a necessidade de calibração futura dos modelos, de forma a incorporar o crescente efeito dos números pequenos, ou mesmo do número zero, por unidade territorial, particularmente na população com menos de 65 anos de idade, população-alvo neste estudo.

2.5 Conclusão

Tal como previamente mencionado, a relação entre AVC e factores socioeconómicos é muito complexa, não sendo definitivamente de fácil avaliação. Neste sentido, este estudo fornece várias contribuições para uma melhor compreensão da natureza desta relação. Estas contribuições centram-se no segmento crítico da população abaixo dos 65 anos de idade, raramente analisado separadamente. Fornece também uma visão mais diversificada, ao segmentar a análise em várias e diferentes dimensões, nomeadamente as dimensões espacial e temporal, bem como a segmentação por sexos. Ao ir para lá da tradicional estrutura usada neste tipo de estudos, geralmente baseada em modelos de regressão global, tentou-se não só contribuir para um melhor entendimento da real natureza da relação entre AVC e determinantes socioeconómicos, mas também para

Análise espaçotemporal da relação entre factores socioeconómicos e acidente vascular cerebral na população de Portugal Continental abaixo de 65 anos de idade

melhorar o enfoque geográfico de intervenção das iniciativas de Saúde Pública, particularmente as que se desenvolvam com intuítos preventivos, mas não só.

Tendo estes objectivos em mente, será muito interessante não só melhorar esta análise com a implementação de algumas das melhorias previamente sugeridas na discussão, mas também alargar esta análise ao período de 2012/16. Como este estudo se centrou em 2 períodos temporais anteriores à crise económica iniciada em 2008, acredita-se que um retrato espaçotemporal que inclua o período mais crítico desta crise poderá incrementar bastante os benefícios desta análise.

3. Condições precárias de habitação e mortalidade por acidente vascular cerebral ao nível da freguesia em Portugal Continental, na população com mais de 65 anos de idade

3.1 Introdução

Muitos estudos indicam que as condições de habitação podem ter impactos significativos na saúde humana, tanto do ponto de vista da saúde física como mental (Marsh et al., 1999). Num estudo focado na evidência de excesso de ocorrências de enfarte agudo miocárdio durante a época mais fria em Portugal, fundamentado em questionários sobre as condições das suas residências a pacientes hospitalares com síndromas coronários, concluiu-se pela existência de um peso significativo das habitações deficientemente aquecidas durante o inverno (Vasconcelos et al., 2011). Noutro estudo sobre a incidência da tuberculose pulmonar em Portugal, a superlotação das habitações é apontada como um dos principais factores de risco de ocorrência, inclusive identificando regiões do país onde este determinante socioeconómico é relevante (Couceiro et al., 2011). Do ponto de vista da saúde mental (Marsh et al., 1999) referem a residência em apartamentos como tendo essencialmente impacto negativo em mulheres, contribuindo para o isolamento social e o desenvolvimento de perturbações psiquiátricas.

Não obstante as melhorias ocorridas na qualidade dos espaços interiores nas últimas décadas nos países da União Europeia, permanece assim um conjunto bem delimitado de riscos para a saúde, decorrentes das condições de habitabilidade das residências, entre os quais a temperatura inadequada, poluição atmosférica, humidade, desenvolvimento de bolores, ausência de equipamentos sanitários e a superlotação das habitações.

É também reconhecido que os problemas com a qualidade das habitações afectam de forma desigual os segmentos da população mais vulneráveis em termos da sua situação socioeconómica ou da faixa etária a que pertencem (WHO, 2013); (Eurostat, 2014b). Assim, na priorização de intervenções por parte dos decisores em Saúde Pública, é importante ter-se em conta que determinados tipos de desigualdades nas condições de

Condições precárias de habitação e mortalidade por acidente vascular cerebral ao nível da freguesia em Portugal Continental, na população com mais de 65 anos de idade

habitação parecem afectar mais as crianças e o seu desenvolvimento, enquanto outros tipos afectam com mais intensidade as faixas etárias mais idosas (Marsh et al., 1999). A discussão relativa ao impacto nas faixas etárias mais idosas será retomada mais à frente neste texto.

As tabelas 6 e 7 identificam um conjunto de indicadores de má qualidade na habitação e as suas consequências para a saúde física e mental, respectivamente.

Tabela 6 – Indicadores de má qualidade na habitação e consequências para a saúde física. Adaptado de Marsh et al. (1999)

Sobrelotação das habitações	- Aumento do risco de infecções ou de problemas respiratórios - Reduzida estatura dos indivíduos
Humidade e bolores	- Problemas respiratórios (por exemplo respiração arquejante) - Asma, rinite e alveolite - Equizema
Poluentes no espaço interior da habitação e infestações	- Asma
Frio	- Diminuição da resistência às infecções respiratórias - Hipotermia - Broncoespasmo - Doença cardíaca isquémica, enfarte de miocárdio e AVC
Ausência permanente de habitação (Sem Abrigo)	- Problemas resultantes da falta de protecção contra os elementos
Ausência temporária de habitação (Realojamento temporário)	- Problemas resultantes da sobrelotação, ruído, alimentação e higiene inadequadas

Tabela 7 – Indicadores de má qualidade na habitação e consequências para a saúde mental. Adaptado de Marsh et al. (1999)

Pessoas a viver nas habitações de pior qualidade relativamente a outras unidades de habitação próximas	- Redução do bem-estar mental dos residentes
Residir em habitações localizadas em áreas menos desejadas	- Menor qualidade de vida emocional por comparação com residentes em áreas mais aprazíveis
Humidade	- Depressão nas mulheres
Sobrelotação nas habitações	- Problemas emocionais e atrasos no desenvolvimento, maiores níveis de abandono escolar precoce e dificuldade de inserção social - Tensões sociais, irritabilidade, enfraquecimento das relações sociais
Residir em apartamentos	Maior reflexo nas mulheres: - aumento das consultas de clinica geral por problemas emocionais - isolamento social e perturbações psiquiátricas

Mais recentemente, a importância e prevalência de muitos destes indicadores foi revista para os países da União Europeia do ponto de vista estatístico, e constata-se que muitos dos problemas indicados em Marsh et al. (1999) continuam a estar presentes e a ser relevantes nestes países. Em Eurostat (2014b), indica-se que em 2012, um em cada dez europeus tinha dificuldade em manter a sua habitação adequadamente aquecida devido a dificuldades económicas. Outros factores considerados relevantes neste relatório, do ponto de vista das condições de habitação, foram a superlotação das habitações, a

Condições precárias de habitação e mortalidade por acidente vascular cerebral ao nível da freguesia em Portugal Continental, na população com mais de 65 anos de idade

habitação em apartamentos, o facto de se ser ou não proprietário da habitação (a propriedade da habitação constitui, segundo esta análise, um factor de vantagem socioeconómica, na medida em que o direito de propriedade parece aumentar a capacidade monetária do agregado familiar) e os problemas estruturais da unidade habitacional (entre os quais se inclui a ausência ou deficiência de capacidade de aquecimento).

Para além da análise das tendências globais na União Europeia, existem também desigualdades marcadas nestes indicadores entre países. Assim, a percentagem de população com dificuldade em manter a sua habitação adequadamente aquecida devido a dificuldades económicas atingia, em 2013, valores tão elevados como 44,9% (Bulgária), 30,5% (Chipre), 29,5% (Grécia), 29,2% (Lituânia). Embora a situação portuguesa atingisse 36,3% em 2004, valor muito superior quando comparado com 14 países com dados para esse ano, foi melhorando gradualmente até 2013, mantendo-se no entanto ainda o 4º país com maior percentagem de população com dificuldade para manter a residência aquecida (27,9%), muito acima das médias para a União Europeia (EU) a 27 (10,8%) e para a Zona Euro (10,0%).

Estas desigualdades parecem agravar-se na população em risco de pobreza, sugerindo também uma relação marcada entre estatuto socioeconómico e estes factores, particularmente a ausência de conforto térmico (Eurostat, 2014b).

Desta forma, a proporção de população abaixo do limiar de pobreza (conjunto de população com rendimento inferior a 60% do rendimento mediano por adulto equivalente) sem capacidade de custear o aquecimento apropriado da sua habitação excedeu entre 2004 e 2013 os 30% em nove dos estados-membros, seis dos quais se localizam no sul e no leste da Europa. Portugal inclui-se neste conjunto, com 56,9% em 2004 e 44,6% em 2013, juntamente com a Bulgária, Chipre, Grécia, Itália, Lituânia, Letónia, Hungria e Malta. Mas apesar da redução verificada entre 2004 e 2013, a situação portuguesa regista ainda valores bastante superiores às médias para a UE a 27 e Zona Euro, respectivamente de 24,1% e 23,6%.

Noutros indicadores relevantes, Portugal aparece em posições mais diversas no contexto europeu.

Se analisarmos a superlotação das habitações, a maioria dos países do Leste europeu registam os valores mais elevados em 2013, entre os 40% e os 50%, embora com

Condições precárias de habitação e mortalidade por acidente vascular cerebral ao nível da freguesia em Portugal Continental, na população com mais de 65 anos de idade

tendência geral para decréscimo entre 2008 e 2013, enquanto a maior parte dos países mais desenvolvidos da UE registam valores abaixo dos 10%, numa média de 17,1% a 27 países, e de 11,8%, na Zona Euro (para 2013). Em 2004, Portugal registava um valor de 15,3% mas a evoluir tendencialmente para uma diminuição até 2013, com 11,4%, próximo do valor médio da Zona Euro. Se se considerar o mesmo indicador para a população abaixo do limiar de pobreza, os valores disparam para uma média, a 27 países, de 30,1% (2013) e a Zona Euro com uma média estabilizada entre 23% e 24%, entre 2004 e 2013. É de referir que em alguns países europeus considerados mais desenvolvidos se registou aumento entre 2012 e 2013, como na Dinamarca (de 21% para 36%) ou em Itália (38,8% para 41,7%) ou ainda na Holanda (de 9,2% para 13,8%). Em Portugal, a percentagem de população em risco de pobreza a viver em casas superlotadas aumentou de 2004 (21,8%) para 2008 (25,1%) mas inverte a tendência em 2013 para 19,7%, percentagem inferior aos valores médios para a UE e Zona Euro.

Acerca da percentagem de população a residir em apartamentos no total da população residente, Portugal registava um valor de 35,5% em 2004, aumentando ao longo de 9 anos e atingindo 43,5% em 2013, valor acima da média da UE a 27 (41,3%) mas abaixo da média da Zona Euro (46,1%). Para a maioria dos países europeus, se as variações temporais entre 2004 e 2013 não são significativas, no entanto há diferenças relevantes a nível regional, entre países como a Alemanha, Espanha ou Itália, com mais de 50% da população a viver em apartamentos, e a Finlândia (33% em 2013) ou o Reino Unido (14,3% em 2013). Se a análise recair sobre o universo da população abaixo do limiar de pobreza, os valores percentuais são significativamente mais baixos, com médias de 7,6% na UE a 27 e de 9,3 na Zona Euro, em 2013, e uma tendência, na maioria dos países, para um ligeiro aumento entre 2004 e 2013. Portugal segue a tendência, com 4,7% em 2004 e 7,2% em 2013.

Comparando os indicadores percentagem de arrendatários e percentagem de proprietários do alojamento, para os 2 universos populacionais estão disponíveis valores para o ano de 2001. Nesse ano, havia 23% de arrendatários e 77% de proprietários no universo do total de população portuguesa, registando-se na UE a 15 uma média de 25% e 41%, respectivamente. No universo da população abaixo do limiar de pobreza 28% eram arrendatários e 72% proprietários, sendo a média da UE a 15 de 41% e 59%, respectivamente. Portugal destaca-se dos restantes estados membros para o conjunto de

Condições precárias de habitação e mortalidade por acidente vascular cerebral ao nível da freguesia em Portugal Continental, na população com mais de 65 anos de idade

população abaixo do limiar de pobreza, com uma percentagem de arrendatários consideravelmente inferior à média europeia, e uma percentagem bastante superior de proprietários. Numa década, até 2013, e em relação ao total da população portuguesa, o valor percentual de arrendatários registou um fraco aumento para os 25,8%, e, inversamente, uma ligeira diminuição na proporção de proprietários para os 74,2%, com a média da UE a 27 de 30,1% e 69,9%, e da Zona Euro de 33,6 e 66,4%, respectivamente. Já no universo da população abaixo do limiar de pobreza, o aumento percentual dos arrendatários foi significativo (de 28% em 2001 para 44% em 2013), mas mantendo-se sempre inferior em relação a todos os valores médios da UE a 27 e da Zona Euro, respectivamente, 49,1% e 56,3% e o peso de proprietários a registar variação inversa, com redução para os 56,0% em 2013, aproximando-se da média da UE a 27 (50,9%) mas ainda distante da média da Zona Euro (43,7%). Numa breve panorâmica na UE, num conjunto de 18 países, as diferenças são significativas no universo do total de população para 2001, com a Holanda a apresentar os valores mais elevados no peso de arrendatários (43,0%), seguida da França com 38%, países a evoluírem para uma diminuição até 2013, mais notória para o primeiro; em contrapartida, países do Leste Europeu como a Roménia, Eslovénia, Hungria e Lituânia, registaram os valores mais baixos de arrendatários (entre 3% e 6% do total de população), seguidos a pouca distância da Espanha (8%), todos a duplicarem ou triplicarem a respectiva percentagem passado uma década (excepto a Roménia); na proporção de proprietários a situação invertia-se em 2001, com a maioria dos países do leste europeu a apresentarem valores superiores a 90%, somente acompanhados pela Grécia e Espanha do lado ocidental e a Holanda a apresentar o valor mais baixo (57%), ao passo que, uma década depois, a tendência na maioria dos países é de decréscimo, mantendo-se apenas a Roménia, Lituânia e Eslováquia acima de 90%. No universo da população em risco de pobreza a percentagem de arrendatários, nos 20 países com dados para 2001, também era liderada pela Holanda (76%), logo seguida pela Suécia (74%) e Dinamarca (61%), verificando-se valores inferiores a 10% para a Roménia, Grécia e Lituânia; a situação inverte-se para os proprietários, semelhante ao quadro apresentado para o universo da população total. Em 2013 a tendência é para aumento da proporção de arrendatários na maioria de países (excepto para os 3 países líderes em 2001), diminuindo conseqüentemente a percentagem de proprietários. Em resumo, os

Condições precárias de habitação e mortalidade por acidente vascular cerebral ao nível da freguesia em Portugal Continental, na população com mais de 65 anos de idade

países com mais arrendatários que proprietários no universo da população com baixos rendimentos são os considerados mais desenvolvidos, da parte ocidental da Europa, estando os países do sul da Europa, Portugal incluído, numa situação inversa mas com evolução tendencial para a paridade. Constatou-se de facto um peso elevado de proprietários de casas nos países menos desenvolvidos.

Analisando agora algumas situações de privação severa das condições de habitação, estas apresentam-se relativamente díspares entre indicadores.

Assim, verificou-se uma disparidade significativa no indicador falta de instalação de banho ou duche entre 5 países do leste europeu (Roménia, Bulgária, Estónia, Letónia e Lituânia) e os restantes países da UE, quer no peso desta privação na população total quer na população abaixo do limiar de pobreza. A média da UE a 27 para o total de população é de 2,5% em 2013 e para a Zona Euro de 0,4%, enquanto a proporção na população mais pobre é de 7% e 1%, respectivamente e para o mesmo ano. Portugal registou decréscimos percentuais entre 2004 e 2013, quer para a população total quer para a população abaixo do limiar de pobreza, passando, respectivamente, de 4,4% para 1,1%, e de 9,8% para 1,6%, esta última uma redução mais significativa, mas ainda acima da média da Zona Euro em ambos os universos.

Os mesmos 5 países do leste europeu voltam de novo a apresentar valores consideravelmente mais elevados no indicador ausência de sanita com autoclismo (i.e. Roménia com 37% da população total), ultrapassando significativamente os valores médios da UE a 27, que se situam, em 2013, em 2,7% da população total e 7,2% do conjunto de população com rendimento abaixo do limiar de pobreza. Mas tal como no indicador anterior, estes 5 países registam decréscimo entre 2004 e 2013. Para a Zona Euro, registam-se médias, em 2013, de 0,5% da população total e 0,8% para o conjunto da população pobre. Em Portugal, regista-se também neste período uma descida para os dois universos populacionais, mas com médias sempre acima dos valores para a Zona Euro.

Todavia, constata-se no indicador tecto que deixa passar água e outras deficiências estruturais na habitação uma situação bastante diferente, em que Portugal se assume, para 2013, como o campeão no “ranking” dos países da UE com maior percentagem, 31,9% do total de população, o que totaliza cerca de 1/3 da população portuguesa e corresponde ao dobro da média na UE a 27 (15,7%) e na Zona Euro (16,4%). Temos

Condições precárias de habitação e mortalidade por acidente vascular cerebral ao nível da freguesia em Portugal Continental, na população com mais de 65 anos de idade

assim Portugal bastante acima dos valores percentuais da maioria dos países da UE, que registam valores inferiores a 20% e com uma subida muito significativa entre 2012 e 2013 da proporção de população afectada. Se o universo populacional for restrito à população abaixo do limiar de pobreza, Portugal, com 40,1%, é apenas ultrapassado em 2013 pela Hungria e pela Letónia e mantém-se bastante acima dos valores médios para a UE e Zona Euro (cerca de 24,0%), principalmente devido também ao expressivo crescimento entre 2012 e 2013.

Sintetizando a posição de Portugal face aos restantes estados membros da UE, um dos indicadores que mostram a posição muito desvantajosa do país é precisamente a proporção de pessoas sem capacidade de manter a habitação aquecida, particularmente se for considerado o universo da população portuguesa, e não apenas o peso no total de população em risco de pobreza. Nos indicadores de privação severa das condições de habitação, Portugal não acompanha a situação da maioria dos países da parte ocidental da Europa, com destaque para o peso significativo das habitações com deficiências estruturais, tais como aquelas cujo tecto deixa passar água.

Elaborando um pouco mais sobre alguns estudos focados apenas em Portugal, as más condições de habitabilidade e em particular o problema da incapacidade de aquecer adequadamente a habitação têm já sido objecto de investigação, especificamente na sua relação com o risco de doença cardiovascular. No contexto Europeu, os países com Invernos mais amenos são indicados como aqueles que apresentam mortalidade mais elevada por este tipo de doenças nesta época, e Portugal é de facto um dos países do sul da Europa com mais mortalidade por doença cardiovascular no Inverno (Vasconcelos et al., 2011). Estes autores levantaram a suspeita desta situação estar associada com as más condições de habitabilidade (especialmente ao nível de um deficiente isolamento térmico das habitações), tendo realizado um inquérito junto de pacientes que deram entrada em hospitais durante o Inverno, apresentando sintomas agudos de doença coronária. Entre os resultados deste inquérito, cerca de 42% dos entrevistados classificavam a sua habitação como “fria”. Aquando do aparecimento dos sintomas, cerca de 76% encontravam-se num espaço fechado, e destes, apenas metade indicaram possuir algum tipo de aquecimento. Mais ainda, destes últimos, apenas 54% tinham utilizado o aquecimento no Inverno anterior. Estes autores concluem que as más

Condições precárias de habitação e mortalidade por acidente vascular cerebral ao nível da freguesia em Portugal Continental, na população com mais de 65 anos de idade

condições de habitabilidade e/ou a ausência de protecção contra a exposição ao frio são frequentes em Portugal.

Todavia, a associação multivariada entre mortalidade por doenças cardiovasculares, exposição ao frio nas habitações e outras condições de habitabilidade permanece ainda pouco conhecida, sendo difícil de determinar. A este nível, alguns autores indicam que a influência das más condições de habitação pode fazer-se sentir de várias formas. Vasconcelos et al. (2011) referem alguns estudos realizados no Reino Unido, um dos quais conclui que as pessoas a viver em piores condições de habitabilidade e em áreas mais frias apresentavam valores de hipertensão 45% mais elevados. Noutro estudo, foi encontrada uma associação positiva entre a idade dos edifícios e o excesso de mortalidade no Inverno, sendo que a população residente em edifícios construídos antes de 1850 apresentava, comparativamente, um excesso de mortalidade nesta estação do ano. Por outras palavras, embora os efeitos da exposição a baixas temperaturas sejam relativamente conhecidos, não parece existir ainda uma percepção exacta dos efeitos de outros factores relacionados com a má qualidade das habitações, na mortalidade por doenças cardiovasculares em geral e por AVC em particular, quer em conjugação com a temperatura, quer observados isoladamente.

Se, como discutido até este ponto, residir numa habitação com condições deficientes de habitabilidade é um factor negativo para a saúde da população em geral e se encontra em geral relacionado com situações de reduzido estatuto socioeconómico, determinados tipos de problemas parecem afectar mais intensamente as faixas etárias mais idosas (Marsh et al., 1999). Tendo em conta que é nestas faixas que ocorre a maior porção de mortalidade e incidência de AVC, a população idosa poderá merecer especial atenção, no que diz respeito a estudos e intervenções de saúde pública focados em eliminar ou mitigar as situações de deficientes condições de habitabilidade.

Alguns autores têm abordado esta problemática do ponto de vista da saúde da população idosa em geral. Num estudo realizado no Reino Unido, na área do país de Gales, Windle et al. (2006) referem as dificuldades na habitação tais como ser-se arrendatário, a percepção de frio resultante de meios de aquecimento ineficazes e o excesso de horas passadas em casa como preditores de um pior estado de saúde. Costa-Font (2008) aborda a questão da habitação como um recurso na saúde dos idosos sob vários pontos de vista (por exemplo o efeito de ser ou não proprietário da habitação).

Condições precárias de habitação e mortalidade por acidente vascular cerebral ao nível da freguesia em Portugal Continental, na população com mais de 65 anos de idade

Outros estudos referem-se directamente à relação entre AVC e deficientes condições de habitabilidade na população idosa. Um relatório de uma organização não governamental sediada no Reino Unido (Age UK, 2012) refere especificamente a contribuição das baixas temperaturas na habitação para a ocorrência de episódios de AVC e outros problemas cardiovasculares, colocando especial ênfase no risco acrescido desta situação na população idosa (especialmente aquela com uma condição cardiovascular pré-existente), pois embora a diminuição da temperatura tenda a provocar a subida da pressão arterial em pessoas de todas as idades, na população idosa este aumento tem tendência a prolongar-se por muitas horas após a exposição a baixas temperaturas. Esta questão agrava-se quando se estima que as pessoas idosas passem cerca de 70 a 90% do seu tempo nas suas casas (Windle et al., 2006).

Do ponto de vista metodológico, vários autores têm recorrido a técnicas de análise de clusters no estudo da relação entre factores socioeconómicos e mortalidade e/ou incidência de AVC. Schieb et al. (2013) recorreram a técnicas de clustering espacial (métodos LISA) para delinear clusters de municípios (counties) nos Estados Unidos da América com valores altos (clusters alto-alto) ou baixos (clusters baixo-baixo) de taxas de hospitalização por AVC, tendo depois calculado estatísticas descritivas sobre os valores de diversas variáveis representativas das características socioeconómicas dos municípios incorporados em cada tipo de cluster e sua rede de cuidados de saúde. Este estudo pretendeu, desta forma, quantificar, com carácter exploratório, associações entre o risco de hospitalização devido a AVC e vários factores socioeconómicos medidos ao nível do município (e.g. o rendimento médio do agregado familiar ou a população abaixo do limiar de pobreza). Pedigo et al. (2011) recorreram, por seu lado, a técnicas de clustering não-espacial (técnicas de *Fuzzy Clustering* e análise de discriminantes) para identificar clusters de unidades censitárias (censos tracts) com características distintas, com o objectivo de fornecer apoio à decisão, no sentido de melhorar o direccionamento e priorização das intervenções de Saúde Pública ao nível do AVC e Enfarte Agudo do Miocárdio.

Decorrente do exposto, este estudo tem como objectivos: i) identificar um conjunto de variáveis representativas das más condições de habitabilidade que possam constituir determinantes do AVC na população geral ou com mais de 64 anos ao nível da Freguesia, em Portugal Continental (ii) analisar a distribuição espacial conjunta destes

determinantes e da mortalidade por AVC numa perspectiva univariada (iii) identificar padrões de clustering univariado e multivariado dos determinantes e sua relação com a mortalidade por AVC, como primeiro esforço para direccionar esforços de intervenção dos decisores em Saúde Pública.

3.2 Materiais e métodos

3.2.1 Área de estudo e fontes de dados

Este estudo foi operacionalizado no território continental Português, adoptando como unidade estatística a freguesia. As freguesias constituem Unidades Administrativas Locais, nível 2, de acordo com a Comissão Europeia (Eurostat, 2014a), podendo igualmente ser entendidas como unidades estatísticas de carácter territorial. O número destas unidades territoriais em Portugal continental tem variado ao longo do tempo devido a sucessivas reconfigurações administrativas, existindo 4005 freguesias em 1991, 4037 em 2001 e 4050 em 2011 (INE, 2014). A mais recente reconfiguração administrativa, aplicada em 2013, veio alterar radicalmente esta situação, pois em 2014 existiam já apenas 3223 destas unidades (DGT, 2015), situação que irá constituir um desafio para futuros estudos multitemporais. A base cartográfica adoptada foi a Base Geográfica de Referenciação de Informação (BGRI) de 2001, obtida a partir do Instituto Nacional de Estatística (INE, 2014), bem como todos os dados referentes aos indicadores de qualidade da Habitação utilizados neste estudo. Todas as variáveis causais relativas à qualidade da habitação, assim como a população das freguesias, são referentes ao ano censitário de 2001. Os dados de mortalidade por AVC ao nível da freguesia foram gentilmente cedidos pela Direcção-Geral de Saúde (DGS) e abrangem um período entre 1 de Janeiro de 1998 a 31 de Dezembro de 2004. Neste período de 7 anos, foram registados um total de 134900 óbitos por AVC nas freguesias do território continental.

3.2.2 Cálculo do indicador de risco de AVC, variáveis explicativas e análise espacial visual

Numa primeira fase observou-se a distribuição espacial do risco de morte por AVC nas freguesias, no período de estudo considerado. Para representação do risco de morte por AVC, e devido ao confundimento que pode advir da existência de diferentes estruturas etárias entre freguesias, optou-se aqui por recorrer ao Rácio Padronizado de Mortalidade ou *Standard Mortality Ratio* (SMR), o qual constitui uma alternativa ao método indirecto de ajustamento de taxas por estrutura etária, sendo frequentemente utilizado em estudos que empregam dados socioeconómicos ou realizam análises comparativas sobre um grande número de áreas de pequena dimensão (Siegel, 2012). Desta forma, este indicador constitui o factor na fórmula de ajustamento indirecto de taxas que representa a mortalidade comparativa de duas populações, assumindo-se uma distribuição etária comum de referência. O cálculo deste indicador obedece à seguinte fórmula:

$$SMR = \frac{d}{\sum M_a p_a} * 100$$

na qual M_a representa as taxas específicas de mortalidade na população de referência, d o total de óbitos e p_a a população existente em cada escalão etário na população em estudo. Por outras palavras, o SMR corresponde, para cada área, à razão entre casos observados e casos esperados no período em estudo, constituindo um estimador do risco relativo associado a essa área, ou seja, o risco de morte em relação a um grupo considerado como referência.

Este indicador foi calculado para seis coortes, combinando pessoas com mais de 64 anos de idade ou de todas as idades, homens ou mulheres ou ambos os sexos, no período de estudo de 7 anos entre 1998 e 2004, tendo-se assumido como população em risco a população residente em cada freguesia de Portugal Continental em 2001 (ano escolhido por constituir o ano mediano do período de estudo) e como população padrão a estrutura etária da população de Portugal Continental nesse ano.

Condições precárias de habitação e mortalidade por acidente vascular cerebral ao nível da freguesia em Portugal Continental, na população com mais de 65 anos de idade

Todavia, o SMR de cada área (SMR_i) possui algumas limitações, não obstante ser utilizado com frequência na produção de cartografias ilustrando a distribuição espacial da mortalidade. Uma das limitações mais importantes prende-se com a sua dependência face à dimensão populacional das áreas para o qual são calculados. Sendo a variância dos SMR_i inversamente proporcional aos valores *esperados*, as áreas com pouca população irão apresentar estimadores com grande variabilidade, sendo, por consequência nestas áreas que surgem com mais frequência valores extremos que, para além de serem estimados com pouca precisão, têm também tendência para dominar o padrão geográfico percebido nos mapas.

Mais ainda, a variabilidade dos casos *observados* pode ser bastante maior do que a esperada para uma distribuição de Poisson, situação geralmente designada como sobredispersão, e que diminui igualmente o rigor da estimação de mortalidade.

Revisitando a questão das áreas com pouca população, estas podem ser entendidas como *áreas pequenas*, embora não exista para estas uma definição universalmente aceite, na medida em que dependerá do contexto do estudo e do número de ocorrências da doença. Nesse sentido, alguns autores consideram como pequena qualquer área cujo tamanho amostral no contexto do estudo é demasiado pequeno para conseguir produzir estimações com um nível de precisão aceitável. Outros são mais específicos, considerando-as como áreas com menos de 20 ocorrências da doença numa população com pelo menos 10000 habitantes. Ainda assim, apesar do inconveniente que representa o pequeno número de ocorrências, estas áreas têm a vantagem de serem em geral mais homogêneas em termos de exposição aos factores determinantes da doença.

É necessário ainda ter-se em conta que, na presença de dados espaciais, devem ser consideradas duas causas potenciais de variabilidade excessiva, sendo a mais relevante a chamada dependência espacial que decorre da correlação da unidade espacial com as suas unidades vizinhas (autocorrelação espacial), geralmente as unidades contíguas. Este efeito leva a que os SMR_i de unidades mais próximas ou mesmo contíguas seja mais similar que os SMR_i de unidades mais distantes, embora parte desta dependência possa ter origem na existência de variáveis (igualmente com estrutura espacial) não incluídas na análise. A segunda fonte potencial deve-se à existência de variabilidade excessiva independente e não correlacionada espacialmente, também designada

Condições precárias de habitação e mortalidade por acidente vascular cerebral ao nível da freguesia em Portugal Continental, na população com mais de 65 anos de idade

heterogeneidade não espacial, e que se deve à existência de variáveis não observadas, mas desta vez sem estrutura espacial (Barceló et al., 2008).

Com o fim de mitigar os problemas que resultam da utilização dos SMR_i sem qualquer tipo de ajuste, foram propostas várias soluções para diminuir as flutuações extremas da sua variância, em geral designadas como formas de “suavizar” este indicador. Os métodos de suavização geralmente utilizados recorrem a modelos hierárquicos Bayesianos, sendo os dois mais frequentes utilizados (Faltin et al., 2012), o modelo Empírico Bayesiano, também designado modelo Poisson-Gamma (Clayton e Kaldor, 1987) e o modelo de Besag, York e Mollié (BYM) (Besag et al., 1991).

Estes dois métodos apresentam diferenças importantes na sua forma de aplicação. No caso do Modelo Empírico Bayesiano, para além de se assumir uma distribuição de probabilidade para os casos observados (normalmente uma distribuição de Poisson), assume-se que o risco relativo constitui igualmente uma variável aleatória, o que torna possível assumir que obedece também a uma determinada distribuição de probabilidade (sendo geralmente adoptada uma distribuição Gamma).

Por outro lado, no modelo BYM, a solução adoptada consiste em incorporar no modelo dois efeitos aleatórios capazes de “absorver” a variabilidade não explicada, sendo que um destes efeitos absorve a dependência espacial e o outro recolhe a sobredispersão de origem não espacial, ou seja, que não possui estrutura espacial. Esta capacidade do modelo BYM de lidar com a sobredispersão que decorre da dependência espacial confere-lhe desta forma vantagem sobre o modelo Empírico Bayesiano, o que faz com que o modelo BYM seja actualmente o mais frequentemente na suavização dos valores dos SMR_i (Barceló et al., 2008).

Em suma, não só as freguesias, pelas suas características (gerais) no que respeita ao quantitativo populacional e número de casos observados, parecem aproximar-se singularmente da noção de *área pequena* previamente discutida, como representam também dados espaciais e como tal, potencialmente sujeitos a dependência espacial. Por estas razões, optou-se pela suavização dos vários SMR_i atrás referidos através do método BYM, sendo que toda a análise (do ponto de vista da mortalidade) se desenvolveu com base nos valores suavizados deste indicador.

No que respeita às variáveis explicativas representativas dos determinantes, foram investigados neste estudo um conjunto de indicadores relacionados com a qualidade da

Condições precárias de habitação e mortalidade por acidente vascular cerebral ao nível da freguesia em Portugal Continental, na população com mais de 65 anos de idade

habitação, cuja escolha foi orientada essencialmente por obedecerem (na medida do possível e disponível) a quatro condicionantes, nomeadamente: os resultados da análise bibliográfica previamente apresentados; a sua disponibilidade imediata a partir do Instituto Nacional de Estatística; a sua construção que, ao incluir na sua maioria (excepto em dois dos indicadores) a população residente (dado encontrarem-se todas referidas à data dos censos 2001), permite estabelecer aqui uma relação mais estreita com a mortalidade e a população em risco; e finalmente por terem sido recolhidas ao nível de desagregação territorial da freguesia. Este estudo focou-se assim em sete variáveis explicativas, nomeadamente: percentagem de residentes sem qualquer tipo de aquecimento; percentagem de alojamentos familiares sem aquecimento central; percentagem de alojamentos superlotados; percentagem de residentes a viver em edifícios construídos até 1960; peso da população residente a viver em alojamentos clássicos arrendados na população residente proprietária do alojamento (em percentagem); percentagem de residentes a viver em alojamentos familiares não clássicos; percentagem de residentes a viver em apartamentos.

De seguida, e como forma inicial de análise espacial, foram gerados também gerados mapas coropletos representativos da distribuição espacial tanto dos rácios padronizados de mortalidade como das variáveis explicativas. Estes mapas, tal como aqueles que ilustram a distribuição do risco de AVC nos vários coortes, foram produzidos em ESRI ArcMap (ESRI, 2014), tendo-se utilizado como escala de classificação o método de Jenks (Longley et al., 2010), com o objectivo de se obter intervalos de valores apropriados para a visualização espacial das variáveis. Desta forma, foi possível uma primeira análise comparativa da distribuição espacial das variáveis em estudo.

3.2.3 Análise univariada de clustering espacial

O nível de *clustering espacial* das variáveis ao nível da freguesia foi examinado, tanto do risco de morte por AVC como das variáveis representativas das más condições de habitabilidade. Isto permitiu ter uma primeira percepção da tendência que as freguesias apresentam para se aglomerarem em determinadas áreas do território nacional (ou não), com base nos valores dos indicadores registados a este nível de desagregação administrativa.

Condições precárias de habitação e mortalidade por acidente vascular cerebral ao nível da freguesia em Portugal Continental, na população com mais de 65 anos de idade

Esta análise apoiou-se nos índices I de Moran Global e Local (Pfeiffer et al., 2008), calculados em ambiente Geoda (GeoDa Center, 2015), e permitiu não só uma exploração da existência de autocorrelação espacial nas variáveis, mas também (tal como no uso prévio dos mapas coropletos) realizar uma análise comparativa da magnitude do clustering espacial (através do índice I de Moran Global) e da localização e padrões dos clusters espaciais das variáveis explicativas *versus* os rácios padronizados de mortalidade (através do índice I de Moran Local, também denominado métodos LISA - *Local Indicators of Spatial Autocorrelation*), tentando-se não apenas perceber áreas de aglomeração mas também entender até que ponto coincidem no espaço os valores dos SMRi e das variáveis explicativas.

O índice I de Moran Local tem como objectivo avaliar a presença de autocorrelação espacial e identifica desta forma unidades espaciais (ou seja, freguesias) com valores de SMRi estatisticamente similares ou dissimilares aos valores das unidades espaciais vizinhas. Este método assume como hipótese nula que não existe associação entre os SMRi de freguesias vizinhas, e como hipótese alternativa que existe clustering espacial, por outras palavras que freguesias vizinhas apresentam valores similares de SMRi.

Para a definição das freguesias vizinhas foi adoptado aqui um critério de contiguidade baseado na existência de qualquer ponto de contacto entre os limites administrativos das freguesias (*Queen Contiguity*).

Este método detecta quatro tipos de fenómenos relacionados com o clustering, nomeadamente clusters de valores altos ou baixos (tipicamente designados clusters do tipo alto-alto ou baixo-baixo) ou localizações do tipo alto-baixo ou baixo-alto. Enquanto os primeiros dois tipos denotam autocorrelação espacial positiva e representam clusters no verdadeiro sentido da palavra, os segundos consistem em localizações isoladas, representam autocorrelação espacial negativa e são geralmente designados *outliers espaciais*.

É importante reter que, no mapa resultante da aplicação deste método, são apenas assinaladas como clusters espaciais as freguesias assinaladas aquelas que se encontram no centro do cluster, sendo que o cluster propriamente dito, para além da freguesia central, engloba também as suas vizinhas, definidas de acordo com o critério de contiguidade adoptado (Anselin, 2005).

Condições precárias de habitação e mortalidade por acidente vascular cerebral ao nível da freguesia em Portugal Continental, na população com mais de 65 anos de idade

Esta situação gera frequentemente situações de sobreposição de clusters de tipos opostos, devido à possibilidade de ocorrência de múltiplas comparações (Rainey et al., 2007). Os clusters sobrepostos são geralmente excluídos, pois a sua ocorrência não permite estabelecer as fronteiras exactas do cluster (Kosfeld e Lauridsen, 2012). Uma das formas de minimizar a ocorrência deste fenómeno consiste em regular o nível de pseudo-significância necessário para a aceitação de um cluster como válido. Desta forma, optou-se neste estudo por obter p-values através do método de Monte Carlo, com base em 9999 randomizações condicionais e adoptando-se um nível de significância $p < 0.01$. Adicionalmente, foram excluídos da análise quaisquer clusters que apresentassem sobreposições entre si.

Após a determinação dos clusters espaciais, uma primeira medida de similaridade espacial consistiu em determinar clusters de valores altos e de valores baixos para cada variável explicativa e calcular estatísticas da mortalidade nas freguesias dentro dos clusters pertencentes aos dois tipos (similarmente a Schieb et al., 2013), mais especificamente a média e a mediana dos SMR_i, o que permitiu obter uma primeira aproximação às diversas tendências da mortalidade nas freguesias pertencentes a cada tipo de cluster, embora ainda numa perspectiva univariada.

Antes de se prosseguir para a análise multivariada das variáveis explicativas, efectuou-se uma avaliação prévia da presença de multicolinearidade entre estas variáveis explicativas. Para tal, recorreu-se ao cálculo do coeficiente de correlação linear de Pearson (Webster e Oliver, 2007) entre pares de variáveis para tentar detectar variáveis independentes altamente correlacionadas entre si. Impunha-se igualmente definir critérios que guiassem a aceitação ou exclusão de variáveis com base na presença de multicolinearidade. Desta forma, adoptou-se o critério de reter apenas uma das variáveis de um par altamente correlacionado para a fase subsequente de análise multivariada, assumindo-se como limiar um valor do coeficiente $r \geq 0,6$ (Weisent et al., 2012).

3.2.4 Análise de clustering multivariada

Tendo em conta que a realidade é mais complexa do que a descrita por cada uma das variáveis em estudo, ou por outras palavras, que as diversas variáveis explicativas não actuam no território da freguesia de forma isolada mas sim simultaneamente, é

importante obter-se uma percepção da sua importância relativa num determinado território e, no contexto desse território, qual o comportamento da mortalidade por AVC. Nesse sentido, recorreu-se a técnicas de análise de clustering multivariadas, no sentido de procurar aglomerações de freguesias no território de Portugal Continental que sejam determinadas pelo efeito combinado das variáveis explicativas.

Antes de mais, é importante referir que, ao contrário do método de clustering univariado referido no ponto anterior, que constitui uma técnica de análise de clustering espacial, as técnicas de análise multivariada de clusters empregues nesta fase não lidam directamente com a componente espacial dos dados. Todavia, um dos objectivos desta análise é que as aglomerações de freguesias possuam (na medida do possível) estrutura espacial, por outras palavras, expressão e coerência espacial ao longo do território. Este efeito foi induzido através da inclusão na análise de duas variáveis adicionais, nomeadamente as coordenadas X e Y dos pontos representativos dos centróides (ou centros geométricos) das freguesias. Estas localizações pontuais funcionam como uma referência artificial que fornece uma medida descritiva da área em causa, constituindo uma medida bi-dimensional equivalente à média (Longley et al., 2005). Estas coordenadas encontram-se expressas em metros e referidas ao sistema de coordenadas ETRS89 / Portugal TM06 (EPSG, 2015).

Esta é uma solução frequente para lidar com dados espaciais representados no contexto dos métodos de clustering hierárquicos não-espaciais, sendo inclusive a mais simples, na medida em que estes métodos permitem, pela sua grande flexibilidade, codificar não apenas as localizações geográficas mais simples como as coordenadas x e y, mas também muitos outros tipos de relações espaciais, tais como distâncias numa rede de estradas, entre outros (Guo e Gahegan, 2006).

De um modo geral, as técnicas de análise de clusters constituem ferramentas de análise exploratória de dados cujo propósito se foca em resolver problemas de classificação, sendo este propósito universal, quer nos estejamos a referir a técnicas de análise de clusters espaciais ou não espaciais. Pretende-se desta forma ordenar dados representativos de algo que se pretende estudar (fenómenos, pessoas, objectos...) em grupos, designados *clusters*, sendo que estes grupos devem apresentar, por um lado, o máximo de homogeneidade interna (dentro do cluster) e por outro o máximo de heterogeneidade externa (entre clusters) (Salah et al., 2012); (Ahmad et al., 2013).

As formas de análise de clusters mais utilizadas dividem-se em *técnicas de optimização* ou em *técnicas hierárquicas*, estas últimas mais frequentemente utilizadas. Quanto às técnicas de optimização, estas apoiam-se simultaneamente num determinado critério de agrupamento de casos e na predefinição de um número de grupos (clusters). Desta maneira, a optimização (quer por minimização ou maximização do critério) levará a que cada caso venha a ser incluído num dos grupos predefinidos. No que respeita às técnicas hierárquicas, estas separam-se em divisivas e aglomerativas. As técnicas divisivas começam por um único grupo de partida e dividem-no sucessivamente em subgrupos até chegarem ao ponto em que o número de grupos iguala o número de indivíduos, sendo geralmente menos utilizadas por serem normalmente mais exigentes do ponto de vista computacional.

Quanto às técnicas hierárquicas aglomerativas, estas têm como ponto de partida a definição de uma matriz de distâncias entre indivíduos, que corresponde a n grupos de partida, tantos quantos o número de indivíduos na análise, e vão alocando sucessivamente os indivíduos a grupos, até ser atingido um único grande grupo que aglomera todos os n elementos. O processo de agrupamento desenrola-se através do sucessivo relaxamento da proximidade absoluta (aquela que existe entre cada indivíduo e si próprio), sendo assim cada indivíduo agregado ao indivíduo que lhe está mais próximo, de seguida cada grupo resultante é agrupado ao grupo mais próximo e assim sucessivamente, num processo que define uma hierarquia, na medida em que cada nível de agregação irá constituir no fundo um nível hierárquico. Desta forma, um requisito incontornável nestas técnicas consiste na escolha de um critério de agregação que permita definir a proximidade entre grupos e controlar desta forma o processo de aglomeração (Mendes e Bação, 1998).

Após consideração das opções, optou-se neste estudo por recorrer a uma técnica hierárquica aglomerativa, mais apropriada quando não existe (como é o caso aqui) uma definição *a priori* do número de clusters em que se pretende agrupar os casos. Com o objectivo de calibrar o método, optou-se por uma abordagem similar a um teste de sensibilidade, pois no decurso da aplicação destas técnicas, o analista enfrenta em geral três desafios, nomeadamente a escolha da medida de distância entre vectores de dados (ou seja, os atributos), do critério de agregação e do número final de clusters (Carvalho et al., 2009).

Condições precárias de habitação e mortalidade por acidente vascular cerebral ao nível da freguesia em Portugal Continental, na população com mais de 65 anos de idade

No que respeita ao critério de agregação, optou-se por utilizar aqui o critério de agregação de Ward (Ward, Jr, 1963). Este critério parece adaptar-se relativamente bem aos objectivos deste estudo do ponto de vista de determinar clusters com alguma continuidade e/ou coerência espacial, pois este método assume que os pontos podem ser representados no espaço euclidiano para interpretação geométrica, para além de ter tendência a encontrar clusters de tamanhos similares, de configuração compacta, aproximadamente esférica (Everitt et al., 2011). Esta característica, aliada ao facto de, como atrás referido, se recorrer adicionalmente à utilização das coordenadas dos centros geométricos das freguesias na tentativa de induzir continuidade espacial ao mecanismo da análise, favorece o recurso a este critério, neste contexto.

Este critério recorre já, em cada passo, a uma análise de variância (ANOVA) para avaliar a distância entre grupos, no sentido de minimizar a soma dos quadrados de cada par de grupos em cada passo (Salah et al., 2012). Nesta técnica, a distância entre dois grupos é medida pela soma dos quadrados dos desvios entre as observações de cada indivíduo face às médias dos grupos a que os indivíduos foram alocados (Mendes e Bação, 1998). Adicionalmente, possui um bom grau de robustez, sendo o mais robusto de entre os que utilizam matrizes de similaridade Townsend (1993), e sendo também superior no que respeita à capacidade de lidar com o “ruído” presente nos dados (Balcan et al., 2013).

Foram assim testadas as combinações do critério de agregação de Ward com várias distâncias disponíveis no pacote de software estatístico utilizado, nomeadamente as distâncias de Gower, Euclideana, Manhattan, Minkowski e Maximum (Everitt et al., 2011). Estes testes permitiram testar combinações baseadas em diferentes níveis de complexidade, recorrendo-se tanto a medidas relativamente populares (como a euclidiana) como a outras menos frequentes (como a de Gower). Esta diferença é evidente na comparação entre as distâncias Euclidiana e de Gower. Assim, a distância Euclideana entre dois elementos (i e j) pode ser calculada de forma relativamente simples, pela extracção da raiz quadrada do somatório dos quadrados das diferenças entre os valores de i e j para todas as variáveis ($v=1,2, \dots, p$) (Mendes e Bação, 1998), como pode ser observado de seguida:

$$d_{ij} = \sqrt{\sum_{v=1}^p (X_{iv} - X_{jv})^2}$$

Por seu lado, a distância de Gower (1971) consistiu uma medida mais complexa, na medida em que opera uma forma de standardização sobre os dados, sendo os valores das distâncias entre as observações de cada variável convertidos previamente numa escala entre 0 e 1 e de seguida somadas todas as distâncias padronizadas entre as observações i e j para cada variável.

Desta forma, a distância entre dois valores i e j em cada variável é função da fórmula de padronização, na qual X_{if} representa, na fórmula seguinte, o valor de cada variável f no objecto i e R_f representa a soma de todos os valores possíveis da variável f ,

$$d_{ij}^{(f)} = \frac{|x_{if} - x_{jf}|}{R_f}$$

sendo o último passo a agregação das distâncias entre i e j para p variáveis:

$$d(i, j) = \frac{1}{p} \sum_{f=1}^p d_{ij}^{(f)}$$

Adicionalmente, a estrutura da distância de Gower permite lidar com outro condicionante, relacionado com as escalas de valores em que se encontram registadas as diversas variáveis envolvidas neste processo de análise de clusters. Assim, embora todas as variáveis incluídas nesta análise sejam numéricas e estejam expressas numa escala de rácio (Gravetter e Wallnau, 2013), os seus intervalos de valores variam grandemente, o que faria com que, caso fossem utilizadas sem qualquer tipo de ajustamento, tivessem o efeito de atribuir maior peso às variáveis com intervalos de valores mais elevados. Este não é um efeito desejado aqui, visto que as variáveis com maior intervalo são as variáveis X e Y contendo as coordenadas dos centróides das freguesias, correndo-se o risco de anular totalmente a contribuição específica de cada variável explicativa em função do efeito do próprio posicionamento das freguesias, quando pelo contrário se deseja que todas as variáveis tenham o mesmo peso na análise. Mais ainda, como se depreende, as variáveis expressam diferentes unidades de medida, pois enquanto as variáveis explicativas medem proporções e estão expressas em percentagens, os valores das coordenadas expressam distâncias expressas em metros.

Devido a este contexto, procedeu-se neste estudo à standardização das variáveis (X), previamente aos testes de clustering hierárquico (excepto naqueles envolvendo a

Condições precárias de habitação e mortalidade por acidente vascular cerebral ao nível da freguesia em Portugal Continental, na população com mais de 65 anos de idade

distância de Gower), um processo que recorre à média (μ) e desvio-padrão (σ) das variáveis, consistindo na sua transformação em novas variáveis (Z) com média nula e desvio-padrão igual à unidade:

$$Z = \frac{X - \mu}{\sigma}$$

No caso específico aos testes envolvendo a distância de Gower, optou-se por utilizar as variáveis sem aplicar previamente este processo, na medida em que o cálculo desta distância inclui já uma forma de standardização.

Uma das formas mais utilizadas para descrever e observar os resultados desta técnica de clustering consiste em recorrer a um dendrograma (ou árvore binária), que permite fornecer uma descrição visual dos passos do processo de agregação, ao ilustrar os grupos e a sua proximidade (Salah et al., 2012).

Assim, para cada combinação do critério de Ward com cada uma das medidas de distância, foi produzido um dendrograma, o qual foi utilizado para uma primeira avaliação visual dos resultados e do desempenho comparativo de cada teste.

Após a obtenção dos resultados de cada combinação distância/critério de agregação, recorreu-se a vários critérios para a escolha da melhor combinação distância/critério de agregação e também para a escolha do melhor número de clusters.

Neste processo de decisão, a primeira fase consistiu em produzir e observar os dendrogramas referentes a diversas combinações distância/critério de agregação, obtendo-se desde logo uma percepção de partida do desempenho de cada teste, visto ser possível avaliar visualmente se o dendrograma apresenta ou não uma estrutura equilibrada de distribuição das freguesias pelos clusters propostos, e avaliar igualmente se o dendrograma permite logo à partida obter uma sugestão do número ideal de clusters ou se pelo contrário esta escolha se apresenta confusa.

Neste passo, recorreu-se também ao cálculo do coeficiente de correlação cofenética (R_c) para cada combinação distância/critério de agregação. Este método, proposto por (Sokal e Rohlf, 1962), consiste no valor da correlação entre os elementos da matriz de distâncias entre os indivíduos e os elementos da matriz dos seus coeficientes de fusão, por outras palavras, das distâncias a que os indivíduos se juntam pela primeira vez para formar grupos. Ou seja, fornece uma medida da relação entre cada valor da matriz de distâncias e um valor recolhido a partir do dendrograma, indicando assim em que

Condições precárias de habitação e mortalidade por acidente vascular cerebral ao nível da freguesia em Portugal Continental, na população com mais de 65 anos de idade

medida o dendrograma que resulta da aplicação do método hierárquico é representativo dos valores da matriz de distâncias inicial. Embora não existam respostas absolutas na escolha da combinação distância/critério de agregação mais apropriada, este coeficiente constitui presentemente o critério mais utilizado (Mendes e Bação, 1998).

Após a observação do dendrograma, tentou-se determinar, para cada combinação, um número apropriado de clusters, o que pode aqui ser entendido como proceder ao “corte” do dendrograma num determinado nível de agregação.

Nesta decisão, foram observados vários critérios. Um destes consistiu em recorrer ao método sugerido por Mendes e Bação (1998) o qual consiste em comparar graficamente o número de clusters com o coeficiente de fusão, ou seja, a medição de distância para a qual diversos casos se juntam para dar origem a um grupo. Desta forma, tenta-se detectar quebras substanciais na curva do gráfico, procurando-se as situações em que a transição entre os valores dos coeficientes de fusão é feita de forma irregular e abrupta, o que irá provocar grandes descontinuidades na curva, levando esta a verticalizar-se. Por outras palavras, quando o valor do coeficiente de fusão aumenta de forma substancial pela agregação de um novo grupo, poder-se-á considerar essa partição como sendo ótima.

Outro critério consistiu, após uma primeira determinação do número de partições, em avaliar o equilíbrio na distribuição das freguesias pelos clusters, contabilizando-se o número de freguesias em cada cluster após cada decisão de corte do dendrograma e procurando-se desta forma evitar situações de grande desequilíbrio na distribuição das freguesias tais como, por exemplo, a existência de um cluster que concentra a maioria das freguesias, enquanto os outros incluem, inversamente, um número muito reduzido de freguesias (no limite apenas uma).

Algumas situações de avaliação do gráfico que compara distâncias de agregação com o número de clusters sugeriam a possibilidade de várias distâncias de corte. Nestas situações, testou-se por vezes mais do que um corte para cada dendrograma e observou-se de novo o balanceamento na distribuição das freguesias dentro dos clusters resultantes do respectivo corte.

Para resolver situações de indecisão acerca do número final de clusters, aplicou-se também um critério baseado na determinação do grau de significância estatística da distribuição dos valores de cada variável explicativa, associados às freguesias

Condições precárias de habitação e mortalidade por acidente vascular cerebral ao nível da freguesia em Portugal Continental, na população com mais de 65 anos de idade

englobadas pelos clusters de freguesias escolhidos, após cada iteração efectuada na escolha do número de clusters. Este teste apoiou-se na realização de análises de variância (ANOVA) sobre a associação entre a pertença da freguesia a um determinado cluster (representada pelo identificador de cada freguesia) e a respectiva distribuição dos valores das variáveis. Assumiu-se assim que, quanto melhores os valores de significância obtidos, mais apropriada seria uma determinada escolha do número de clusters. Este teste foi igualmente realizado para os valores dos diversos SMR_i, tentando-se avaliar de igual forma a significância estatística da distribuição diferenciada dos vários indicadores de mortalidade testados dentro dos clusters extraídos em cada iteração. Ou, por outras palavras, se as (sub)distribuições das variáveis nas freguesias dentro de um cluster se apresentavam significativamente diferentes, do ponto de vista estatístico, das (sub)distribuições obtidas a partir das freguesias incluídas nos outros clusters.

Durante o processo de determinação do número mais apropriado de clusters, foram produzidos interactivamente mapas coropletos com a distribuição espacial das freguesias por cada cluster em cada iteração, no sentido de possibilitar também uma avaliação visual de cada escolha.

Após operada a escolha final da combinação distância/critério de agregação e do número de clusters mais adequado, foram calculadas medidas de estatística descritiva (nomeadamente a média e a mediana) das variáveis explicativas, previamente estandardizadas, associadas às freguesias contidas em cada cluster final. Este processo, estruturalmente similar ao utilizado por (Schieb et al., 2013), é aqui aplicado com o fito de determinar, para cada cluster, qual das variáveis explicativas apresenta uma média mais alta, assumindo-se essa variável, como “dominante” no contexto das freguesias desse cluster. Este procedimento permitiu também obter mais uma medida da bondade da escolha final de clusters, ao permitir, em conjunto com a observação dos mapas coropletos de distribuição dos clusters, verificar a existência potencial de clusters contíguos com a mesma variável dominante.

No último passo da metodologia, procedeu-se ao cálculo das médias e medianas dos SMR_i correspondentes aos vários coortes nas freguesias incluídas em cada cluster final, com o objectivo de determinar quais os clusters em que a mortalidade é comparativamente mais elevada, o que, em conjunção com a variável explicativa

entendida como dominante em cada cluster, permitiu formular várias hipóteses acerca da importância relativa dos determinantes para a ocorrência de AVC ao longo do território de Portugal Continental no período abrangido por este estudo.

Na que respeita às ferramentas de software utilizadas, toda a produção de mapas coropletos foi efectuada em ambiente ESRI ArcMap (ESRI, 2014), as análises de clustering espacial univariado foram efectuadas em ambiente GeoDa (GeoDa, 2014), e as análises de clustering multivariado bem como todo o restante processamento estatístico, foram efectuadas com recurso à plataforma R (R, 2014a) e em ambiente RStudio (RStudio, 2015).

3.3 Resultados

3.3.1 Distribuição espacial do risco de AVC e dos determinantes relacionados com a qualidade da habitação

Os valores estimados por freguesia dos Rácios Padronizados de Mortalidade para cada município i (SMR $_i$) suavizados apresentavam alguma variação entre as 4037 freguesias do território de Portugal Continental no período em estudo de 1998-2004. É relevante notar que em todas as freguesias ocorreriam mais mortes que o esperado, em todos os 6 diferentes coortes populacionais estudados, caso a estrutura etária da população residente na freguesia fosse a mesma de Portugal Continental em 2001. Dentro dos diversos coortes, a variação é mais acentuada entre os mínimos e máximos dos SMR $_i$ do que entre coortes e também (embora com uma diferença menos apreciável) entre os valores registados por sexos. A maior amplitude de valores de risco ocorre na mortalidade em pessoas do sexo feminino com mais de 64 anos, variando entre um excesso mínimo de mortalidade de 2,94 e um valor máximo de 24,05, sendo neste coorte que se regista também o valor mais elevado de excesso de mortalidade de entre as freguesias. Nos dois coortes femininos verificam-se limiares superiores acima do ocorrido nos coortes masculinos, onde o valor mais elevado é de 15,81 nos homens com mais de 64 anos. Inversamente, em ambos os coortes masculinos se verificam limites mínimos mais elevados do que nos coortes femininos. É de notar alguma proximidade

Condições precárias de habitação e mortalidade por acidente vascular cerebral ao nível da freguesia em Portugal Continental, na população com mais de 65 anos de idade

entre os limiares superiores que ocorrem nos coortes englobando todos os sexos e os valores ocorridos nos coortes femininos. Verifica-se igualmente, comparando os coortes inclusivos de todas as idades com os coortes de pessoas com mais de 64 anos, que estes últimos apresentam sempre valores mais elevados nos limiares superiores. Ao nível dos valores médios e medianos, a similaridade é grande entre todos os coortes (tabela 8).

Tabela 8 – Estatísticas descritivas dos rácios padronizados de mortalidade por coorte

	Homens, mais de 64 anos de idade	Homens, todas as idades	Mulheres, mais de 64 anos de idade	Mulheres, todas as idades	Homens e Mulheres, mais de 64 anos de idade	Homens e Mulheres, todas as idades
Mínimo	3,12	3,16	2,94	2,89	2,76	3,18
Máximo	15,81	15,17	24,05	22,34	22,41	21,91
Média	7,49	7,43	7,42	7,40	7,46	7,43
Mediana	7,27	7,22	7,02	7,03	7,08	7,07
Amplitude	12,69	12,01	21,11	19,45	19,65	18,73
Desvio-Padrão	1,60	1,55	2,15	2,08	2,06	1,99

No que respeita à distribuição geográfica dos SMRi, nota-se a existência de um padrão espacial relativamente bem delineado nas freguesias ao longo do território de Portugal, o qual é similar entre os 6 coortes representados pelos vários SMRi, embora com algumas diferenças, particularmente entre sexos. Assim, na mortalidade em homens acima de 64 anos é notória uma zona de valores mais elevados de excesso de mortalidade na área noroeste e centro norte interior do território, surgindo ainda uma outra área com valores elevados, englobando a região norte do distrito de Lisboa e distrito de Leiria, e finalmente algumas manchas menos concentradas em vários pontos do território, como por exemplo no interior sudeste (figura 65a). Quanto ao excesso de mortalidade em homens de todas as idades, este apresenta um padrão quase idêntico ao do coorte com mais de 64 anos de idade, com ligeiras diferenças, entre as quais um pequeno aumento da compacidade das manchas de valores altos na área interior sudeste alentejano (figura 65b).

Condições precárias de habitação e mortalidade por acidente vascular cerebral ao nível da freguesia em Portugal Continental, na população com mais de 65 anos de idade

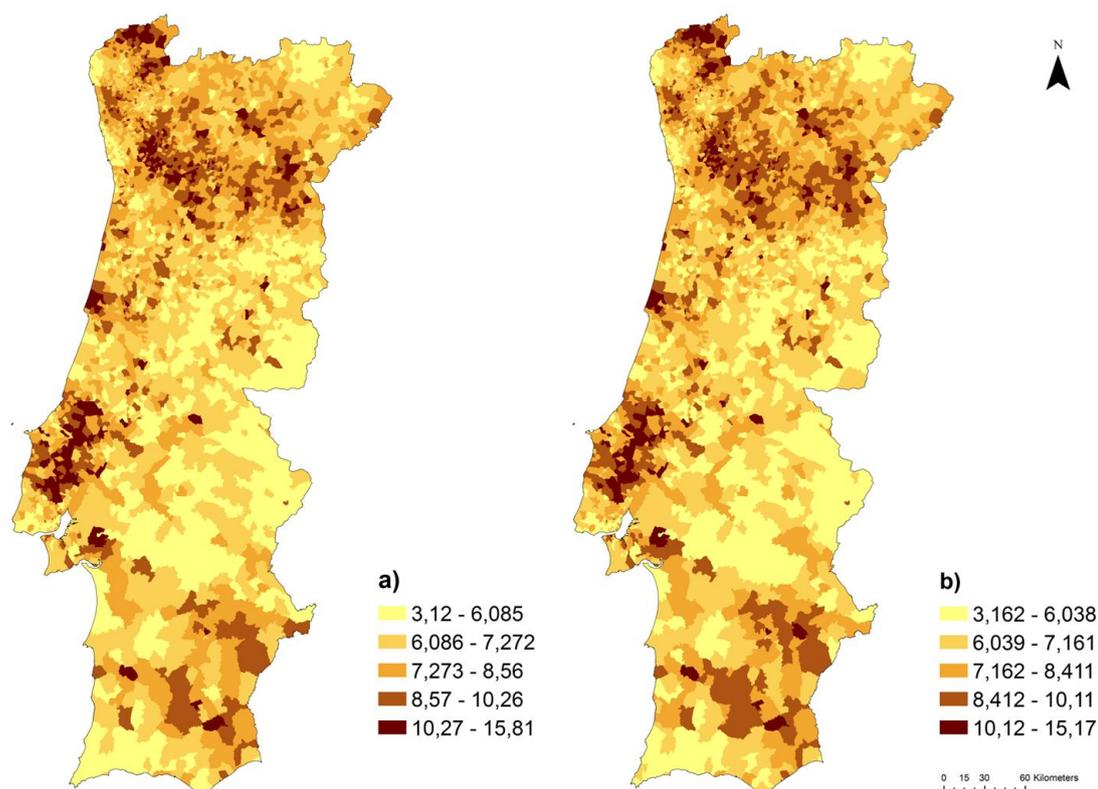


Figura 65 – Rácios padronizados de mortalidade por AVC em Homens com mais de 64 anos de idade (a) e Homens de todas as idades (b) em municípios de Portugal Continental, 1998-2004.

Na análise dos padrões dos SMRi para o sexo feminino não surgem grandes alterações em termos de estrutura espacial, sendo que a diferença essencial para os coortes masculinos, tanto no coorte das mulheres com mais de 64 anos (figura 66a) como nas mulheres de todas as idades (figura 66b), reside na concentração de valores mais elevados no norte do país se deslocar mais para a área noroeste e não tanto para o interior da região norte, tal como acontece no padrão da mortalidade masculina, para além de ser aparente que os valores altos se encontram menos concentrados do que nos coortes masculinos. Finalmente, nos dois coortes agrupando ambos os sexos (figuras 67a e 67b), é de notar a similaridade com a distribuição espacial do coorte feminino de pessoas com mais de 64 anos.

Passando agora aos determinantes relacionados com a qualidade da habitação avaliados neste estudo, a sua análise através de estatística descritiva evidencia muito mais diversidade do que o que se verifica nos indicadores do risco de mortalidade por AVC, observando-se acima de tudo importantes disparidades entre os seus valores máximos e

Condições precárias de habitação e mortalidade por acidente vascular cerebral ao nível da freguesia em Portugal Continental, na população com mais de 65 anos de idade

mínimos registados ao longo das freguesias de Portugal Continental no ano censitário de 2001.

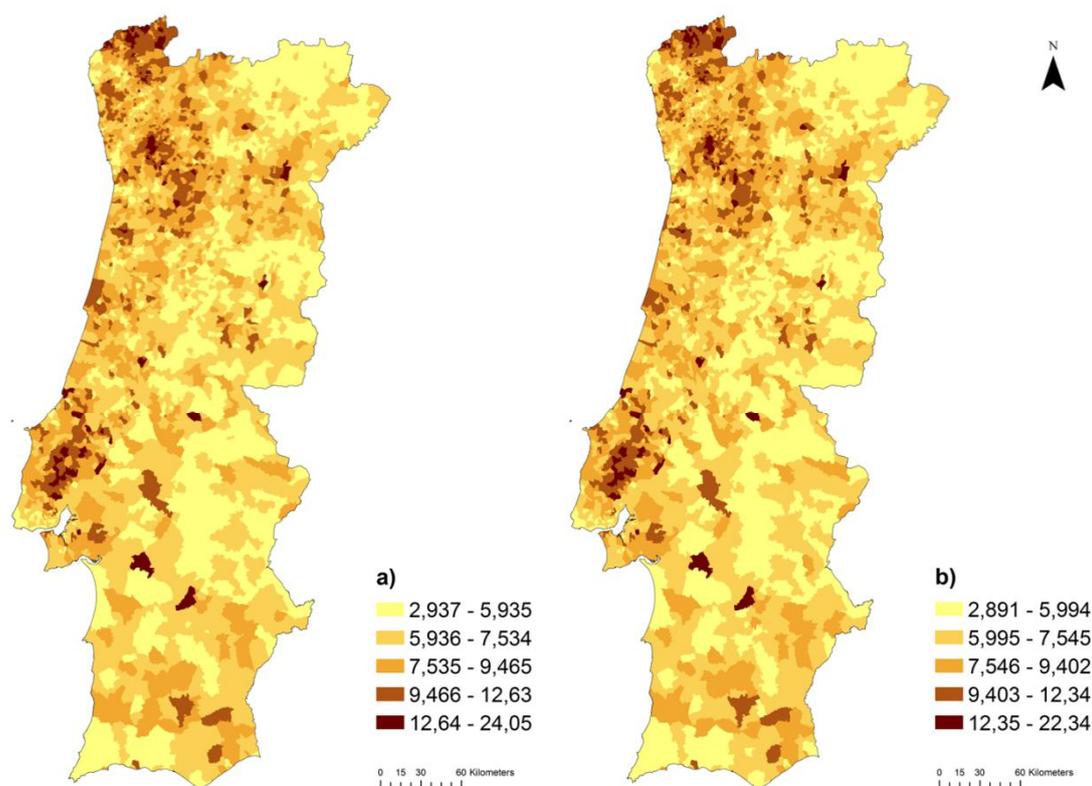


Figura 66 – Rácios padronizados de mortalidade por AVC em Mulheres com mais de 64 anos de idade (a) e Mulheres de todas as idades (b) em municípios de Portugal Continental, 1998-2004.

São notórios os casos em que uma grande percentagem (num dos casos mais de 90%) da população residente (em média 10% da população residente nas freguesias) declara não possuir qualquer meio de aquecimento da habitação, o facto de em 50% das freguesias mais de 95,7% dos alojamentos não possuírem aquecimento central, de chegarem a existir freguesias em que perto de 100% da população reside em edifícios construídos antes da década de 1960, e ainda em média, cerca de 15% dos alojamentos nas freguesias se encontrarem em situação de superlotação (tabela 9).

A análise da distribuição espacial mostra também grandes diferenças entre os padrões de distribuição espacial das variáveis, reforçando a ideia de grande heterogeneidade já transmitida pela estatística descritiva, sendo de notar a existência de desigualdades bem marcadas ao longo do território.

Condições precárias de habitação e mortalidade por acidente vascular cerebral ao nível da freguesia em Portugal Continental, na população com mais de 65 anos de idade

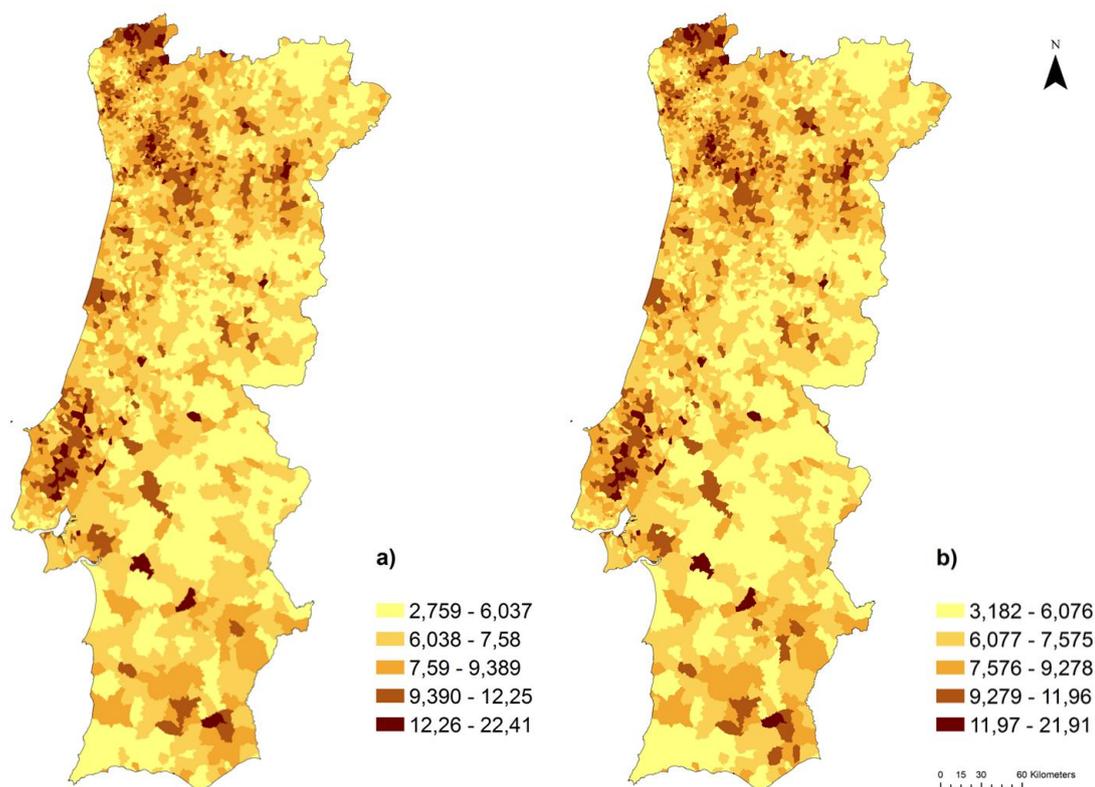


Figura 67 – Rácios padronizados de mortalidade por AVC, ambos os sexos, mais de 64 anos de idade (a) e ambos os sexos, todas as idades (b) em municípios de Portugal Continental, 1998-2004.

Tabela 9 – Estatísticas descritivas das variáveis explicativas

	Alojamentos superlotados (%)	População residente a viver em edifícios construídos até 1960 (%)	Alojamentos sem aquecimento central (%)	População Residente sem qualquer tipo de aquecimento (%)	Peso da população residente arrendatária face à população residente proprietária (%)	População residente a viver em apartamentos (%)	População Residente em alojamentos familiares não clássicos (%)
Mínimo	0	0	58,5	0	0	0	0
Máximo	42,65	99,64	100	90,35	983,54	99,75	26,4
Média	14,58	23,37	94,09	10,78	21,21	16,62	0,73
Mediana	13,52	20,69	95,70	4,36	9,96	6,97	0,1
Amplitude	42,65	99,64	41,50	90,35	983,54	99,75	26,40
Desvio-Padrão	6,96	13,3	5,71	13,83	41,07	22,16	1,74

A percentagem de residentes sem qualquer tipo de aquecimento na freguesia é um exemplo evidente, com duas manchas de valores mais elevados localizados em áreas distintas, nomeadamente numa área que abarca a área metropolitana do Porto, estendendo-se ao longo do litoral para norte até à foz do rio Minho e também para leste diminuindo para o interior, e noutra área localizada no extremo sul, coincidindo em grande parte com a NUT2 do Algarve (figura 68). Quanto à percentagem de alojamentos sem aquecimento central, o padrão é bem definido, com um notório diferencial entre as zonas a norte e a sul do rio Tejo, sendo que a maioria dos

Condições precárias de habitação e mortalidade por acidente vascular cerebral ao nível da freguesia em Portugal Continental, na população com mais de 65 anos de idade

alojamentos nas freguesias a sul do Tejo não possui este tipo de aquecimento, mas para norte do Tejo a situação revela muito menos homogeneidade, coexistindo tanto freguesias em que praticamente não existem alojamentos dotados de aquecimento central, com outras (embora pouco frequentes e geograficamente dispersas) em que mais de 40% dos alojamentos possuem este equipamento (figura 69). É de notar também que na região transmontana, onde se registam temperaturas mais baixas durante grande parte do ano com excepção do verão, existem muitas freguesias sem este tipo de equipamento.

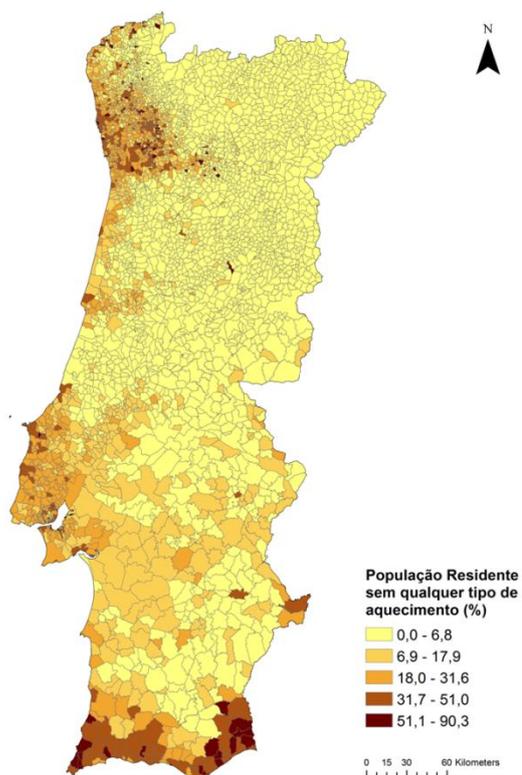


Figura 68

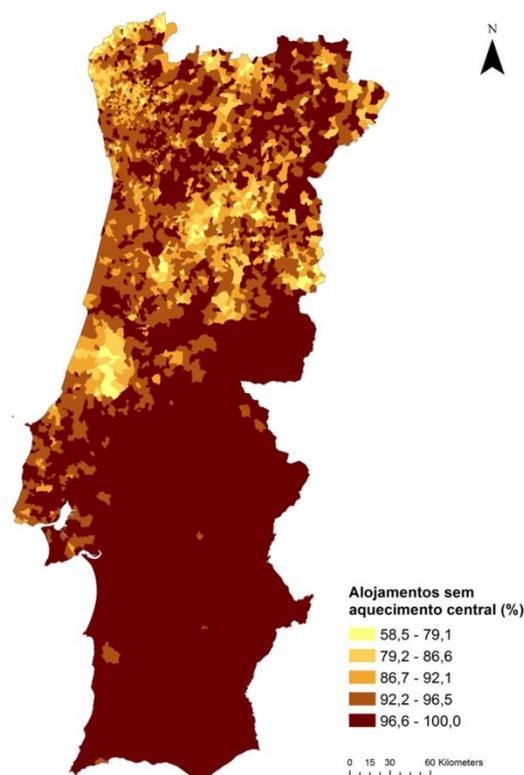


Figura 69

Quanto à distribuição espacial da percentagem de alojamentos superlotados por freguesia, esta apresenta, à semelhança de outras variáveis, uma mancha de valores elevados na área noroeste do território (figura 70), algo que também se verifica com a percentagem de população residente a viver em apartamentos (figura 71) e, de forma mais difusa, com o peso percentual dos residentes em alojamentos arrendados face aos residentes proprietários dos alojamentos (figura 72), sendo que nestas duas variáveis este padrão ocorre também noutras áreas do território continental, particularmente em outras áreas urbanas, tal como na área metropolitana de Lisboa e no Algarve, embora na

Condições precárias de habitação e mortalidade por acidente vascular cerebral ao nível da freguesia em Portugal Continental, na população com mais de 65 anos de idade

figura 72 seja também visível uma mancha de valores um pouco mais elevados que se estende de Lisboa para leste até à fronteira com Espanha. Quanto à percentagem de residentes a viver em edifícios construídos antes de 1960, é de notar um aumento progressivo dos valores do litoral para o interior, ocorrendo também valores elevados nas cidades de Lisboa e Porto (figura 73).

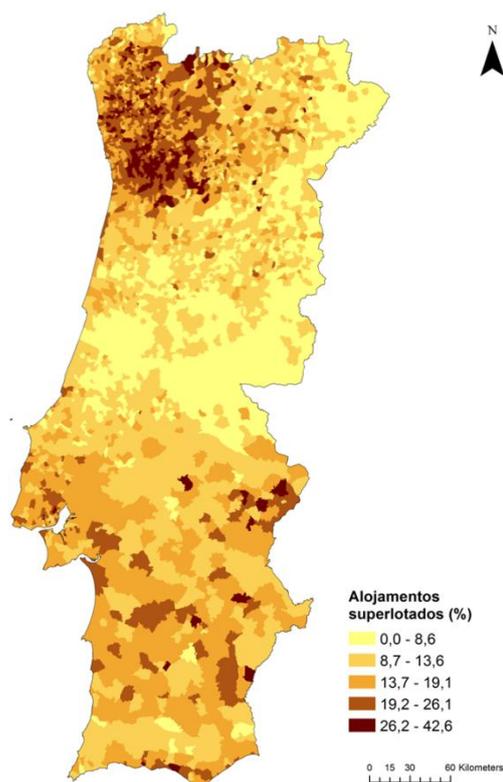


Figura 70

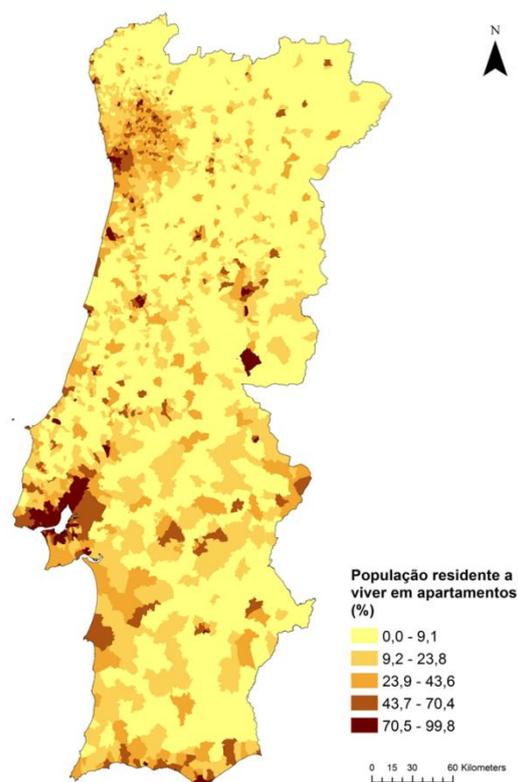


Figura 71

Finalmente, a percentagem de população residente a viver em alojamentos não clássicos aparenta ser a variável com uma distribuição espacial dos valores elevados menos concentrada, sendo que a maior parte das freguesias do território continental apresentam valores muito reduzidos (figura 74).

Condições precárias de habitação e mortalidade por acidente vascular cerebral ao nível da freguesia em Portugal Continental, na população com mais de 65 anos de idade

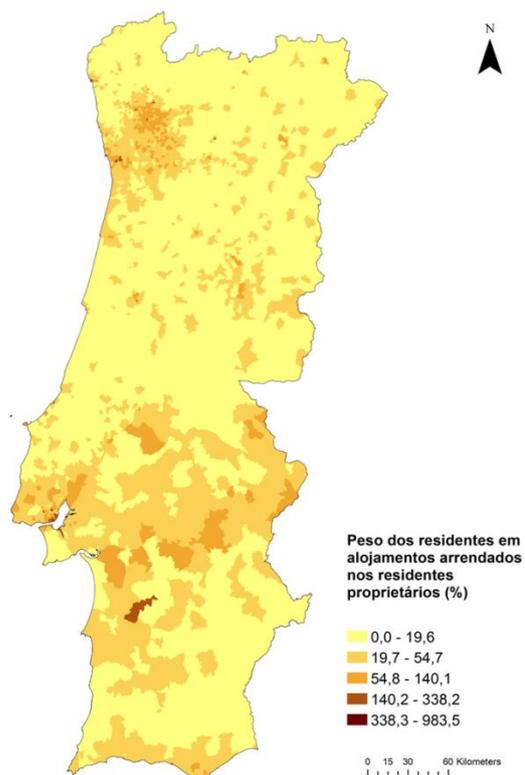


Figura 72

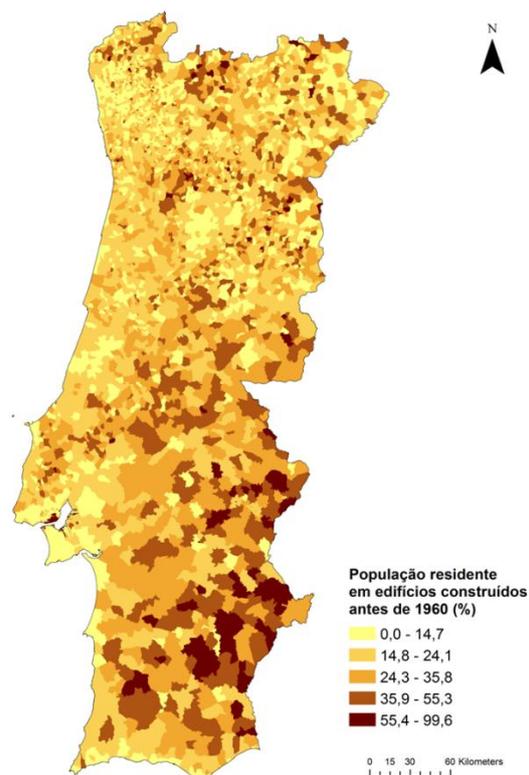


Figura 73

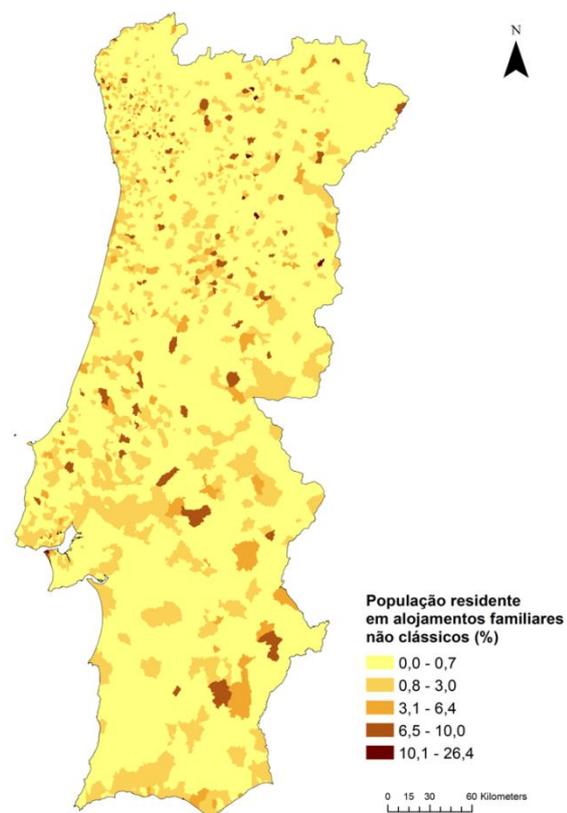


Figura 74

3.3.2 Análise univariada de clustering espacial do risco de AVC e determinantes relacionados com a qualidade da habitação

O cálculo do Índice I de Moran permitiu constatar que todas as variáveis em estudo apresentam tendência para o clustering espacial, tendo sido detectada a presença de autocorrelação espacial positiva na estrutura espacial de todas elas. Por outras palavras, em todas as variáveis existe uma tendência generalizada para valores similares ocorrerem em freguesias mais próximas entre si, sendo esta tendência reforçada confirmada pela obtenção de um nível alto de pseudo-significância ($p < 0,001$) em todos os testes efectuados (Anexo 7).

Após a avaliação da tendência para a ocorrência de clustering, a detecção da distribuição espacial dos próprios clusters, com recurso ao Índice I de Moran Local (LISA), veio revelar a presença dos padrões espaciais de clustering em cada uma das variáveis. Assim, no que respeita aos padrões de clustering espacial dos indicadores de risco de mortalidade, todos apresentam clusters espaciais relativamente bem definidos de agregações de freguesias com valores altos (clusters alto-alto) e de valores baixos (clusters baixo-baixo), como se observa nas figuras 75 a 80.

Assim, os padrões de clustering espacial dos vários SMR_i observados nas figuras 75 a 80 apresentam, como esperado, tendências similares aos padrões detectados visualmente na análise prévia através de mapas coropletos dos vários SMR_i, notando-se, tal como nesses mapas, a presença de um padrão geral relativamente comum, por exemplo no que diz respeito à ocorrência de valores elevados de excesso de risco de mortalidade no sector noroeste e norte/centro do território de Portugal Continental. Também no que respeita à distribuição dos clusters espaciais das variáveis explicativas (figuras 81 a 87), são igualmente detectados padrões de clustering espacial, e embora, consoante o determinante em causa, a dimensão e a quantidade de clusters variem de forma mais substancial do que nos indicadores de mortalidade, de um modo geral a distribuição espacial acompanha aquela detectada visualmente nos mapas coropletos das variáveis explicativas. Entre os vários exemplos, é de notar a confirmação de clusters do tipo Alto-Alto de freguesias com valores elevados da percentagem de população residente sem qualquer meio de aquecimento, especialmente na área noroeste do território continental (figura 81).

Condições precárias de habitação e mortalidade por acidente vascular cerebral ao nível da freguesia em Portugal Continental, na população com mais de 65 anos de idade

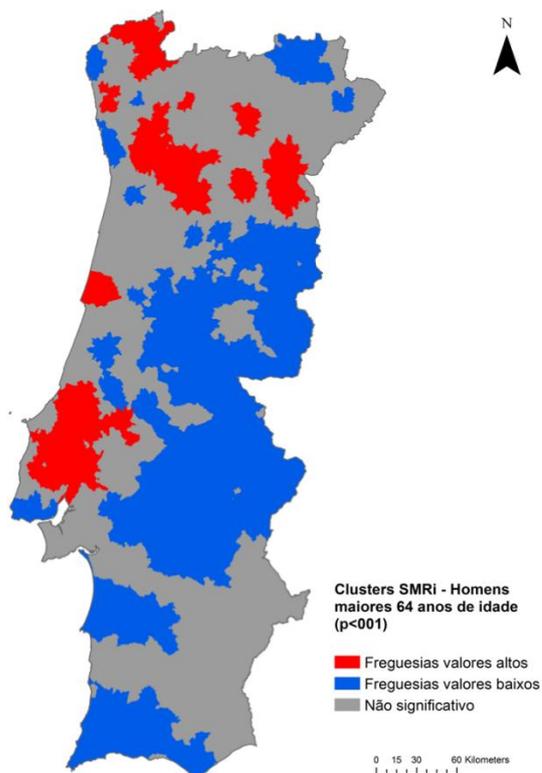


Figura 75

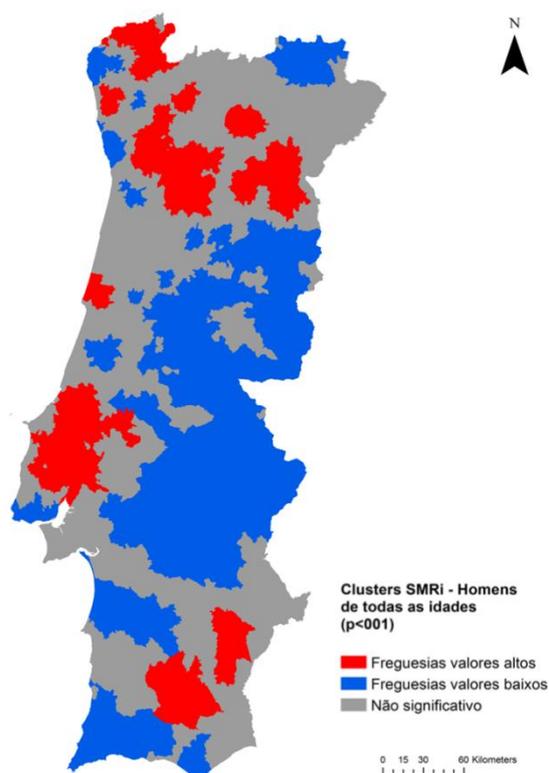


Figura 76

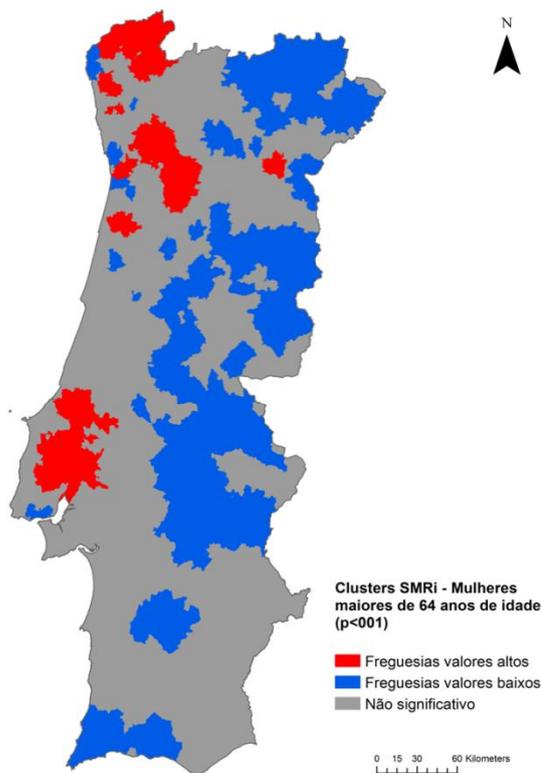


Figura 77

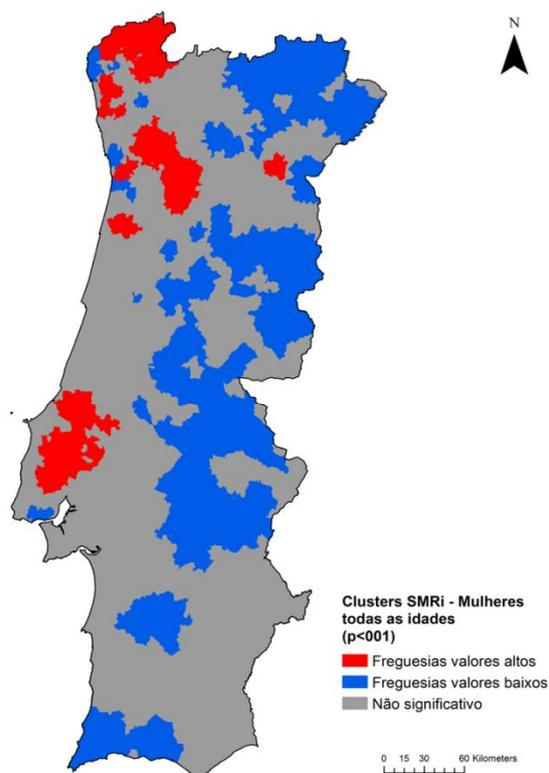


Figura 78

Condições precárias de habitação e mortalidade por acidente vascular cerebral ao nível da freguesia em Portugal Continental, na população com mais de 65 anos de idade

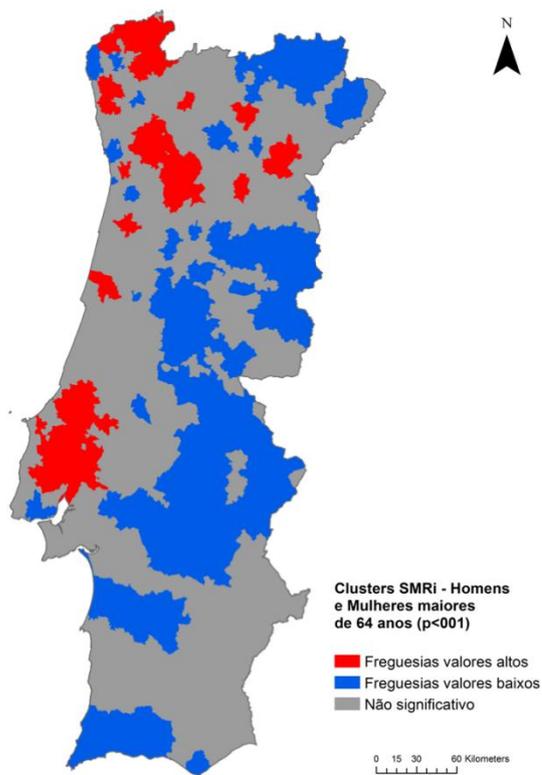


Figura 79

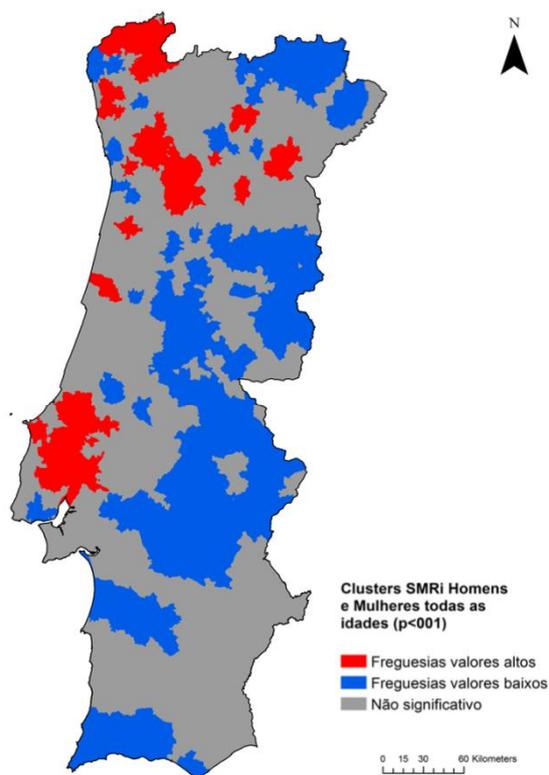


Figura 80

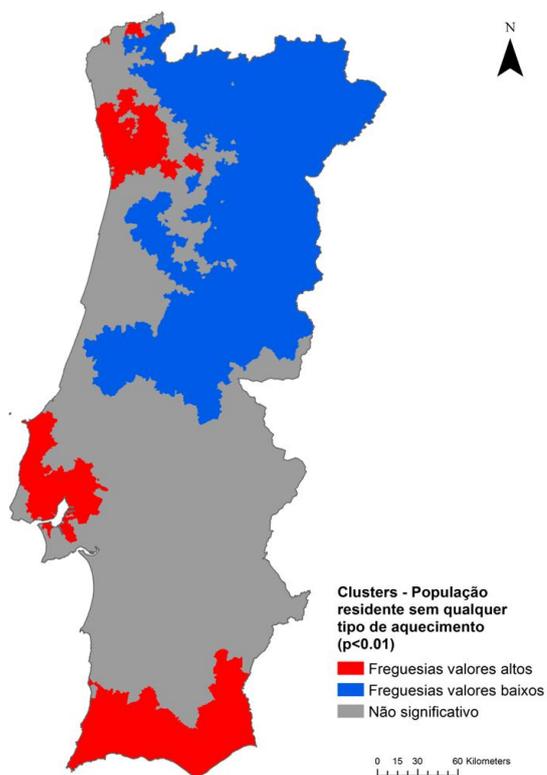


Figura 81

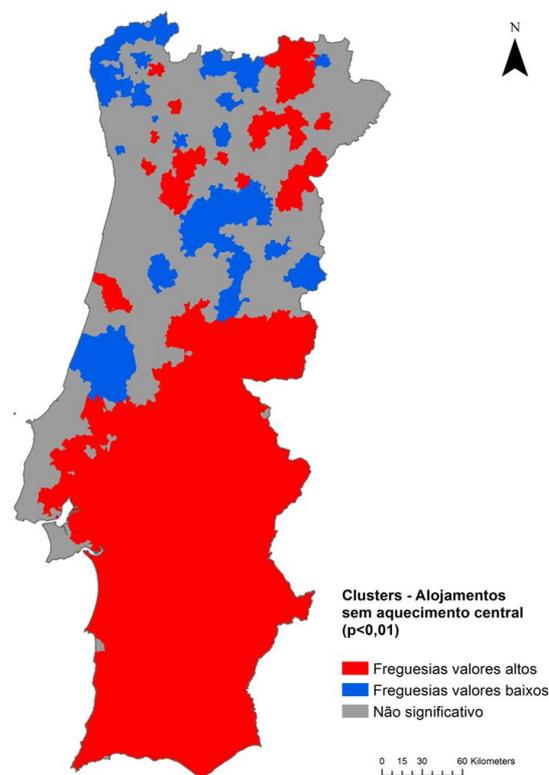


Figura 82

Condições precárias de habitação e mortalidade por acidente vascular cerebral ao nível da freguesia em Portugal Continental, na população com mais de 65 anos de idade

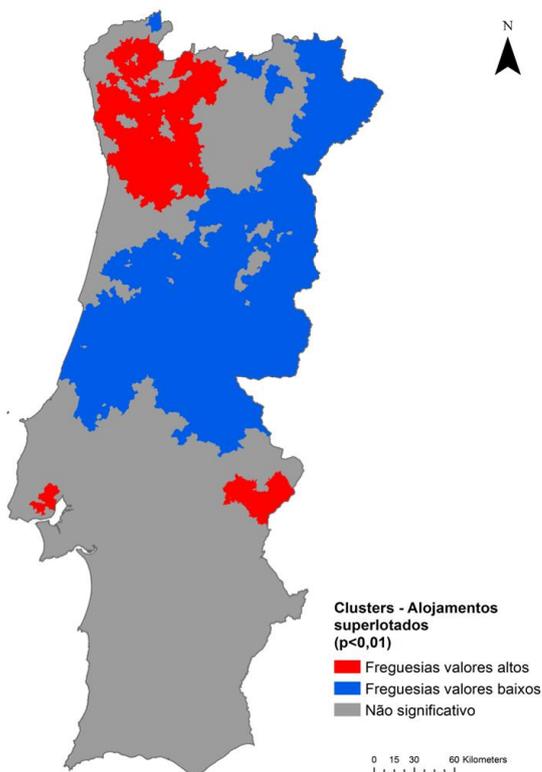


Figura 83

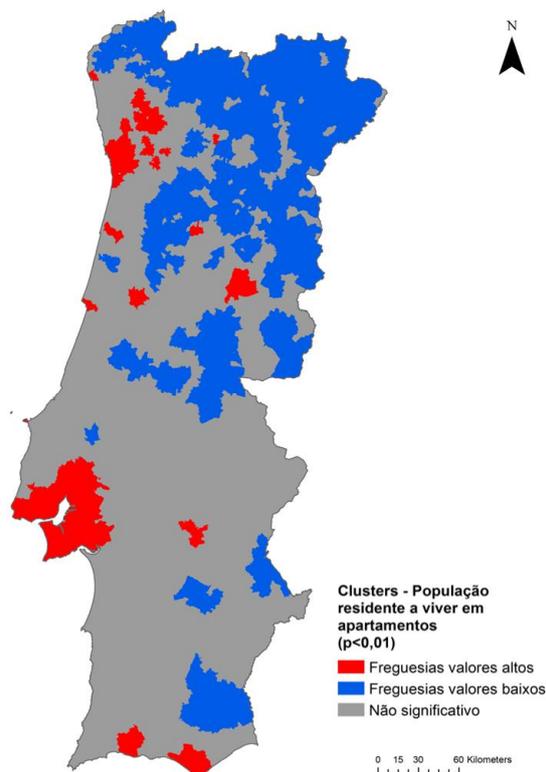


Figura 84

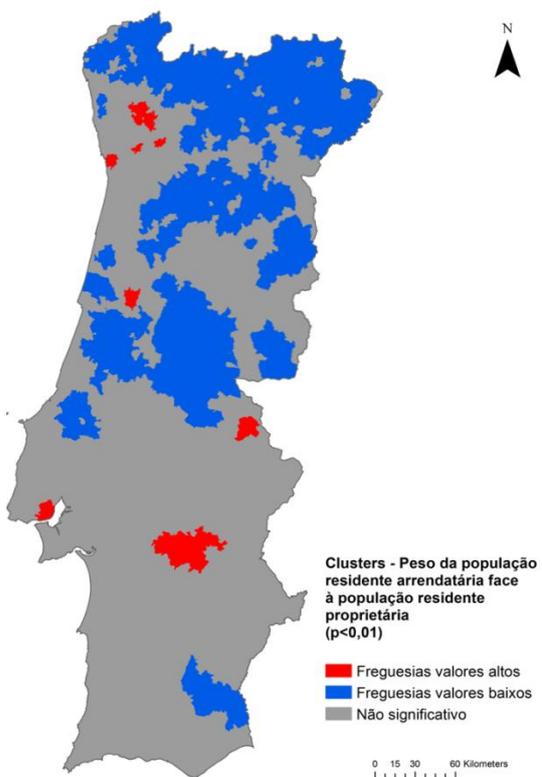


Figura 85

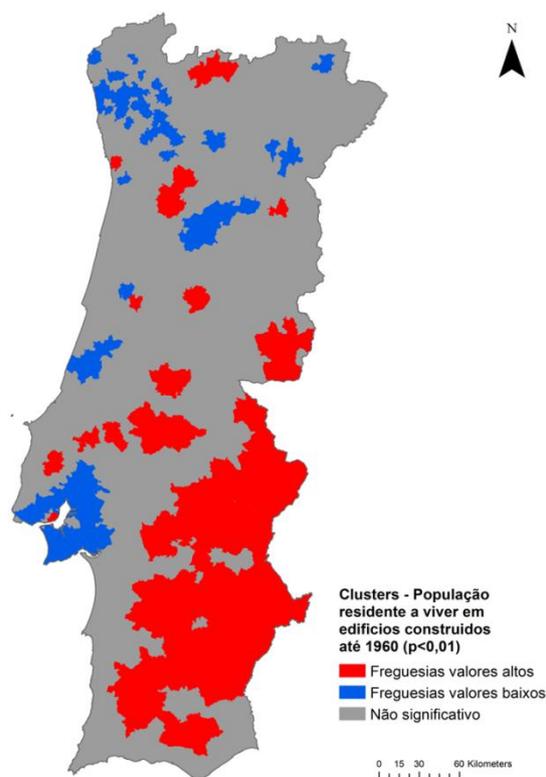


Figura 86

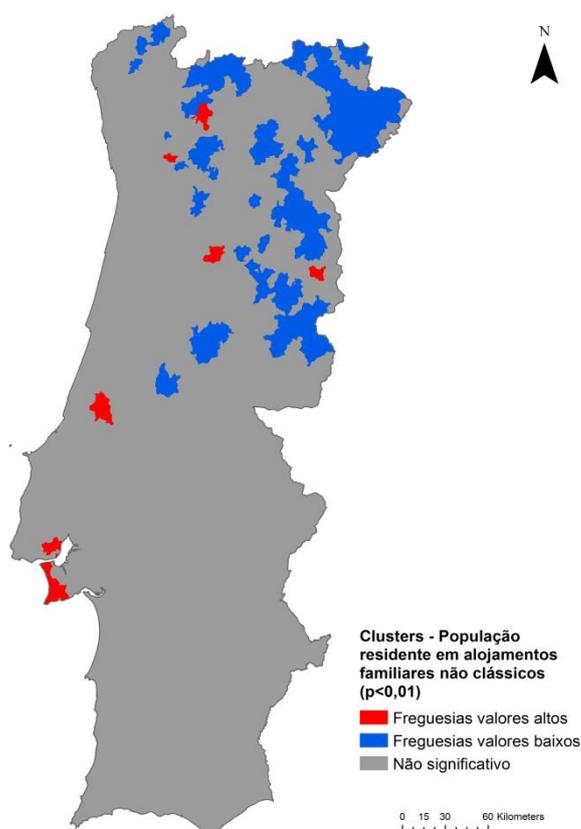


Figura 87

Após a obtenção da delimitação dos clusters espaciais Alto-Alto e Baixo-Baixo, verificou-se, ao efectuar o cálculo dos valores das médias e medianas dos vários SMRi nas freguesias incluídas dentro dos dois tipos de clusters para cada variável explicativa, que em algumas destas variáveis ocorre um padrão consistente, no sentido das médias e medianas das freguesias nos clusters alto-alto apresentarem valores mais elevados do que nos clusters baixo-baixo (tabela 10). Esta situação é particularmente importante na percentagem de população residente sem qualquer tipo

de aquecimento, na medida que é transversal a todos os coortes de mortalidade analisados. A mesma situação ocorre na relação com a percentagem de alojamentos superlotados. Finalmente, esta situação ocorre também de forma similar embora menos consistente na percentagem de população residente em apartamentos, na medida em que acontece apenas nos dois coortes femininos e no coorte que engloba todas as pessoas com mais de 64 anos, e ainda na percentagem de população residente em alojamentos não clássicos, em que acontece também apenas nos dois coortes femininos.

No que respeita à detecção de situações de multicolinearidade que pudessem determinar a exclusão de variáveis explicativas antes da transição para fase de análise multivariada, não foram detectados valores de correlação entre variáveis em que o coeficiente r excedesse o limiar previsto de 0,6 (Anexo 8), pelo que todas as variáveis explicativas em estudo transitaram para a próxima fase de análise.

Condições precárias de habitação e mortalidade por acidente vascular cerebral ao nível da freguesia em Portugal Continental, na população com mais de 65 anos de idade

Tabela 10 – Médias e medianas dos SMRi por tipo de cluster de variável explicativa

Indicadores de risco mortalidade AVC (SMRi) por coorte e tipo cluster	Estatística	Tipo de cluster	Homens todas as idades	Homens mais de 64 anos	Mulheres todas as idades	Mulheres mais de 64 anos	Homens e Mulheres todas as idades	Homens e Mulheres mais de 64 anos
Alojamentos superlotados (%)	Média	Alto-Alto	8,13	8,26	8,55	8,60	8,40	8,50
		Baixo-Baixo	6,86	6,92	6,53	6,52	6,66	6,67
	Mediana	Alto-Alto	8,04	8,10	8,28	8,27	8,20	8,27
		Baixo-Baixo	6,68	6,74	6,30	6,28	6,45	6,45
População residente a viver em edifícios construídos até 1960 (%)	Média	Alto-Alto	6,96	6,90	6,95	6,92	6,95	6,91
		Baixo-Baixo	7,57	7,67	7,71	7,76	7,67	7,74
	Mediana	Alto-Alto	6,83	6,72	6,50	6,47	6,64	6,61
		Baixo-Baixo	7,37	7,49	7,50	7,50	7,36	7,43
Alojamentos sem aquecimento central (%)	Média	Alto-Alto	7,24	7,24	7,06	7,05	7,15	7,14
		Baixo-Baixo	7,35	7,45	7,50	7,53	7,43	7,48
	Mediana	Alto-Alto	6,92	6,83	6,56	6,55	6,69	6,66
		Baixo-Baixo	7,13	7,21	7,02	7,05	7,00	7,05
População residente sem qualquer tipo de aquecimento (%)	Média	Alto-Alto	7,77	7,89	8,18	8,25	8,05	8,16
		Baixo-Baixo	7,22	7,24	6,78	6,76	6,96	6,96
	Mediana	Alto-Alto	7,56	7,64	7,85	7,87	7,62	7,72
		Baixo-Baixo	7,06	7,06	6,50	6,48	6,65	6,63
Peso da população residente arrendatária face à população residente proprietária (%)	Média	Alto-Alto	6,96	7,00	7,05	7,08	7,02	7,06
		Baixo-Baixo	7,58	7,63	7,34	7,34	7,48	7,49
	Mediana	Alto-Alto	6,87	6,83	6,65	6,67	6,81	6,75
		Baixo-Baixo	7,38	7,42	6,91	6,87	7,07	7,05
População residente a viver em apartamentos (%)	Média	Alto-Alto	7,16	7,26	7,41	7,46	7,32	7,41
		Baixo-Baixo	7,54	7,56	7,22	7,21	7,38	7,38
	Mediana	Alto-Alto	6,89	6,92	7,03	7,06	6,91	6,98
		Baixo-Baixo	7,38	7,38	6,84	6,79	7,00	6,98
População Residente em alojamentos familiares não clássicos (%)	Média	Alto-Alto	7,05	7,01	7,20	7,20	7,13	7,13
		Baixo-Baixo	7,55	7,57	7,16	7,15	7,37	7,36
	Mediana	Alto-Alto	6,84	6,80	6,88	6,74	6,84	6,97
		Baixo-Baixo	7,31	7,29	6,78	6,74	7,06	6,98

3.3.3 Análise de clustering multivariada

Após o teste das combinações previstas do critério de agregação de Ward com as medidas de distância (ou similaridade), os resultados mais equilibrados foram obtidos na utilização do critério de agregação de Ward com a distância de Gower, a qual aparenta conjugar melhor os vários critérios empregues na escolha do padrão de clustering espacial multivariado mais apropriado. Assim, a análise do dendrograma resultante permitiu constatar visualmente a presença de uma estrutura de clustering equilibrada, no sentido das freguesias se encontrarem distribuídas de forma relativamente uniforme pelos clusters. Sugere também de forma intuitiva a presença de três clusters relativamente bem destacados (e distanciados) em relação a outros padrões de clustering mais fragmentados, sugerindo assim o corte do dendrograma ao nível dos

Condições precárias de habitação e mortalidade por acidente vascular cerebral ao nível da freguesia em Portugal Continental, na população com mais de 65 anos de idade

três clusters. A figura 88 representa o dendrograma resultante com a delimitação dos três clusters sugeridos (a vermelho) em sobreposição ao diagrama.

Também o cálculo do coeficiente de correlação cofenética para todas as combinações testadas confirma a preferência por esta combinação, na medida em que esta é a escolha que apresenta o valor de correlação mais elevado (tabela 11).

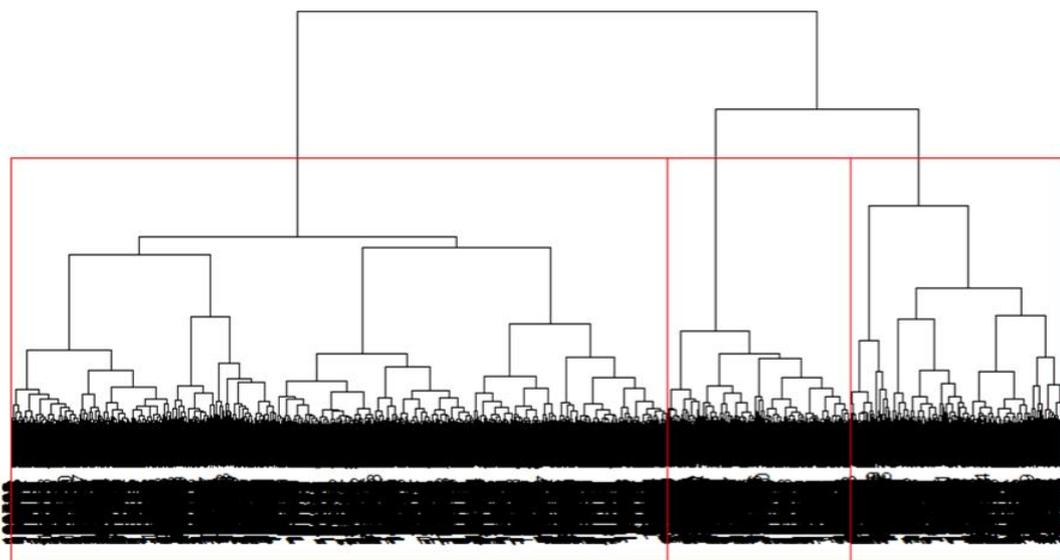


Figura 88 – Dendrograma final

Tabela 11 – Coeficientes de correlação cofenética por teste

Critério de agregação	Tipo de distância	Coeficiente de correlação cofenética (R_c)
Ward	Euclidiana	0,4
	Minkowski	0,4
	Máxima	0,4
	Manhattan	0,46
	Gower	0,53

Quanto à análise do número de clusters através da comparação gráfica do número de clusters com o coeficiente de fusão, esta parece confirmar a sugestão inicial de “corte” do dendrograma ao nível de três clusters (figura 89). A contagem do número de freguesias abrangidas por cada cluster apresentou-se, como esperado, relativamente equilibrada, contendo os clusters 1, 2 e 3 um total de 2511, 825 e 701 freguesias, respectivamente.

Condições precárias de habitação e mortalidade por acidente vascular cerebral ao nível da freguesia em Portugal Continental, na população com mais de 65 anos de idade

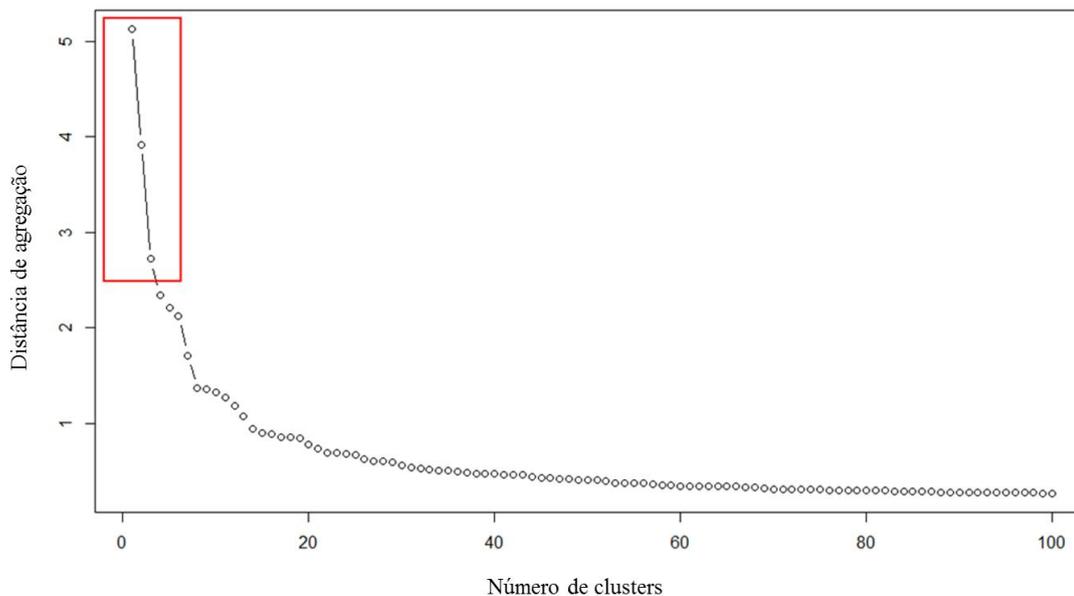


Figura 89 – Coeficientes de fusão, número de clusters e clusters a adoptar

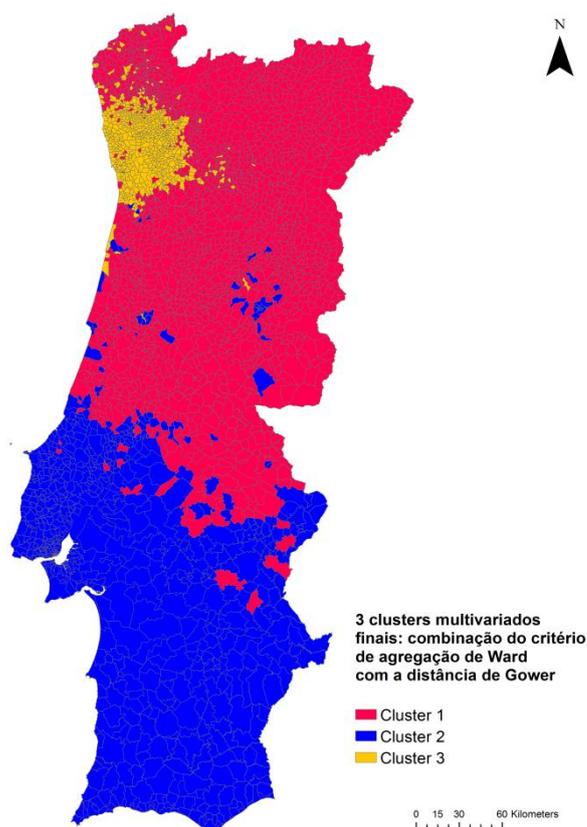


Figura 90

A distribuição espacial dos três clusters finais resultantes da aplicação do método hierárquico combinando o critério de Ward com a distância de Gower podem ser observados na figura 90, sendo visível a presença de um cluster bem definido (cluster 3) na área norte do país, o qual incide em grande parte na área metropolitana do Porto, mas também que o cluster 1 engloba grande parte do restante território norte e centro/norte do país. Quanto ao cluster 2, este abarca a maior parte do sul do território continental.

No que respeita aos testes de significância estatística referentes à

diferenciação entre as distribuições dos valores de todas as variáveis em estudo pelos 3 clusters, esta diferenciação revelou-se estatisticamente significativa de forma

Condições precárias de habitação e mortalidade por acidente vascular cerebral ao nível da freguesia em Portugal Continental, na população com mais de 65 anos de idade

consistente para todas as variáveis explicativas e SMRi dos coortes femininos e todas as idades ($p < 0,0001$), tendo apenas o grau de significância descido um pouco nos coortes homens de todas as idades ($p < 0,01$) e homens com menos de 64 anos ($p < 0,05$), embora não perdendo significância estatística (assumindo-se aqui $p < 0,05$ como o mínimo aceitável).

O cálculo da média das variáveis explicativas estandardizadas dentro dos clusters permitiu determinar qual a variável dominante em cada um dos 3 clusters (Tabela 12), estabelecendo-se assim como dominantes nos clusters 3, 2 e 1 as variáveis: população residente sem qualquer tipo de aquecimento (%), população residente a viver em apartamentos (%) e população residente a viver em edifícios construídos até 1960 (%), respectivamente.

Tabela 12 – Médias dos valores padronizados das variáveis explicativas nos 3 clusters finais

Variáveis causais	Clusters		
	1	2	3
Alojamentos superlotados (%)	-0,299	0,042	1,021
População residente a viver em edifícios construídos até 1960 (%)	-0,026	0,419	-0,398
Alojamentos sem aquecimento central (%)	-0,196	0,620	-0,027
População residente sem qualquer tipo de aquecimento (%)	-0,511	0,535	1,201
Peso da população residente arrendatária face à população residente proprietária (%)	-0,305	0,506	0,499
População residente a viver em apartamentos (%)	-0,440	0,741	0,704
População Residente em alojamentos familiares não clássicos (%)	-0,058	0,137	0,045

Na última fase da análise multivariada, o cálculo das médias e medianas dos SMRi dos vários coortes em estudo nas freguesias dentro dos 3 clusters determinados permitiu concluir que o cluster 3 é aquele que apresenta, em média, valores mais altos de excesso de risco de morte por AVC, seguido pelo cluster 1 e por último pelo cluster 2, sendo este o que apresenta, em média, valores de excesso de risco mais reduzidos (Tabela 13).

Condições precárias de habitação e mortalidade por acidente vascular cerebral ao nível da freguesia em Portugal Continental, na população com mais de 65 anos de idade

Tabela 13 – Valores SMRi médios e medianos por cluster multivariado

Cluster	Estatística	Homens todas as idades	Homens mais de 64 anos	Mulheres todas as idades	Mulheres mais de 64 anos	Homens e Mulheres todas as idades	Homens e Mulheres mais de 64 anos
Cluster 3	Média	8,04	8,22	8,48	8,57	8,48	8,48
	Mediana	7,93	8,01	8,15	8,24	8,15	8,17
Cluster 1	Média	7,35	7,40	7,18	7,19	7,18	7,27
	Mediana	7,15	7,21	6,87	6,83	6,87	6,90
Cluster 2	Média	7,17	7,15	7,16	7,16	7,16	7,17
	Mediana	6,88	6,79	6,73	6,77	6,73	6,76

Estas tendências aparentam ser consistentes, na medida em que sucedem em todos os coortes em estudo. Parece assim verificar-se um padrão espacial na distribuição do risco de morte por AVC, na medida em que o excesso de risco parece diminuir progressivamente de norte para sul ao longo do território nacional continental.

3.4 Discussão

Embora os métodos de clustering hierárquico ou espaciais utilizados neste estudo existam já há algum tempo, a metodologia aqui ensaiada, parcialmente inspirada em Schieb et al. (2013) nunca foi, tanto quanto sabemos, testada em território Português. Em parte, a sua novidade reside em procurar clusters espaciais com base na análise multivariada das variáveis explicativas e verificar depois o comportamento dos indicadores de risco de AVC dentro desses mesmos clusters, e não o inverso, como em Schieb et al. (2013). Por outro lado, a análise da relação entre excesso de mortalidade por AVC e os determinantes relacionados com a qualidade da habitação não foi até agora, tanto quanto sabemos, ensaiada em Portugal empregando uma unidade estatística ao nível de desagregação territorial da freguesia.

A metodologia de análise aqui ensaiada constitui uma abordagem essencialmente exploratória e permitiu, como tal, detectar várias tendências na relação entre o excesso de mortalidade por AVC e os indicadores da qualidade de habitação nos vários coortes populacionais em estudo, e formular, neste contexto, várias hipóteses.

Do ponto de vista dos vários coortes populacionais em estudo, os resultados da comparação das estatísticas descritivas entre os vários Rácios Padronizados de

Condições precárias de habitação e mortalidade por acidente vascular cerebral ao nível da freguesia em Portugal Continental, na população com mais de 65 anos de idade

Mortalidade (SMR_i) na análise conjunta de todas as freguesias do território continental sugerem (particularmente pela análise dos valores máximos) que as duas influências dominantes no excesso de mortalidade parecem ser a mortalidade ocorrida na população e em pessoas com mais de 64 anos, o que não é surpreendente, tendo em conta a própria distribuição dos valores absolutos de mortalidade por AVC durante o período de estudo entre coortes acompanha estas tendências.

Por outro lado, quando se observam os valores das médias e medianas entre os coortes populacionais, as diferenças surgem mais esbatidas, pois o excesso médio de mortalidade varia entre 7,40 em mulheres de todas as idades e 7,49 em homens com mais de 64 anos e o excesso mediano varia entre 7,02 em mulheres com mais de 64 anos e 7,27 em homens com mais de 64 anos de idade. Também a visualização da distribuição espacial dos vários SMR_i em todas as freguesias através dos respectivos mapas coropletos parece apontar mais para a similaridade do que para a diferença entre coortes, parecendo emergir um padrão comum de valores mais elevados de excesso de mortalidade na área norte/noroeste do território (incluindo grande parte da área metropolitana do Porto) e na área imediatamente no norte do distrito de Lisboa e distrito de Leiria, surgindo por vezes algumas manchas menos concentradas em vários pontos do território, notavelmente no interior sudeste.

No que respeita às variáveis explicativas, algumas parecem apresentar, nos mapas coropletos, um padrão de distribuição espacial similar ao apresentado pelos indicadores do excesso de mortalidade, o que permite colocar a hipótese de existir um nexo de causalidade entre estas variáveis e os padrões de excesso de mortalidade. Desta forma, na percentagem de população residente sem qualquer tipo de aquecimento observa-se também uma concentração de valores mais elevados na área norte/noroeste do território (incluindo a área metropolitana do Porto). Adicionalmente, o facto dos indicadores do excesso de mortalidade não apresentarem o mesmo padrão de valores elevados na região do Algarve que se nota nesta variável levanta desde logo a hipótese da presença de meios de aquecimento poder ter um impacto diferencial entre os extremos norte e sul do território, possivelmente por questões ligadas aos diferentes regimes climáticos destas duas regiões, na medida em que no extremo norte do território continental as temperaturas médias anuais são mais baixas do que na região do Algarve. É de notar que o mesmo padrão de concentração na área norte (mas não no Algarve) ocorre

Condições precárias de habitação e mortalidade por acidente vascular cerebral ao nível da freguesia em Portugal Continental, na população com mais de 65 anos de idade

igualmente na percentagem de alojamentos superlotados na freguesia. Tendo em conta que residir num alojamento superlotado constitui um indicador de um estatuto socioeconómico mais reduzido, poderá acontecer que a similaridade destes padrões espaciais tenha a ver com a hipótese das pessoas que vivem nestes alojamentos sejam justamente aquelas que (devido à sua capacidade económica mais reduzida) possuem mais dificuldade em dispor de meios de aquecer devidamente a sua habitação.

No que diz respeito à percentagem de alojamentos sem aquecimento central por freguesia, esta variável possui um padrão espacial diferente da população sem qualquer meio de aquecimento, com uma concentração de valores altos na maior parte do sul do país e um padrão menos bem definido por todo o norte. Todavia, tendo em conta que este meio de aquecimento não é ainda uma opção muito frequente nos alojamentos em Portugal (por ser em geral considerado como tendo custos mais elevados de instalação e utilização), e que a sua ausência não exclui a posse de outro qualquer meio de aquecimento da habitação, a contribuição desta variável será provavelmente menos importante. Nas restantes variáveis, é ainda de realçar que tanto a percentagem de população residente em apartamentos, como o peso percentual da população residente em alojamentos arrendados, apresentam também uma concentração de valores mais elevados na área norte do país (em particular na Área Metropolitana do Porto), embora este padrão se note também na Área Metropolitana de Lisboa, aparentando ser mais frequente em áreas de grande concentração populacional e não apenas na Área Metropolitana do Porto.

Para além da análise visual de partida, focada na coincidência de padrões espaciais de clustering, o facto das médias e medianas dos vários SMR_i nas freguesias incluídas dentro dos clusters espaciais Alto-Alto e Baixo-Baixo apresentarem um padrão consistente em algumas das variáveis, no sentido das médias e medianas das freguesias nos clusters alto-alto se apresentarem mais altas do que nos clusters baixo-baixo, vem reforçar a hipótese da existência de associações relevantes de sentido positivo entre o aumento do risco de mortalidade por AVC e algumas das variáveis explicativas. Esta situação é especialmente relevante na percentagem de população residente sem qualquer tipo de aquecimento e na percentagem de alojamentos superlotados, por ser transversal a todos os cortes em estudo, sugerindo, com carácter univariado, que valores elevados em ambas as variáveis possam estar associados ao aumento do risco de morte por AVC.

Condições precárias de habitação e mortalidade por acidente vascular cerebral ao nível da freguesia em Portugal Continental, na população com mais de 65 anos de idade

Esta hipótese é coerente com o estado do conhecimento acerca da relação entre as baixas temperaturas e o aumento de risco de morte por AVC e também acerca da relação entre a superlotação dos alojamentos e o estatuto socioeconómico

O facto de esta situação ocorrer com a percentagem de população residente a viver em apartamentos, mas apenas nos dois coortes femininos e no coorte de todas as pessoas com mais de 64 anos, levanta uma nova hipótese, visto que o estado do conhecimento refere-se até agora ao aumento do risco de depressão em mulheres, sendo algo inesperada a hipótese de existir uma associação positiva com o aumento do risco de AVC.

Quanto aos valores médios de excesso de risco mais elevado nos clusters alto-alto da percentagem de população residente em alojamentos não clássicos, que ocorrem também apenas nos dois coortes femininos, especula-se que esta situação poderá dever-se apenas ao facto da residência nestes alojamentos ser, em geral, um indicador de baixos níveis de estatuto socioeconómico, o que, aliado à maior longevidade geral do sexo feminino, poderá ter alguma influência nesta situação. Todavia, esta situação poderá também ser de carácter espúrio, visto esta ser a variável explicativa que apresenta um padrão mais escasso de clustering espacial, algo já previamente notado pela observação da sua distribuição espacial.

Após as primeiras hipóteses, decorrentes da análise univariada, impunha-se dar o passo seguinte e avaliar a resposta conjunta das várias variáveis explicativas na formação de clusters de freguesias. Nesse sentido, os métodos de clustering multivariado ensaiados aparentam ser adequados ao caso de estudo, especialmente no que diz respeito à utilização do critério de agregação de Ward, que, como esperado, permitiu obter clusters de configuração mais compacta do que seria esperado com outros critérios. Todavia, é importante realçar que essa configuração poderia ocorrer apenas num espaço multidimensional abstracto definido pelos eixos que representam as várias variáveis, ou por outra palavras, não apresentar ainda assim padrões espaciais coerentes. Para tal, a opção de incluir na análise as coordenadas X e Y dos centros geométricos das freguesias permitiu assim adaptar com sucesso a técnica de clustering aglomerativo hierárquico, transformando-a numa metodologia de clustering “semi-espacial” e permitindo obter padrões coerentes de clustering espacial multivariado.

Condições precárias de habitação e mortalidade por acidente vascular cerebral ao nível da freguesia em Portugal Continental, na população com mais de 65 anos de idade

Quanto ao padrão de clustering espacial multivariado obtido, é de realçar a importância do cluster 3, que permanece coerente com os resultados já obtidos previamente na análise univariada, na medida em que a variável dominante neste cluster é a percentagem de população residente sem qualquer meio de aquecimento, a qual predomina, mesmo quando colocada “em competição” com todas as outras variáveis explicativas, algumas não apresentando padrões espaciais que evidenciem uma relação coerente com o excesso de risco de morte por AVC. É igualmente de realçar que a segunda variável com mais peso em média neste cluster é a percentagem de alojamentos superlotados. Assim, tendo em conta que as freguesias pertencentes ao cluster 3 apresentam de forma consistente valores de excesso de risco mais altos em todos os coortes em estudo, é lícito reforçar a hipótese já colocada (e coerente com os resultados prévios da análise univariada) de que o factor determinante no excesso de mortalidade na área noroeste do território continental seja a incapacidade de aquecer adequadamente a habitação. O facto da percentagem de alojamentos superlotados ser aqui o segundo factor dominante vem apenas reforçar esta hipótese, se assumirmos que a incapacidade de aquecer adequadamente a habitação decorre acima de tudo de uma situação de baixo estatuto socioeconómico, por outras palavras da incapacidade financeira dos habitantes. Quanto ao cluster 1, o facto da variável dominante neste ser a percentagem de edifícios construídos antes da década de 1960 levanta aqui a hipótese desta situação se relacionar (embora de forma menos intensa do que não possuir nenhum meio de aquecimento) com uma menor capacidade de manter a habitação devidamente aquecida, assumindo que as habitações situadas nestes edifícios possuem (em geral) menor eficiência energética em termos de isolamento térmico, do que aquelas baseadas em edifícios de construção mais recente.

Também a diminuição progressiva dos valores médios e medianos do excesso de risco de morte por AVC (em todos os coortes analisados) do cluster 3 para o cluster 2, ou por outras palavras de norte para sul, vem reforçar a hipótese da incapacidade de aquecer devidamente a habitação poder ter mais impacto no aumento do excesso de risco na área norte do país, sendo esta condição ampliada pelos padrões climáticos do território continental Português, na medida em que a área mais a norte apresenta, em média, temperaturas mais baixas durante esta época.

Condições precárias de habitação e mortalidade por acidente vascular cerebral ao nível da freguesia em Portugal Continental, na população com mais de 65 anos de idade

Retomando parte do início desta discussão, tendo em conta que a influência dominante na mortalidade geral por AVC parece ser exercida pelos coortes de pessoas com mais de 64 anos, bem como por pessoas do sexo feminino, coloca-se aqui também a hipótese da incapacidade de aquecer devidamente a habitação ter especial impacto na área dos clusters 3 e 1, dado que as pessoas pertencentes a estes dois coortes terão, em princípio, por força das certas circunstâncias (como por exemplo a contingência das pessoas de mais de 64 anos se encontrarem em boa parte em situação de reforma), tendência para passarem mais tempo na habitação, aumentando assim o seu grau de exposição a este tipo de situação adversa.

Existem, obviamente, algumas limitações inerentes a este estudo. Uma parte destas prende-se com a natureza exploratória não apenas da metodologia em geral, mas das próprias técnicas empregues. Todavia, o facto de algumas destas técnicas incorporarem já a indicação do grau de significância estatística associado aos resultados (mesmo que, por exemplo, no caso dos métodos LISA se esteja, em rigor, a falar de pseudo-significância) e de terem sido efectuados testes de significância relativos ao grau de diferenciação das (sub)distribuições das variáveis dentro dos clusters, permitiu, na medida do possível, aumentar o grau de confiança nos resultados obtidos. É também óbvio que, dado estarmos em presença de um estudo ecológico, os seus resultados abrirão sempre portas a estudos com metodologias mais específicas, que permitam alterar o enfoque na população como um todo e eventualmente passar ao nível individual. Não obstante, os estudos ecológicos permitem muitas vezes estreitar o foco de investigação, introduzindo um primeiro “filtro” (por exemplo ao nível dos coortes ou do âmbito territorial a abordar nos estudos seguintes) que se pode traduzir numa racionalização dos recursos de Saúde Pública empregues.

3.5 Conclusão

Do ponto de vista dos objectivos previstos, é possível afirmar que estes foram atingidos, na medida em que foram identificadas e investigadas um conjunto de variáveis representativas das más condições de habitabilidade, tendo-se determinado, dentro das limitações já debatidas, a sua contribuição relativa enquanto determinantes ambientais

Condições precárias de habitação e mortalidade por acidente vascular cerebral ao nível da freguesia em Portugal Continental, na população com mais de 65 anos de idade

tanto em sentido lato (como a residência em apartamentos, cujo impacto na saúde humana é maioritariamente psicológico) como em sentido restrito (como a incapacidade de aquecer devidamente a habitação no Inverno) do AVC em vários coortes populacionais ao nível da freguesia, em Portugal Continental. Também a análise conjunta da distribuição espacial destes determinantes e da mortalidade por AVC nos coortes em foco permitiu com sucesso obter uma primeira avaliação das possíveis associações entre estes determinantes e o risco de morte por AVC ao longo do território continental Português. Todavia, foi porventura a busca de clusters espaciais através de uma abordagem multivariada que permitiu dar um passo para lá da interpretação isolada dos determinantes. Os resultados obtidos parecem, assim, reforçar a ideia, sustentada por alguns autores, de que, no que respeita à contribuição conjunta dos vários factores ligados às condições deficientes de habitabilidade para o risco no AVC, a incapacidade de aquecer o local onde se habita parece (no conjunto limitado de variáveis analisadas) constituir o determinante mais crítico.

Mas mais importante do que constatar a importância das baixas temperaturas na habitação, foi possível, partindo de um nível de desagregação territorial que representa uma escala geográfica de análise razoavelmente grande, delimitar áreas do território continental Português que poderiam assim constituir alvos primordiais de intervenções de Saúde Pública para mitigar ou prevenir o efeito deste determinante, focadas preferencialmente, neste caso, no coorte populacional das pessoas com mais de 64 anos de idade.

4. Discussão geral e conclusões

O cruzamento entre determinantes socioeconómicos e risco de AVC (e em particular o risco de morte, nos dois casos de estudo que integram esta tese nos capítulos 2 e 3) tem vindo a ser estudado desde há já algum tempo, tanto internacionalmente como em Portugal, como se observa em vários estudos citados ao longo desta tese, tais como Cox et al. (2006) e Addo et al. (2012), e no contexto português Vasconcelos et al. (2011), entre outros. Todavia, o objectivo primordial desta tese remete, não para a percepção geral da associação entre estes determinantes e o risco de AVC, mas sim para a avaliação de como a distribuição espacial dos determinantes socioeconómicos pode influir na distribuição espacial do risco. Reportando especificamente ao AVC, conhecem-se poucas investigações neste âmbito, citando-se a título de exemplo os trabalhos de Soljak et al. (2011), abordando a variação espacial dos casos de sub-diagnóstico de doenças cardiovasculares (incluindo o AVC) no Reino Unido, de Odoi e Busingye (2014), que investigam a associação entre disparidades geográficas registadas ao nível dos factores socioeconómicos entre pequenas áreas censitárias e a mortalidade por AVC, e de Schieb et al. (2013), que utilizaram índices LISA para detectar clusters de municípios (counties) com valores altos ou baixos de taxas de hospitalização por AVC, tendo investigado de seguida a relação entre os clusters dos dois tipos e a caracterização socioeconómicos dos municípios por eles englobados, nos Estados Unidos da América.

Neste contexto, a primeira questão a investigar consiste em perceber se existem, ao longo do território de Portugal Continental, importantes disparidades na distribuição geográfica tanto dos determinantes económicos como do risco de morte. A visualização de ambos, através de mapas coropletos, permitiu, desde logo, nos dois estudos, constatar a existência destas disparidades nos valores de ambos os conjuntos de variáveis, registados nas unidades estatísticas espaciais (municípios ou freguesias, consoante o estudo do capítulo 2 ou 3, respectivamente) ao longo da área de estudo correspondente a Portugal Continental.

A segunda questão remete para a segunda fase dos estudos em Epidemiologia Espacial, fase de exploração. Será a existência (e a localização) destas disparidades significativa

do ponto de vista estatístico? O recurso a técnicas de análise de clusters espaciais, recorrendo ao índice I de Moran e índices LISA permitiu constatar, por um lado, a existência de autocorrelação espacial positiva e estatisticamente significativa na distribuição espacial das variáveis em ambos os casos estudados, e por outro lado, detectar a localização de clusters de municípios ou freguesias com valores altos ou baixos, e também estatisticamente significativos.

A terceira questão surge um passo mais à frente e constitui aquela que responde, na prática, ao primeiro objectivo geral: será que estamos em presença de variação espacial na associação entre determinantes e risco de morte por AVC, e qual a intensidade com que essa variação se verifica; por outras palavras, se existe não-estacionaridade espacial no processo de associação e se esta é expressiva ao longo da área de estudo ou, pelo contrário, representa um fenómeno pontual e de pequena magnitude. Ambos os casos de estudo parecem sugerir que esta variação existe e tem uma força razoável.

Assim, no estudo do capítulo 2, a presença desta variação constata-se pela existência de deriva espacial nos coeficientes locais (municipais) dos modelos de regressão geograficamente ponderada (GWR). Adicionalmente, os padrões espaciais formados pela distribuição espacial destes coeficientes são, na maior parte, bem expressivos, como se pode observar no mapa coropleto dos coeficientes que expressam a associação entre a percentagem de mulheres que atingiram o ensino básico e o risco de AVC em mulheres com menos de 65 anos no período de 1992/96, os quais são de sinal negativo na maior parte do território continental português (constituindo assim o nível de escolaridade um factor de protecção), mas apresentam-se positivos numa área genericamente centrada na Área Metropolitana do Porto (o que sugere que nesta área este determinante constitui antes um factor de aumento do risco). Esta situação permite aferir da importância de utilizar modelos locais, capazes de espacializar estas associações, no sentido de aprofundar os resultados dos modelos globais de regressão, cujo resultado é aplicável a toda a área de estudo. Quando se recorre apenas a modelos globais, corre-se o risco de serem ignoradas potenciais variações territoriais relevantes na associação entre determinantes socioeconómicos e risco de morte por AVC. A este respeito, note-se que o modelo global da associação entre a percentagem de mulheres que atingiram o ensino básico e o risco de AVC em mulheres com menos de 65 anos no

período de 1992/96, apontava para uma associação negativa ao longo de toda a área de estudo.

O modelo de análise ensaiado no estudo do capítulo 3, embora bastante diferente, permite também colocar a hipótese da presença de não-estacionaridade espacial, particularmente na associação entre excesso de risco de morte por AVC em vários coortes populacionais (entre os quais as pessoas com mais de 64 anos) e a percentagem de pessoas que não possuem meios de aquecer a sua habitação durante o inverno, ao nível da freguesia. Assim, não só a conjugação de vários determinantes relacionados com a má qualidade da habitação, numa análise de clusters multivariada, revelou 3 clusters de freguesias bem delimitados no território continental português, como num dos clusters os valores médios e medianos do excesso de risco de morte por AVC, em todos os coortes populacionais em estudo, foram superiores aos valores registados nos dois clusters restantes. Mais ainda, o determinante com mais expressão nesse cluster (cujos valores médios e medianos estandardizados apresentavam valores mais altos nas freguesias contidas no cluster) era justamente a percentagem de pessoas que não possuem meios de aquecer a sua habitação durante o inverno.

A conjugação das várias análises efectuadas permitiu avaliar, para além dos resultados anteriores, quais os determinantes que se revelaram mais importantes em cada estudo e a força da sua associação com o risco de morte por AVC. No estudo do capítulo 2, os resultados sugerem que, ao nível global (em Portugal Continental com um todo), o nível de escolaridade parece ser um importante factor de protecção face ao risco de morte por AVC, pois os municípios com elevadas percentagens de pessoas tendo atingido o ensino básico tenderam para um risco de morte por AVC consistentemente baixo, nos modelos de ambos os sexos no período de 2002/06, e pelo menos no modelo para mulheres no período temporal de 1992/96. Nota-se também consistência na distribuição espacial da força de associação, no sentido em que, mesmo nos modelos locais, na maior parte dos municípios e em ambos os períodos de estudo (nos três modelos acima referidos), este determinante era um factor de protecção, exceptuando nos casos (minoritários) dos municípios em que esta situação se inverte, casos estes que poderão merecer uma análise mais aprofundada em estudos subsequentes. Já os resultados do segundo estudo (capítulo 3) permitem colocar a hipótese de que, de entre as várias variáveis representativas das más condições de habitabilidade incluídas no estudo, a incapacidade

de aquecer a habitação durante o inverno representa o mais importante factor de aumento do excesso de risco de morte por AVC, e que a sua influência poderá ser mais sentida num cluster de freguesias da área noroeste do território continental, centrado na Área Metropolitana do Porto.

No que respeita ao desenho epidemiológico dos dois estudos, ambos do tipo ecológico, é relevante debater e avaliar algumas das escolhas efectuadas. No que respeita aos coortes populacionais investigados nos dois estudos, pretendeu-se investigar em separado determinantes relevantes do risco de morte por AVC em coortes etários opostos da população e, simultaneamente, enriquecer o conteúdo da tese em geral, tentando abranger, mesmo que em estudos separados, a população em geral.

Desta forma, o estudo do capítulo 2 aborda o risco de morte por AVC na população abaixo de 65 anos de idade, alicerçando-se na discussão conjunta dos estudos de Krishnamurthi et al. (2013), que sinalizam entre 1990 e 2010 um aumento à escala mundial e estatisticamente significativo na incidência em pessoas com idades compreendidas entre os 20 e os 64 anos (para o qual contribuíram substancialmente os países de baixo e médio rendimento) e de Cox et al. (2006) e Addo et al. (2012), que referem que o risco de AVC apresenta padrões diversos no que respeita à sua distribuição geográfica, não só entre países mas também entre regiões do mesmo país. Assim, a junção destas contribuições suscitou interesse na investigação das disparidades geográficas na associação entre determinantes socioeconómicos e risco de AVC no território continental português na população na população abaixo de 65 anos de idade, a qual se encontra, na sua maior parte, em idade laboral.

Por outro lado, como sugerido em WHO (2014a), não é menos verdade que esta é uma doença que ocorre com mais frequência em adultos de meia-idade ou em idosos, como sugerem os resultados que decorrem da comparação dos indicadores do excesso de risco de morte por AVC, calculados para coortes de todas as idades e também de mais de 64 anos de idade. Nestes resultados (e nos valores absolutos de óbitos) é patente que os coortes populacionais de mais de 64 anos são os que mais contribuem para a mortalidade por AVC no contexto da população em geral. E assim, procurou-se, no estudo do capítulo 3, investigar um conjunto de determinantes socioeconómicos ligados às más condições de habitabilidade, assumindo-se à partida que estas podem contribuir de forma mais importante para o excesso de risco de AVC. Entre outros motivos, esta

situação deve-se ao maior grau de exposição da população idosa a estes determinantes, a qual, por ter na sua maioria atingido a idade da reforma, apresenta uma tendência natural para permanecer na sua habitação durante períodos de tempo mais longos.

Outra questão relevante, ainda no desenho epidemiológico dos dois estudos, teve a ver com as unidades estatísticas espaciais utilizadas em cada um. Assim, com a escolha do município como unidade estatística no estudo do capítulo 2, tentou-se minimizar a ocorrência de números pequenos e seus efeitos nefastos (Barceló et al., 2008), pois mesmo agregando 5 anos de mortalidade em cada período temporal deste estudo, o reduzido número de óbitos por AVC em pessoas em idade activa levaria a um brutal aumento da magnitude deste problema, caso se tivesse recorrido antes à freguesia como unidade estatística. A mesma discussão, mas em sentido inverso, justifica a adopção da freguesia como unidade estatística espacial no estudo do capítulo 3. Em qualquer dos casos, o recurso a qualquer um destes níveis de desagregação espacial no contexto dos dois estudos efectuados é, tanto quanto foi possível apurar na análise prévia da bibliografia, inédito em Portugal. Alguns dos estudos internacionais já citados ao longo desta tese utilizaram, de facto, unidades estatísticas espaciais a um nível de desagregação espacial muito grande. Em estudos relacionados com o AVC, Schieb et al. (2013) recorrem ao município, mas Odoi e Busingye (2014) recorrem já a unidades censitárias de maior desagregação (census tracts), ambos os estudos no contexto norte-americano.

De facto, se nos abstrairmos do problema dos números pequenos, e se for possível a recolha de todos os dados ao nível de desagregação pretendido, o recurso a unidades estatísticas espaciais de grande escala (no sentido cartográfico) possui a vantagem de minimizar o problema da unidade areal modificável (MAUP) (Longley et al., 2010), na medida em que os resultados dependem da divisão do espaço.

Como segundo objectivo primordial deste trabalho, pretendia-se ensaiar estudos que pudessem contribuir para melhorar a tomada de decisões ao nível das medidas de prevenção e mitigação dos casos de AVC em Portugal. A este nível, não obstante o seu carácter exploratório, o estudo do capítulo 3 permitiu delinear uma área onde se coloca a hipótese de uma intervenção de Saúde Pública focada na população que não possui a capacidade para aquecer devidamente a sua habitação no inverno, podendo contribuir assim para prevenir ou mitigar a ocorrência de episódios de AVC na área em causa.

Consequentemente, e admitindo a prossecução deste rumo de investigação, a metodologia utilizada poderia (salvaguardadas as devidas calibrações), vir a contribuir para a escolha de áreas de intervenção prioritárias no território de Portugal Continental, podendo, em última análise, ser integrada na vigilância epidemiológica do AVC.

Os estudos ensaiados nestes trabalhos possuem, obviamente, algumas limitações. Uma destas prende-se com o reduzido número de indicadores socioeconómicos relevantes para estes estudos que se encontra disponível para recolha imediata através dos produtores oficiais de dados estatísticos, em particular a escalas de desagregação espacial muito grandes. Esta situação contrasta com o que acontece, por exemplo, nos Estados Unidos da América. Um exemplo flagrante consiste na disponibilização, a estas escalas, de estatísticas da população abaixo do limiar de pobreza, uma variável que teria sido sem dúvida interessante incluir no estudo do capítulo 2. Outra limitação prende-se com o desenho ecológico dos dois estudos ensaiados. Não obstante, estes estudos permitem obter conhecimento útil, possibilitando, entre outras vantagens, estreitar o foco de investigação, ao introduzir um primeiro filtro (por exemplo ao nível dos coortes ou do âmbito territorial a abordar nos estudos seguintes), o que se pode traduzir numa racionalização dos recursos de Saúde Pública empregues, especialmente se, como foi o caso aqui, os custos de recolha de dados e processamento forem diminutos.

Os estudos de aplicação ensaiados no âmbito desta tese despertaram interesse para vários desenvolvimentos futuros. Entre estes, uma possibilidade seria a investigação de associações estatísticas em ambos estudos com recurso a modelos hierárquicos bayesianos, tal como nos estudos de Odoi e Busingye (2014) e Hu et al. (2008). Outra hipótese seria o recurso a modelos de regressão do tipo Zero-Inflated (Sartorius et al., 2011), especialmente dotados para lidar com a presença de um grande número de unidades espaciais com zero ocorrências registadas, por exemplo na eventualidade do ensaio do estudo do capítulo 2, partindo da freguesia como unidades estatística espacial. Um outro desenvolvimento possível, que não foi efectuado por questões de gestão do tempo disponível, seria prosseguir o estudo do capítulo 3 através da utilização de modelos de regressão. Finalmente, tendo em conta que os estudos ensaiados se focaram em períodos temporais anteriores à crise económica iniciada em 2008, a operacionalização de ambos com dados mais actuais (por exemplo a repetição do estudo do capítulo 2 com dados de mortalidade do período 2012-2016 comparados com dados

socioeconómicos referidos aos Censos de 2011) e comparação de resultados entre períodos, poderia incrementar bastante os benefícios destas análises e ajudar a perceber melhor os efeitos da presente situação de crise económica e financeira no risco de AVC.

5. Referências bibliográficas

- ACSS [WWW Document], 2014. URL <http://www.acss.min-saude.pt/> (acedido 5.11.14).
- Addo, J., Ayerbe, L., Mohan, K.M., Crichton, S., Sheldenkar, A., Chen, R., Wolfe, C.D.A., McKeivitt, C., 2012. Socioeconomic status and stroke: an updated review. *Stroke J. Cereb. Circ.* 43, 1186–1191. doi:10.1161/STROKEAHA.111.639732
- Age UK [WWW Document], 2012. The Cost of Cold. URL http://www.ageuk.org.uk/Documents/EN-GB/Campaigns/The_cost_of_cold_2012.pdf (acedido 7.11.14).
- Ahmad, N.H., Othman, I.R., Deni, S.M., 2013. Hierarchical Cluster Approach for Regionalization of Peninsular Malaysia based on the Precipitation Amount. *J. Phys. Conf. Ser.* 423, 012018. doi:10.1088/1742-6596/423/1/012018
- Ahrens, W., Pigeot, I. (Eds.), 2004. *Handbook of Epidemiology*, 1st ed. 2005. Corr. 2nd printing edition. ed. Springer, Berlin.
- Anselin, L., 2005. *Exploring Spatial Data with GeoDa: A Workbook*.
- Anselin, L., 2001. Spatial Econometrics, in: *PALGRAVE HANDBOOK OF ECONOMETRICS: VOLUME 1, ECONOMETRIC THEORY*. Kluwer, pp. 310–330.
- Anselin, L., Lozano-Gracia, N., Deichmann, U., Lall, S., 2010. Valuing access to water—a spatial hedonic approach, with an application to Bangalore, India. *Spat. Econ. Anal.* 5, 161–179.
- ArcGIS Explorer, 2015. ArcGIS Explorer | GIS Viewer | Free GIS Software & Maps [WWW Document]. URL <http://www.esri.com/software/arcgis/explorer> (acedido 3.14.15).
- Arlinghaus, S.L., Griffith, D.A.--S.A.E., Arlinghaus, W.C.--A.E., Drake, W.D.--A.E., Nystuen, J.D.--A.E., Arlinghaus, S.L., Nystuen, J.D., Griffith, D.A., Vasiliev, I.R., Stehman, S.V., Overton, W.S., Wong, D.W.S., Li, B., Brown, D.G., Feng, H.M., Can, A., Long, D.S., Arlinghaus, S.L., 1996. *Practical Handbook of Spatial Statistics*. CRC Press.
- Avendaño, M., Kunst, A.E., Huisman, M., van Lenthe, F., Bopp, M., Borrell, C., Valkonen, T., Regidor, E., Costa, G., Donkin, A., Borgan, J.-K., Deboosere, P., Gadeyne, S., Spadea, T., Andersen, O., Mackenbach, J.P., 2004. Educational level and stroke mortality: a comparison of 10 European populations during the 1990s. *Stroke J. Cereb. Circ.* 35, 432–437. doi:10.1161/01.STR.0000109225.11509.EE

- Baço, F., Lobo, V., Painho, M., 2005. The self-organizing map, the Geo-SOM, and relevant variants for geosciences. *Comput. Geosci., Geospatial Research in Europe: AGILE 2003* 31, 155–163. doi:10.1016/j.cageo.2004.06.013
- Balamurugan, A., Delongchamp, R., Bates, J.H., Mehta, J.L., 2013. The neighborhood where you live is a risk factor for stroke. *Circ. Cardiovasc. Qual. Outcomes* 6, 668–673. doi:10.1161/CIRCOUTCOMES.113.000265
- Balcan, M.-F., Liang, Y., Gupta, P., 2013. Robust Hierarchical Clustering. *ArXiv14010247 Cs*.
- Barceló, A.M., Saez, M., Cano-Serral, G., Martínez-Beneito, M.Á., Martínez, J., Borrell, C., Ocaña-Riola, R., Montoya, I., Calvo, M., López-Abente, G., Rodríguez-Sanz, M., Toro, S., Alcalá, J.T., Saurina, C., Sánchez-Villegas, P., Figueiras, A., 2008. Métodos para la suavización de indicadores de mortalidad: aplicación al análisis de desigualdades en mortalidad en ciudades del Estado español (Proyecto MEDEA). *Gac. Sanit.* 22, 596–608. doi:10.1016/S0213-9111(08)75362-7
- Barlow, C.F., 1984. *Headaches and Migraine in Childhood*. Cambridge University Press.
- Bendixen, B.H., Posner, J., Lango, R., 2001. Stroke in young adults and children. *Curr. Neurol. Neurosci. Rep.* 1, 54–66.
- Besag, J., York, J., Mollié, A., 1991. Bayesian image restoration, with two applications in spatial statistics. *Ann. Inst. Stat. Math.* 43, 1–20. doi:10.1007/BF00116466
- Biller, J., 2009. *Stroke in Children and Young Adults*. Elsevier Health Sciences.
- BioMedware, 2014. BioMedware ClusterSeer | BioMedware [WWW Document]. URL <http://www.biomedware.com/?module=Page&sID=clusterseer> (accedido 6.21.14).
- Blessing, J.N.R., Saposnik, G., Nisenbaum, R., Schweizer, T.A., Reddy, D., Thorpe, K.E., Macdonald, R.L., 2013. Effect of Socioeconomic Status on Inpatient Mortality and Use of Postacute Care After Subarachnoid Hemorrhage. *Stroke* 44, 2842–2847. doi:10.1161/STROKEAHA.113.001368
- Bogousslavsky, J., Pierre, P., 1992. Ischemic stroke in patients under age 45. *Neurol. Clin.* 10, 113–124.
- Boulos, K.M.N., Roudsari, A.V., Carson, E.R., 2001. Health Geomatics: An Enabling Suite of Technologies in Health and Healthcare. *J. Biomed. Inform.* 34, 195–219. doi:10.1006/jbin.2001.1015
- Brainin, M., Heiss, W.-D. (Eds.), 2009. *Textbook of stroke medicine*. Cambridge University Press, Cambridge, UK ; New York.

- Burrough, P.A., McDonnell, R.A., 1998. Principles of Geographical Information Systems. Spatial Information Systems.
- Busingye, D., Pedigo, A., Odoi, A., 2011. Temporal changes in geographic disparities in access to emergency heart attack and stroke care: Are we any better today? *Spat. Spatio-Temporal Epidemiol.* 2, 247–263. doi:10.1016/j.sste.2011.07.010
- Carvalho, A., Albuquerque, P., Almeida Junior, G., Guimarães, R., 2009. Spatial Hierarchical Clustering. *Rev Bras Biom* 27, 411–442.
- Cesana, G., Ferrario, M., Gigante, S., Sega, R., Toso, C., Achilli, F., 2001. Socio-occupational differences in acute myocardial infarction case-fatality and coronary care in a northern Italian population. *Int. J. Epidemiol.* 30 Suppl 1, S53–58.
- Cheng, E.M., Atkinson, P.M., Shahani, A.K., 2011. Elucidating the spatially varying relation between cervical cancer and socio-economic conditions in England. *Int. J. Health Geogr.* 10, 51. doi:10.1186/1476-072X-10-51
- Clayton, D., Kaldor, J., 1987. Empirical Bayes Estimates of Age-Standardized Relative Risks for Use in Disease Mapping. *Biometrics* 43, 671–681. doi:10.2307/2532003
- Colhoun, H.M., Hemingway, H., Poulter, N.R., 1998. Socio-economic status and blood pressure: an overview analysis. *J. Hum. Hypertens.* 12, 91–110.
- Collins, R., MacMahon, S., 1994. Blood pressure, antihypertensive drug treatment and the risks of stroke and of coronary heart disease. *Br. Med. Bull.* 50, 272–298.
- Costa-Font, J., 2008. Housing assets and the socio-economic determinants of health and disability in old age. *Health Place* 14, 478–491. doi:10.1016/j.healthplace.2007.09.005
- Couceiro, L., Santana, P., Nunes, C., 2011. Pulmonary tuberculosis and risk factors in Portugal: a spatial analysis. *Int. J. Tuberc. Lung Dis. Off. J. Int. Union Tuberc. Lung Dis.* 15, 1445–1454, i. doi:10.5588/ijtld.10.0302
- Cox, A.M., McKevitt, C., Rudd, A.G., Wolfe, C.D.A., 2006. Socioeconomic status and stroke. *Lancet Neurol.* 5, 181–188. doi:10.1016/S1474-4422(06)70351-9
- Cromley, E.K., 2003. GIS and Disease. *Annu. Rev. Public Health* 24, 7–24. doi:10.1146/annurev.publhealth.24.012902.141019
- Cunha, G., Martins, M. do R., Sousa, R., Oliveira, F.F. de, 2007. Estatística Aplicada às Ciências e Tecnologias da Saúde. Lidel Edicoes Tecnicas.
- De Smith, M., Goodchild, M., Longley, P., 2015. Geospatial Analysis 5th Edition, 2015 [WWW Document]. URL

- http://www.spatialanalysisonline.com/HTML/index.html?suggested_reading.htm (acedido 4.24.15).
- DGS, 2013. Portugal, Doenças Cérebro-Cardiovasculares em números - 2013. Lisboa: DGS.
- DGT, 2015. DGTerritório - Início [WWW Document]. URL <http://www.dgterritorio.pt/> (acedido 3.15.15).
- Dobson, A.J., 2002. An introduction to generalized linear models, 2nd ed. ed, Chapman & Hall/CRC texts in statistical science series. Chapman & Hall/CRC, Boca Raton.
- Dorling, D., 1995. The visualization of local urban change across Britain. *Environ. Plan. B Plan. Des.* 22, 269–290. doi:10.1068/b220269
- Elliott, P., Wakefield, J., Best, N., Briggs, D., 2001. *Spatial Epidemiology*. Oxford University Press.
- Elliott, P., Wartenberg, D., 2004. *Spatial Epidemiology: Current Approaches and Future Challenges*. *Environ. Health Perspect.* 112, 998–1006. doi:10.1289/ehp.6735
- Engström, G., Jerntorp, I., Pessah-Rasmussen, H., Hedblad, B., Berglund, G., Janzon, L., 2001. Geographic Distribution of Stroke Incidence Within an Urban Population Relations to Socioeconomic Circumstances and Prevalence of Cardiovascular Risk Factors. *Stroke* 32, 1098–1103. doi:10.1161/01.STR.32.5.1098
- EPSG, 2015. ETRS89 / Portugal TM06: EPSG Projection -- Spatial Reference [WWW Document]. URL <http://spatialreference.org/ref/epsg/etrs89-portugal-tm06/> (acedido 3.15.15).
- EROS Center, 2015. Home Page | Earth Resources Observation and Science (EROS) Center [WWW Document]. URL <http://eros.usgs.gov/> (acedido 3.2.15).
- ESO, 2013. European Stroke Organisation [WWW Document]. URL <http://www.eso-stroke.org/eso-stroke/home.html> (acedido 5.25.14).
- ESRI, 2014. ArcGIS for Desktop [WWW Document]. URL <http://www.esri.com/software/arcgis/arcgis-for-desktop> (acedido 12.22.14).
- Eurostat, 2014a. Eurostat [WWW Document]. URL <http://ec.europa.eu/eurostat> (acedido 9.11.14).
- Eurostat, 2014b. Living conditions in Europe [WWW Document]. URL <http://ec.europa.eu/eurostat/en/web/products-statistical-books/-/KS-DZ-14-001> (acedido 10.24.14).
- Everitt, B., Landau, S., Leese, M., Stahl, D., 2011. *Cluster analysis*. Wiley.

- Fahrmeir, L., Kneib, T., Lang, S., Marx, B., 2013. Regression: Models, Methods and Applications. Springer Science & Business Media.
- Falcão, I.V., Carvalho, E.M.F. de, 2004. Acidente vascular cerebral precoce: implicações para adultos em idade produtiva atendidos pelo Sistema Único de Saúde. Rev. Bras. Saúde Materno Infant. doi:10.1590/S1519-38292004000100009
- Faltin, F., Kenett, R., Ruggeri, F., 2012. Statistical Methods in Healthcare. John Wiley & Sons.
- Feigin, V.L., Forouzanfar, M.H., Krishnamurthi, R., Mensah, G.A., Connor, M., Bennett, D.A., Moran, A.E., Sacco, R.L., Anderson, L., Truelsen, T., O'Donnell, M., Venketasubramanian, N., Barker-Collo, S., Lawes, C.M.M., Wang, W., Shinohara, Y., Witt, E., Ezzati, M., Naghavi, M., Murray, C., 2014. Global and regional burden of stroke during 1990–2010: findings from the Global Burden of Disease Study 2010. Lancet 383, 245–254.
- Fotheringham, A.S., Brunson, C., Charlton, M., 2002. Geographically weighted regression: the analysis of spatially varying relationships. Wiley, Chichester, England ; Hoboken, NJ, USA.
- Franks, P.J., Adamson, C., Bulpitt, P.F., Bulpitt, C.J., 1991. Stroke death and unemployment in London. J. Epidemiol. Community Health 45, 16–18.
- Fukuda, Y., Nakamura, K., Takano, T., 2005. Cause-specific mortality differences across socioeconomic position of municipalities in Japan, 1973–1977 and 1993–1998: increased importance of injury and suicide in inequality for ages under 75. Int. J. Epidemiol. 34, 100–109. doi:10.1093/ije/dyh283
- GeoDa, 2014. GeoDa | GeoDa Center [WWW Document]. URL <https://geodacenter.asu.edu/software/downloads> (acedido 12.22.14).
- GeoDa Center, 2015. GeoDa Center | Spatial methods and tools - geodacenter.asu.edu [WWW Document]. URL <https://geodacenter.asu.edu/> (acedido 4.14.15).
- Geography Hunter, 2015. Kernel density calculations [WWW Document]. URL <http://www.geography.hunter.cuny.edu/~jochen/GTECH361/lectures/lecture11/concepts/Kernel%20density%20calculations.htm> (acedido 3.12.15).
- GeoNetwork, 2015. GeoNetwork opensource [WWW Document]. URL <http://geonetwork-opensource.org/> (acedido 3.1.15).
- GeoSOM, 2015. GeoSOM Home [WWW Document]. URL <http://www.novaims.unl.pt/labnt/geosom/> (acedido 4.24.15).
- Google Earth, 2014. Google Earth [WWW Document]. URL <https://www.google.com/earth/> (acedido 12.21.14).

- Gravetter, F.J., Wallnau, L.B., 2013. *Essentials of Statistics for the Behavioral Sciences*, 8 edition. ed. Cengage Learning, Belmont, CA.
- Guo, D., Gahegan, M., 2006. Spatial ordering and encoding for geographic data mining and visualization. *J. Intell. Inf. Syst.* 27, 243–266. doi:10.1007/s10844-006-9952-8
- Haining, R., 2003. *Spatial Data Analysis: Theory and Practice*, 1 edition. ed. Cambridge University Press, Cambridge.
- Hanchate, A.D., Schwamm, L.H., Huang, W., Hylek, E.M., 2013. Comparison of ischemic stroke outcomes and patient and hospital characteristics by race/ethnicity and socioeconomic status. *Stroke J. Cereb. Circ.* 44, 469–476. doi:10.1161/STROKEAHA.112.669341
- Hartigan, J.A., Wong, M.A., 1979. Algorithm AS 136: A K-Means Clustering Algorithm. *J. R. Stat. Soc. Ser. C Appl. Stat.* 28, 100–108. doi:10.2307/2346830
- Haslett, J., Bradley, R., Craig, P., Unwin, A., Wills, G., 1991. Dynamic Graphics for Exploring Spatial Data with Application to Locating Global and Local Anomalies. *Am. Stat.* 45, 234–242. doi:10.2307/2684298
- Hayashi, T., Kobayashi, Y., Yamaoka, K., Yano, E., 1996. Effect of overtime work on 24-hour ambulatory blood pressure. *J. Occup. Environ. Med. Am. Coll. Occup. Environ. Med.* 38, 1007–1011.
- Health Canada, 2012. Health Canada - Home Page [WWW Document]. URL <http://www.hc-sc.gc.ca/index-eng.php> (acedido 2.25.12).
- Helbich, M., Leitner, M., Kapusta, N.D., 2012. Geospatial examination of lithium in drinking water and suicide mortality. *Int. J. Health Geogr.* 11, 19. doi:10.1186/1476-072X-11-19
- Henriques, F.C., Ferreira Rodrigues, T., Fraga O. Martins, M., 2009. Ageing, Education and Health in Portugal: Prospective from the 19th to the 21st Century. *Hygiea Int. Interdiscip. J. Hist. Public Health* 8, 81–96. doi:10.3384/hygiea.1403-8668.098181
- Henriques, R., Bacao, F., Lobo, V., 2012. Exploratory geospatial data analysis using the GeoSOM suite. *Comput. Environ. Urban Syst.* 36, 218–232. doi:10.1016/j.compenvurbsys.2011.11.003
- HFA-DB, 2011. European Health for All database (HFA-DB) [WWW Document]. URL <http://www.euro.who.int/en/data-and-evidence/databases/european-health-for-all-database-hfa-db> (acedido 3.15.11).
- Hilbe, J.M., 2007. *Negative binomial regression*. Cambridge University Press, Cambridge ; New York.

- Hillbom, M., Haapaniemi, H., Juvela, S., Palomäki, H., Numminen, H., Kaste, M., 1995. Recent Alcohol Consumption, Cigarette Smoking, and Cerebral Infarction in Young Adults. *Stroke* 26, 40–45. doi:10.1161/01.STR.26.1.40
- Howard, G., Anderson, R., Johnson, N.J., Sorlie, P., Russell, G., Howard, V.J., 1997. Evaluation of Social Status as a Contributing Factor to the Stroke Belt Region of the United States. *Stroke* 28, 936–940. doi:10.1161/01.STR.28.5.936
- Howard, G., Evans, G.W., Pearce, K., Howard, V.J., Bell, R.A., Mayer, E.J., Burke, G.L., 1995. Is the Stroke Belt Disappearing? An Analysis of Racial, Temporal, and Age Effects. *Stroke* 26, 1153–1158. doi:10.1161/01.STR.26.7.1153
- Hulle, M.M.V., 2012. Self-organizing Maps, in: Rozenberg, G., Bäck, T., Kok, J.N. (Eds.), *Handbook of Natural Computing*. Springer Berlin Heidelberg, pp. 585–622.
- Hu, Z., Liebens, J., Rao, K.R., 2008. Linking stroke mortality with air pollution, income, and greenness in northwest Florida: an ecological geographical study. *Int. J. Health Geogr.* 7, 20. doi:10.1186/1476-072X-7-20
- IAPMEI, 2014. IAPMEI, I.P. - Agência para a Competitividade e Inovação [WWW Document]. URL <http://www.iapmei.pt/> (acedido 2.22.15).
- ICD-10, 2015. International Statistical Classification of Diseases and Related Health Problems 10th Revision [WWW Document]. URL <http://apps.who.int/classifications/icd10/browse/2015/en> (acedido 12.21.14).
- iGEO, 2015. iGEO - Informação Geográfica [WWW Document]. URL <http://www.igeo.pt/> (acedido 3.5.15).
- IJHG, 2015. International Journal of Health Geographics [WWW Document]. URL <http://www.ij-healthgeographics.com/> (acedido 3.14.15).
- INE, 2014. Portal do Instituto Nacional de Estatística [WWW Document]. URL http://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_main (acedido 12.21.14).
- Inspire, 2015. INSPIRE > WELCOME TO INSPIRE [WWW Document]. URL <http://inspire.ec.europa.eu/> (acedido 3.3.15).
- ISO, 2015. ISO - International Organization for Standardization [WWW Document]. URL <http://www.iso.org/iso/home.html> (acedido 3.3.15).
- Jakovljević, D., Sivenius, J., Sarti, C., Torppa, J., Mähönen, M., Immonen-Räihä, P., Kaarsalo, E., Alhainen, K., Tuomilehto, J., Puska, P., Salomaa, V., 2001. Socioeconomic inequalities in the incidence, mortality and prognosis of subarachnoid hemorrhage: the FINMONICA Stroke Register. *Cerebrovasc. Dis. Basel Switz.* 12, 7–13. doi:47674

- Kiely, D.K., Wolf, P.A., Cupples, L.A., Beiser, A.S., Myers, R.H., 1993. Familial aggregation of stroke. The Framingham Study. *Stroke J. Cereb. Circ.* 24, 1366–1371.
- Kleinbaum, D.G., Klein, M., c2010. *Logistic regression : a self-learning text /*, 3rd ed. ed. Springer, New York :
- Kohonen, T., 1990. The self-organizing map. *Proc. IEEE* 78, 1464–1480. doi:10.1109/5.58325
- Kosfeld, R., Lauridsen, J., 2012. Identifying Clusters within R&D Intensive Industries Using Local Spatial Methods (MAGKS Papers on Economics No. 201214). Philipps-Universität Marburg, Faculty of Business Administration and Economics, Department of Economics (Volkswirtschaftliche Abteilung).
- Krishnamurthi, R.V., Feigin, V.L., Forouzanfar, M.H., Mensah, G.A., Connor, M., Bennett, D.A., Moran, A.E., Sacco, R.L., Anderson, L.M., Truelsen, T., O'Donnell, M., Venketasubramanian, N., Barker-Collo, S., Lawes, C.M.M., Wang, W., Shinohara, Y., Witt, E., Ezzati, M., Naghavi, M., Murray, C., 2013. Global and regional burden of first-ever ischaemic and haemorrhagic stroke during 1990–2010: findings from the Global Burden of Disease Study 2010. *Lancet Glob. Health* 1, e259–e281. doi:10.1016/S2214-109X(13)70089-5
- Kunst, A.E., del Rios, M., Groenhouf, F., Mackenbach, J.P., 1998. Socioeconomic inequalities in stroke mortality among middle-aged men: an international overview. European Union Working Group on Socioeconomic Inequalities in Health. *Stroke J. Cereb. Circ.* 29, 2285–2291.
- Labarthe, D.R., 2011. *Epidemiology And Prevention Of Cardiovascular Diseases: A Global Challenge*, 2 edition. ed. Jones & Bartlett Learning, Sudbury, Mass.
- Lambert, C., Vinson, S., Shofer, F., Brice, J., 2013. The relationship between knowledge and risk for heart attack and stroke. *J. Stroke Cerebrovasc. Dis. Off. J. Natl. Stroke Assoc.* 22, 996–1001. doi:10.1016/j.jstrokecerebrovasdis.2012.02.002
- Landsbergis, P., 2004. Long work hours, hypertension, and cardiovascular disease. *Cad. Saúde Pública* 20, 1746–1748. doi:10.1590/S0102-311X2004000600035
- Lanska, D.J., Peterson, P.M., 1995. Geographic Variation in the Decline of Stroke Mortality in the United States. *Stroke* 26, 1159–1165. doi:10.1161/01.STR.26.7.1159
- Lawson, A.B., 2001. *Statistical Methods in Spatial Epidemiology*. Wiley.
- Lawson, A.B., Williams, F.L.R., 2001. *An Introductory Guide to Disease Mapping*, 1 edition. ed. Wiley, Chichester.
- Lawson, A., Denison, D., 2002. *Spatial Cluster Modelling*. Chapman and Hall/CRC.

- LeSage, J., Pace, R.K., 2009. *Introduction to Spatial Econometrics*. CRC Press.
- Li, C., Hedblad, B., Rosvall, M., Buchwald, F., Khan, F.A., Engström, G., 2008. Stroke Incidence, Recurrence, and Case-Fatality in Relation to Socioeconomic Position A Population-Based Study of Middle-Aged Swedish Men and Women. *Stroke* 39, 2191–2196. doi:10.1161/STROKEAHA.107.507756
- Lloyd, C., 2010. *Local Models for Spatial Analysis, Second Edition*. CRC Press.
- Longley, P.A., Goodchild, M.F., Maguire, D.J., Rhind, D.W., 2005. *Geographic Information Systems and Science, 2nd Edition* edition. ed. John Wiley & Sons, Chichester.
- Longley, P.A., Goodchild, M., Maguire, D.J., Rhind, D.W., 2010. *Geographic Information Systems and Science, 3rd Edition* edition. ed. John Wiley & Sons.
- Longley, P., Tobón, C., 2003. Spatial dependence and heterogeneity in patterns of urban deprivation (ERSA conference paper No. ersa03p132). European Regional Science Association.
- Maheswaran, R., Elliott, P., 2003. Stroke Mortality Associated With Living Near Main Roads in England and Wales A Geographical Study. *Stroke* 34, 2776–2780. doi:10.1161/01.STR.0000101750.77547.11
- Maheswaran, R., Pearson, T., Beevers, S.D., Campbell, M.J., Wolfe, C.D., 2014. Outdoor air pollution, subtypes and severity of ischemic stroke – a small-area level ecological study. *Int. J. Health Geogr.* 13, 23. doi:10.1186/1476-072X-13-23
- Marmot, M., Wilkinson, R. (Eds.), 2005. *Social Determinants of Health, Second Edition*. ed. Also available as: eBook.
- Marsh, A., Gordon, D., Pantazis, C., Heslop, P., 1999. *Home sweet home?: the impact of poor housing on health*. Policy Press, University of Bristol, Bristol, UK.
- Matias-Guiu, J., Alvarez, J., Insa, R., Moltó, J.M., Martin, R., Codina, A., Martinez-Vazquez, J.M., 1990. Ischemic stroke in young adults. II. Analysis of risk factors in the etiological subgroups. *Acta Neurol. Scand.* 81, 314–317.
- Meade, M.S., Emch, M., 2010. *Medical Geography*. Guilford Press.
- Mendes, J., Bação, F., 1998. *A Análise de Clusters*.
- Mendes, R., 1988. The impact of occupation on workers' health: I - Morbidity. *Rev. Saúde Pública* 22, 311–326. doi:10.1590/S0034-89101988000400007
- Mennis, J., 2006. Mapping the Results of Geographically Weighted Regression. *Cartogr. J.* 43, 171–179. doi:10.1179/000870406X114658

- Monmonier, M.S., 1996. How to lie with maps. University of Chicago Press, Chicago.
- Monteiro, C.A., Moura, E.C., Conde, W.L., Popkin, B.M., 2004. Socioeconomic status and obesity in adult populations of developing countries: a review. *Bull. World Health Organ.* 82, 940–946.
- Moon, G., 2009. Health Geography, in: Thrift, R.K. (Ed.), *International Encyclopedia of Human Geography*. Elsevier, Oxford, pp. 35–45.
- Moran, P.A.P., 1950. Notes on Continuous Stochastic Phenomena. *Biometrika* 37, 17–23. doi:10.2307/2332142
- Müller-Nordhorn, J., Nolte, C.H., Rossnagel, K., Jungehülsing, G.J., Reich, A., Roll, S., Villringer, A., Willich, S.N., 2006. Knowledge About Risk Factors for Stroke A Population-Based Survey With 28 090 Participants. *Stroke* 37, 946–950. doi:10.1161/01.STR.0000209332.96513.82
- Nakaya, T., Fotheringham, A.S., Brunsdon, C., Charlton, M., 2005. Geographically weighted Poisson regression for disease association mapping. *Stat. Med.* 24, 2695–2717. doi:10.1002/sim.2129
- NCSS, 2014. NCSS - Statistical Software [WWW Document]. URL <http://www.ncss.com/> (acedido 10.6.14).
- Odoi, A., Busingye, D., 2014. Neighborhood geographic disparities in heart attack and stroke mortality: Comparison of global and local modeling approaches. *Spat. Spatio-Temporal Epidemiol.* 11, 109–123. doi:10.1016/j.sste.2014.10.001
- OGC, 2015. Open Geospatial Consortium [WWW Document]. URL <http://www.opengeospatial.org/> (acedido 3.3.15).
- ONS, 2014. Britain's mapping agency | Ordnance Survey [WWW Document]. URL <https://www.ordnancesurvey.co.uk/> (acedido 3.14.15).
- Openshaw, S., 1984. The modifiable areal unit problem. *Geo Abstracts University of East Anglia*.
- Osgood, D.W., 2000. Poisson-Based Regression Analysis of Aggregate Crime Rates. *J. Quant. Criminol.* 16, 21–43. doi:10.1023/A:1007521427059
- Palermo, E., 2013. What Is the Stroke Belt? [WWW Document]. LiveScience.com. URL <http://www.livescience.com/34465-what-is-the-stroke-belt.html> (acedido 4.24.15).
- Pedigo, A., Aldrich, T., Odoi, A., 2011. Neighborhood disparities in stroke and myocardial infarction mortality: a GIS and spatial scan statistics approach. *BMC Public Health* 11, 644. doi:10.1186/1471-2458-11-644

- Pfeiffer, D.U., Hugh-Jones, M., 2002. Geographical information systems as a tool in epidemiological assessment and wildlife disease management. *Rev. Sci. Tech. Int. Off. Epizoot.* 21, 91–102.
- Pfeiffer, D.U., Robinson, T.P., Stevenson, M., Stevens, K.B., Rogers, D.J., Clements, A.C.A., 2008. *Spatial Analysis in Epidemiology*. Also available as: Hardback | eBook.
- Piza, E.L., 2012. Using Poisson and negative binomial regression models to measure the influence of risk on crime incident counts. Rutgers Cent. O N Public Secur.
- Porta, M., 2008. *A Dictionary of Epidemiology*. Oxford University Press, USA.
- QGIS, 2014. Welcome to the QGIS project!
- R, 2014a. R: The R Project for Statistical Computing [WWW Document]. URL <http://www.r-project.org/> (acedido 6.24.14).
- R, 2014b. R: Fitting Generalized Linear Models.
- R, 2014c. R: Dispersion Test [WWW Document]. URL <http://artax.karlin.mff.cuni.cz/r-help/library/AER/html/dispersiontest.html> (acedido 10.18.14).
- R, 2014d. R: Generate All Combinations of n Elements, Taken m at a Time [WWW Document]. URL <https://stat.ethz.ch/R-manual/R-devel/library/utils/html/combn.html> (acedido 10.19.14).
- R, 2014e. R: Fit a Negative Binomial Generalized Linear Model [WWW Document]. URL <https://stat.ethz.ch/R-manual/R-devel/library/MASS/html/glm.nb.html> (acedido 11.2.14).
- R, 2014f. R: Akaike's An Information Criterion [WWW Document]. URL <http://stat.ethz.ch/R-manual/R-patched/library/stats/html/AIC.html> (acedido 11.10.14).
- Rainey, J.J., Omenah, D., Sumba, P.O., Moormann, A.M., Rochford, R., Wilson, M.L., 2007. Spatial clustering of endemic Burkitt's lymphoma in high-risk regions of Kenya. *Int. J. Cancer* 120, 121–127. doi:10.1002/ijc.22179
- Rogerson, P., Yamada, I., 2009. *Statistical detection and surveillance of geographic clusters*. CRC Press, Boca Raton.
- Rossum, C.T.M. van, Mheen, H. van de, Breteler, M.M.B., Grobbee, D.E., Mackenbach, J.P., 1999. Socioeconomic Differences in Stroke Among Dutch Elderly Women The Rotterdam Study. *Stroke* 30, 357–362. doi:10.1161/01.STR.30.2.357
- Rothman, K.J., Greenland, S., Lash, T.L., 2008. *Modern Epidemiology*. Lippincott Williams & Wilkins.

- RStudio, 2015. RStudio [WWW Document]. URL <http://www.rstudio.com/> (acedido 3.16.15).
- Rudd, A., Irwin, P., Penhale, B., 2004. Stroke: The “at Your Fingertips” Guide, 2nd edition. ed. Class Publishing, London.
- Sacco, R.L., Benjamin, E.J., Broderick, J.P., Dyken, M., Easton, J.D., Feinberg, W.M., Goldstein, L.B., Gorelick, P.B., Howard, G., Kittner, S.J., Manolio, T.A., Whisnant, J.P., Wolf, P.A., 1997. Risk Factors. *Stroke* 28, 1507–1517. doi:10.1161/01.STR.28.7.1507
- Salah, E.A.M., Turki, A.M., Al-Othman, E.M., 2012. Assessment of Water Quality of Euphrates River Using Cluster Analysis. *J. Environ. Prot.* 03, 1629–1633. doi:10.4236/jep.2012.312180
- Sá, M.J., 2009. AVC. primeira causa de morte em Portugal.
- Santana, P., 2005. Geografias da Saúde e do Desenvolvimento - Evolução e Tendências em Portugal. Ed Almedina, Coimbra.
- Sartorius, B.K., Sartorius, K., Chirwa, T.F., Fonn, S., 2011. Infant mortality in South Africa - distribution, associations and policy implications, 2007: an ecological spatial analysis. *Int. J. Health Geogr.* 10, 61. doi:10.1186/1476-072X-10-61
- Schieb, L.J., Mobley, L.R., George, M., Casper, M., 2013. Tracking Stroke Hospitalization Clusters Over Time and Associations With County-Level Socioeconomic and Healthcare Characteristics. *Stroke* 44, 146–152. doi:10.1161/STROKEAHA.112.669705
- Shaper, A.G., Phillips, A.N., Pocock, S.J., Walker, M., Macfarlane, P.W., 1991. Risk factors for stroke in middle aged British men. *BMJ* 302, 1111–1115.
- Shekhar, S., Xiong, H., 2008. Encyclopedia of GIS | Springer.
- Shinton, R., Beevers, G., 1989. Meta-analysis of relation between cigarette smoking and stroke. *BMJ* 298, 789–794.
- Siegel, J.S., 2012. The demography and epidemiology of human health and aging. Springer, Dordrecht ; New York.
- Silva, A.R. da, Rodrigues, T.C.V., 2013. Geographically Weighted Negative Binomial Regression—incorporating overdispersion. *Stat. Comput.* 24, 769–783. doi:10.1007/s11222-013-9401-9
- SNIG, 2015. Bem vindo ao SNIG! [WWW Document]. URL <http://snig.dgterritorio.pt/portal/> (acedido 3.14.15).
- Sokal, R.R., Rohlf, F.J., 1962. The Comparison of Dendrograms by Objective Methods. *Taxon* 11, 33–40. doi:10.2307/1217208

- Soljak, M., Samarasundera, E., Indulkar, T., Walford, H., Majeed, A., 2011. Variations in cardiovascular disease under-diagnosis in England: national cross-sectional spatial analysis. *BMC Cardiovasc. Disord.* 11, 12. doi:10.1186/1471-2261-11-12
- Sparks, P.J., Sparks, C.S., 2010. An application of spatially autoregressive models to the study of US county mortality rates. *Popul. Space Place* 16, 465–481. doi:10.1002/psp.564
- Sposato, L.A., Ioli, P., Povedano, G., Esnaola y Rojas, M.M., Saposnik, G., 2012. Unemployment: A Social Risk Factor Associated with Early Ischemic Stroke Mortality? Results from the Argentinean National Stroke Registry (ReNACer). *J. Stroke Cerebrovasc. Dis.* 21, 679–683. doi:10.1016/j.jstrokecerebrovasdis.2011.02.018
- Srinath Reddy, K., Shah, B., Varghese, C., Ramadoss, A., 2005. Responding to the threat of chronic diseases in India. *The Lancet* 366, 1744–1749. doi:10.1016/S0140-6736(05)67343-6
- Steenland, K., Hu, S., Walker, J., 2004. All-Cause and Cause-Specific Mortality by Socioeconomic Status Among Employed Persons in 27 US States, 1984–1997. *Am. J. Public Health* 94, 1037–1042.
- St-Hilaire, S., Mannel, S., Commendador, A., Mandal, R., Derryberry, D., 2010. Correlations between meteorological parameters and prostate cancer. *Int. J. Health Geogr.* 9, 19. doi:10.1186/1476-072X-9-19
- Takahashi, K., Kulldorff, M., Tango, T., Yih, K., 2008. A flexibly shaped space-time scan statistic for disease outbreak detection and monitoring. *Int. J. Health Geogr.* 7, 14. doi:10.1186/1476-072X-7-14
- Tobler, W.R., 1970. A Computer Movie Simulating Urban Growth in the Detroit Region. *Econ. Geogr.* 46, 234–240. doi:10.2307/143141
- Townsend, P., 1993. *The International Analysis of Poverty*. Harvester Wheatsheaf.
- Tsai, P.-J., Perng, C.-H., 2011. Spatial autocorrelation analysis of 13 leading malignant neoplasms in Taiwan: a comparison between the 1995-1998 and 2005-2008 periods. *Health (N. Y.)* 03, 712–731. doi:10.4236/health.2011.312120
- Tsutsumi, A., Kayaba, K., Ishikawa, S., 2011. Impact of occupational stress on stroke across occupational classes and genders. *Soc. Sci. Med.* 72, 1652–1658. doi:10.1016/j.socscimed.2011.03.026
- Vasconcelos, J., Freire, E., Morais, J., Machado, J.R., Santana, P., 2011. The health impacts of poor housing conditions and thermal discomfort. *Procedia Environ. Sci., Urban Environmental Pollution* 2010 4, 158–164. doi:10.1016/j.proenv.2011.03.019

- Waller, L.A., Gotway, C.A., 2004. Frontmatter, in: *Applied Spatial Statistics for Public Health Data*. John Wiley & Sons, Inc., pp. i–xviii.
- Ward, Jr, J.H., 1963. Hierarchical Grouping to Optimize an Objective Function. *J. Am. Stat. Assoc.* 58, 236–244.
- Wardle, J., Steptoe, A., 2003. Socioeconomic differences in attitudes and beliefs about healthy lifestyles. *J. Epidemiol. Community Health* 57, 440–443. doi:10.1136/jech.57.6.440
- Webster, R., Oliver, M.A., 2007. *Geostatistics for environmental scientists*, 2. ed. ed, *Statistics in practice*. Wiley, Chichester.
- Weisent, J., Rohrbach, B., Dunn, J.R., Odoi, A., 2012. Socioeconomic determinants of geographic disparities in campylobacteriosis risk: a comparison of global and local modeling approaches. *Int. J. Health Geogr.* 11, 45. doi:10.1186/1476-072X-11-45
- Welin, L., Svärdsudd, K., Wilhelmsen, L., Larsson, B., Tibblin, G., 1987. Analysis of risk factors for stroke in a cohort of men born in 1913. *N. Engl. J. Med.* 317, 521–526. doi:10.1056/NEJM198708273170901
- Westover, A.N., McBride, S., Haley, R.W., 2007. Stroke in young adults who abuse amphetamines or cocaine: a population-based study of hospitalized patients. *Arch. Gen. Psychiatry* 64, 495–502. doi:10.1001/archpsyc.64.4.495
- WHO, 2014a. STEPwise approach to stroke surveillance [WWW Document]. URL <http://www.who.int/chp/steps/stroke/en/> (acedido 12.21.14).
- WHO, 2014b. The determinants of health [WWW Document]. WHO. URL <http://www.who.int/hia/evidence/doh/en/> (acedido 4.23.15).
- WHO, 2013. Combined or multiple exposure to health stressors in indoor built environments [WWW Document]. URL <http://www.euro.who.int/en/health-topics/environment-and-health/Housing-and-health/publications/2014/combined-or-multiple-exposure-to-health-stressors-in-indoor-built-environments> (acedido 3.9.15).
- Windle, G.S., Burholt, V., Edwards, R.T., 2006. Housing related difficulties, housing tenure and variations in health status: evidence from older people in Wales. *Health Place* 12, 267–278. doi:10.1016/j.healthplace.2004.08.010
- Winston, J., Emch, M., 2013. *Medical geography*. Oxford University Press, New York.
- Wu, S.H., Woo, J., Zhang, X.-H., 2013. Worldwide socioeconomic status and stroke mortality: an ecological study. *Int. J. Equity Health* 12, 42. doi:10.1186/1475-9276-12-42

- Yang, H., Schnall, P.L., Jauregui, M., Su, T.-C., Baker, D., 2006. Work hours and self-reported hypertension among working people in California. *Hypertension* 48, 744–750. doi:10.1161/01.HYP.0000238327.41911.52
- Zucchini, W., 2003. APPLIED SMOOTHING TECHNIQUES Part 1: Kernel Density Estimation.

6. Anexos

	Indicador	Vantagens	Limitações
Educação	Nº de anos de frequência do sistema de ensino a tempo inteiro Idade de saída do sistema de ensino a tempo inteiro	Pode ser facilmente recolhido, facilmente quantificável; não é condicionado por más condições de saúde em adulto; marcador das condições nos primeiros tempos de vida	Culturalmente específico; menor aplicabilidade a determinados grupos ou coortes (por exemplo a relação entre educação e rendimento é mais fraca nas mulheres idosas)
Ocupação profissional	A ocupação profissional é geralmente expressa em categorias de classe social ocupacional. Exemplos: Registrar General Great Britain social class; United States standard occupational classification system; ISCO-88 international standard classification of occupations.	Pode ser indicador do estatuto social, poder, rendimento e nível de educação	Subjectividade dos sistemas de classificação; inconsistência dos sistemas de classificação; limitações na demonstração do estatuto socioeconómico em determinados subgrupos (por exemplo em mulheres e em algumas minorias étnicas); dificuldade na incorporação de novas profissões, de pessoas que não trabalham, de idosos; pode ser condicionado por más condições de saúde em adulto
Rendimento	Rendimento total monetário (por mês ou ano) Rendimento disponível Individual ou do agregado familiar Benefícios não monetários	Boa avaliação da capacidade aquisitiva de bens, educação, cuidados de saúde	Recolha difícil; grande número de não-respondentes, problemas de confidencialidade; pode ser condicionado por más condições de saúde em adulto; pode variar ao longo da vida; membros do agregado familiar podem ter acesso desigual
Posse de bens materiais	Propriedade de habitação, de carro, de electrodomésticos	Elevada correlação com o rendimento e a educação; indicador do nível de vida	Pode ser indicador de diferenças de estilos de vida
Medidas de base espacial (caracterização de áreas)	Compostos de múltiplas variáveis para caracterizar uma área geográfica. As variáveis usadas podem variar e são normalmente retiradas de Censos. São exemplos os índices de privação de Carstairs (desemprego, propriedade de habitação sem carro, sobrelotação da habitação, classe social do chefe de família) e Townsend (desemprego, não possuir carro, não ser proprietário da habitação).	Revela aspectos importantes da área de residência que não seria possível obter de forma individual	Risco de “Falácia ecológica”- as associações entre variáveis ao nível ecológico podem não ser verdadeiras ao nível individual.

Anexo 1 – Tipologia de indicadores usados na medição do estatuto socioeconómico. Adaptado de Cox et al. (2006).

	Homens a viver em áreas predominantemente rurais (%)	Homens que trabalhavam mais de 45 horas por semana (%)	Trabalhadores manuais e agrícolas (%)	Poder de compra municipal (%)	Infraestruturas de saúde municipais (n°)	Valor médio mensal das rendas dos alojamentos (euros)	Distância ao hospital mais próximo (minutos)	Taxa de desemprego - Homens (%)	Homens que atingiram o ensino básico (%)	índice municipal de poder de compra per capita
Homens a viver em áreas predominantemente rurais (%)	1									
Homens que trabalhavam mais de 45 horas por semana (%)	-.085	1								
	.159									
Trabalhadores manuais e agrícolas (%)	.465**	.008	1							
	.000	.889								
Poder de compra municipal (%)	-.282*	.066	-.464**	1						
	.000	.276	.000							
Infraestruturas de saúde municipais (n°)	-.189**	-.029	-.410**	.760**	1					
	.002	.636	.000	.000						
Valor médio mensal das rendas dos alojamentos (euros)	-.211**	.204**	-.247**	.087	.131*	1				
	.000	.001	.000	.149	.028					
Distância ao hospital mais próximo (minutos))	.541**	.082	.460**	-.258**	-.362**	-.242**	1			
	.000	.171	.000	.000	.000	.000				
Taxa de desemprego - Homens (%)	-.096	-.087	-.348**	.251**	.126*	-.168**	-.087	1		
	.109	.148	.000	.000	.036	.005	.148			
Homens que atingiram o ensino básico (%)	-.612**	.022	-.437**	.100	.123*	.281**	-.592**	-.009	1	
	.000	.718	.000	.098	.040	.000	.000	.887		
índice municipal de poder de compra per capita	-.512**	.022	-.802**	.775**	.645**	.246**	-.467**	.240**	.397**	1
	.000	.718	.000	.000	.000	.000	.000	.000	.000	

** Correlação significativa (p<0.01).

* Correlação significativa (p<0.05).

Anexo 2 - Coeficientes de correlação de Pearson entre pares de variáveis independentes investigadas no período 2002-2006 (modelo masculino).

	Mulheres que trabalhavam mais de 45 horas por semana (%)	Trabalhadores manuais e agrícolas (%)	Poder de compra municipal (%)	Infraestruturas de saúde municipais (n°)	Valor médio mensal das rendas dos alojamentos (euros)	Distância ao hospital mais próximo (minutos)	Taxa de desemprego - Mulheres (%)	Mulheres a viver em áreas predominantemente rurais (%)	Mulheres que atingiram o ensino básico (%)	índice municipal de poder de compra per capita
Mulheres que trabalhavam mais de 45 horas por semana (%)	1									
Trabalhadores manuais e agrícolas (%)	,123	1								
	,040									
Poder de compra municipal (%)	,051	-,464**	1							
	,401	,000								
Infraestruturas de saúde municipais (n°)	-,093	-,410**	,760**	1						
	,121	,000	,000							
Valor médio mensal das rendas dos alojamentos (euros)	,029	-,247**	,087	,131	1					
	,635	,000	,149	,028						
Distância ao hospital mais próximo (minutos)	,267**	,460**	-,258**	-,362**	-,242**	1				
	,000	,000	,000	,000	,000					
Taxa de desemprego - Mulheres (%)	,210**	,219**	-,150	-,236**	-,448**	,403**	1			
	,000	,000	,012	,000	,000	,000				
Mulheres a viver em áreas predominantemente rurais (%)	,121	,470**	-,281**	-,192**	-,214**	,546**	,264**	1		
	,044	,000	,000	,001	,000	,000	,000			
Mulheres que atingiram o ensino básico (%)	-,120	-,482**	,214**	,145	,281**	-,535**	-,293**	-,665**	1	
	,045	,000	,000	,016	,000	,000	,000	,000		
índice municipal de poder de compra per capita	-,112	-,802**	,775**	,645**	,246**	-,467**	-,339**	-,515**	,485**	1
	,063	,000	,000	,000	,000	,000	,000	,000	,000	

** Correlação significativa (p<0.01).

* Correlação significativa (p<0.05).

Anexo 3 - Coeficientes de correlação de Pearson entre pares de variáveis independentes investigadas no período 2002-2006 (modelo feminino).

	Homens que trabalhavam mais de 45 horas por semana (%)	Taxa de desemprego - Homens (%)	Poder de compra municipal (%)	índice municipal de poder de compra per capita	Homens que atingiram o ensino básico (%)	Trabalhadores manuais e agrícolas (%)
Homens que trabalhavam mais de 45 horas por semana (%)	1					
Taxa de desemprego - Homens (%)	-,167**	1				
	,006					
Poder de compra municipal (%)	-,214**	,137*	1			
	,000	,023				
índice municipal de poder de compra per capita	-,440**	,120*	,650**	1		
	,000	,047	,000			
Homens que atingiram o ensino básico (%)	-,496**	,196**	,254**	,719**	1	
	,000	,001	,000	,000		
Trabalhadores manuais e agrícolas (%)	,424**	-,258**	-,452**	-,835**	-,851**	1
	,000	,000	,000	,000	,000	

** Correlação significativa (p<0.01).

* Correlação significativa (p<0.05).

Anexo 4 - Coeficientes de correlação de Pearson entre pares de variáveis independentes investigadas no período 1992 -1996 (modelo masculino).

	Mulheres que trabalhavam mais de 45 horas por semana (%)	Taxa de desemprego - Mulheres (%)	Poder de compra municipal (%)	índice municipal de poder de compra per capita	Mulheres que atingiram o ensino básico (%)	Trabalhadores manuais e agrícolas (%)
Mulheres que trabalhavam mais de 45 horas por semana (%)	1					
Taxa de desemprego - Mulheres (%)	-,115	1				
	,056					
Poder de compra municipal (%)	-,191**	-,081	1			
	,001	,182				
índice municipal de poder de compra per capita	-,459**	-,109	,650**	1		
	,000	,070	,000			
Mulheres que atingiram o ensino básico (%)	-,479**	,151*	,277**	,736**	1	
	,000	,012	,000	,000		
Trabalhadores manuais e agrícolas (%)	,433**	-,066	-,452**	-,835**	-,849**	1
	,000	,276	,000	,000	,000	

** Correlação significativa (p<0.01).

* Correlação significativa (p<0.05).

Anexo 5 - Coeficientes de correlação de Pearson entre pares de variáveis independentes investigadas no período 1992 -1996 (modelo feminino).

Homens com menos de 65 anos de idade (1992-1996)

Categoria	Variável	Estimativa (1992-1996) ¹	SE ² (1992-1996)	P-Value (1992-1996)	AIC ³ (1992-1996)
Estatuto Soci	Homens que trabalhavam mais de 45 horas por semana (%)	0.0061 (0.0032, 0.0090)	0.0015	0.0001	1792
Emprego	Taxa de desemprego - Homens (%)	0.0277 (0.0121, 0.0432)	0.0079	0.0005	1797
Nível de escolaridade	Homens que atingiram o ensino básico (%)	-0.0071 (-0.0133, -0.0009)	0.0032	0.0245	1804
Rendimento	Poder de compra municipal (%)	0.0164 (0.0116, 0.0212)	0.0025	0.0000	1767

Mulheres com menos de 65 anos de idade (1992-1996)

Categoria	Variável	Estimativa (1992-1996) ¹	SE ² (1992-1996)	P-Value (1992-1996)	AIC ³ (1992-1996)
Estatuto Soci	Mulheres que trabalhavam mais de 45 horas por semana (%)	0.0122 (0.0076, 0.0167)	0.0023	0.0001	1421
Emprego	Taxa de desemprego - Mulheres (%)	0.0031 (-0.0178, 0.0239)	0.0106	0.7670	1447
Nível de escolaridade	Mulheres que atingiram o ensino básico (%)	-0.0269 (-0.0355, -0.0184)	0.0044	0.0001	1409
Rendimento	Poder de compra municipal (%)	0.0082 (0.0016, 0.0146)	0.0033	0.0134	1441

Homens com menos de 65 anos de idade (2002-2006)

Categoria	Variável	Estimativa (2002-2006) ¹	SE ² (2002-2006)	P-Value (2002-2006)	AIC ³ (2002-2006)
Estatuto Soci	Homens que trabalhavam mais de 45 horas por semana (%)	0.0012 (-0.0060, 0.0083)	0.0037	0.3260	1597
	Trabalhadores manuais e agrícolas (%)	-0.0006 (-0.0030, 0.0019)	0.0012	0.6460	1597
Emprego	Taxa de desemprego - Homens (%)	0.0320 (0.0135, 0.0504)	0.0094	0.0007	1586
Nível de escolaridade	Homens que atingiram o ensino básico (%)	-0.0783 (-0.0984, -0.0582)	0.0103	0.0001	1540
Rendimento	Poder de compra municipal (%)	0.0051 (-0.0039, 0.0138)	0.0045	0.2570	1596
	Valor médio mensal das rendas dos alojamentos (euros)	-0.0030 (-0.0041, -0.0019)	0.0006	0.0001	1568
Ambiente	Distância ao hospital mais próximo (minutos)	0.0059 (0.0036, 0.0082)	0.0012	0.0001	1573

Mulheres com menos de 65 anos de idade (2002-2006)

Categoria	Variável	Estimativa (2002-2006) ¹	SE ² (2002-2006)	P-Value (2002-2006)	AIC ³ (2002-2006)
Estatuto Soci	Mulheres que trabalhavam mais de 45 horas por semana (%)	0.0373 (0.0205, 0.0539)	0.0085	0.0001	1242
	Trabalhadores manuais e agrícolas (%)	-0.0014 (-0.0047, 0.0019)	0.0017	0.4080	1260
Emprego	Taxa de desemprego - Mulheres (%)	0.0352 (0.0225, 0.0475)	0.0064	0.0001	1233
Nível de escolaridade	Mulheres que atingiram o ensino básico (%)	-0.0768 (-0.1050, -0.0484)	0.0144	0.0001	1233
Rendimento	Poder de compra municipal (%)	0.0136 (0.0021, 0.0246)	0.0057	0.0178	1256
	Valor médio mensal das rendas dos alojamentos (euros)	-0.0022 (-0.0037, -0.0007)	0.0008	0.0041	1253
Ambiente	Distância ao hospital mais próximo (minutos)	0.0071 (0.0039, 0.0103)	0.0016	0.0001	1242

¹Intervalo de Confiança 95% ²Standard Error ³Critério de Informação de Akaike

Anexo 6 - Associações univariadas entre risco de AVC e variáveis independentes seleccionadas

Variáveis causais	Pseudo p-value	I de Moran (Global)
Alojamentos superlotados (%)	0,0001	0,6545
População residente a viver em edificios construidos até 1960 (%)	0,0001	0,3958
Alojamentos sem aquecimento central (%)	0,0001	0,5705
População residente sem qualquer tipo de aquecimento (%)	0,0001	0,6646
Peso da população residente arrendatária face à população residente proprietária (%)	0,0001	0,6616
População residente a viver em apartamentos (%)	0,0001	0,7393
População Residente em alojamentos familiares não clássicos (%)	0,0002	0,0454
Indicadores de risco mortalidade AVC (SMRi) por coorte	Pseudo p-value	I de Moran (Global)
Mulheres mais de 64 anos	0,0001	0,6681
Mulheres todas as idades	0,0001	0,6679
Homens mais de 64 anos	0,0001	0,6547
Homens todas as idades	0,0001	0,6716
Homens e Mulheres mais de 64 anos	0,0001	0,6119
Homens e Mulheres todas as idades	0,0001	0,61101

Anexo 7 – Índices I de Moran e níveis significância, para todas as variáveis.

	Alojamentos superlotados (%)	População residente a viver em edifícios construídos até 1960 (%)	Alojamentos sem aquecimento central (%)	População Residente sem qualquer tipo de aquecimento (%)	Peso da população residente arrendatária face à população residente proprietária (%)	População residente a viver em apartamentos (%)	População Residente em alojamentos familiares não clássicos (%)
Alojamentos superlotados (%)	1						
População residente a viver em edifícios construídos até 1960 (%)	-.010	1					
Alojamentos sem aquecimento central (%)	.166**	.305**	1				
População Residente sem qualquer tipo de aquecimento (%)	.416**	-.046**	.079**	1			
Peso da população residente arrendatária face à população residente proprietária (%)	.267**	.273**	.087**	.409**	1		
População residente a viver em apartamentos (%)	.232**	-.049**	.005	.486**	.595**	1	
População Residente em alojamentos familiares não clássicos (%)	.015	-.028	-.015	.105**	.102**	.142**	1
	.339	.070	.334	.000	.000	.000	

** Correlação significativa ($p < 0.01$)

Anexo 8 - Coeficientes de correlação de Pearson entre pares de variáveis independentes.