

FACULDADE DE ENGENHARIA DA UNIVERSIDADE DO PORTO

**Fracturas Osteoporóticas do Colo do Fémur em Portugal
e seus Determinantes Socioeconómicos**

Joana Maria Faria Portela Moreira

Licenciada em Matemática
pela Faculdade de Ciências da Universidade do Porto

Dissertação submetida para satisfação pessoal dos
requisitos do grau de Mestre em
Engenharia Biomédica

Dissertação realizada sobre a supervisão da Professora Doutora
Maria de Fátima de Pina, Professora Associada da Faculdade de Medicina da
Universidade do Porto e Investigadora do Instituto de Engenharia Biomédica –
INEB, da Universidade do Porto

Porto, Março de 2008



FACULDADE DE ENGENHARIA DA UNIVERSIDADE DO PORTO

Fracturas Osteoporóticas do Colo do Fémur em Portugal e seus Determinantes Socioeconómicos

Joana Maria Faria Portela Moreira

Março de 2008

ÍNDICE GERAL

RESUMO	I
ABSTRACT	II
AGRADECIMENTOS	III
ÍNDICE DE TEXTO	IV
ÍNDICE DE FIGURAS	VII
ÍNDICE DE TABELAS	X
ABREVIATURAS	XII
INTRODUÇÃO	1
ESTADO DA ARTE	6
MATERIAIS E MÉTODOS	74
RESULTADOS	93
DISCUSSÃO	114
REFERÊNCIAS	129
ANEXOS	135

RESUMO

A osteoporose é uma doença progressiva e uma das doenças musculoesqueléticas mais graves. É considerada mundialmente como um problema de saúde pública que afecta um grande número de pessoas, em especial a população a partir dos 50 anos de idade e, maioritariamente, mulheres. Por ser uma doença silenciosa, o primeiro sinal geralmente é a ocorrência de fracturas, quando a doença já está instalada. Perante isto, é muito importante reconhecer quais os factores que contribuem para o aumento do risco de sofrer fracturas, de forma a poderem ser tomadas as devidas medidas preventivas.

O padrão da ocorrência da osteoporose numa população pode estar relacionado com vários factores, incluindo factores socioeconómicos, como o rendimento, o nível de escolaridade, etc., e factores relacionados com o estilo de vida, como a dieta, a actividade física e a sua diversificação. Estes e outros aspectos têm sido cada vez mais identificados como determinantes importantes de várias doenças, incluindo a osteoporose.

O objectivo deste trabalho é explicar as variações geográficas da incidência das fracturas osteoporóticas do colo do fémur, encontradas em Portugal Continental, através do estudo da associação destas com factores socioeconómicos. Para tal foi desenvolvido um estudo ecológico que relaciona a incidência destas fracturas com diversos indicadores socioeconómicos, utilizando como ferramentas de análise os Sistemas de Informação Geográfica – SIG, a Análise de Componentes Principais e Análise de *Clusters*. Foram utilizadas dados sócio-demográficos do Censo de 2001 do Instituto Nacional de Estatística – INE, dados sobre a mortalidade infantil do início e meados do século XX, também do INE e dados relativos à incidência das fracturas do colo do fémur, em indivíduos com 50 anos ou mais, (2000 a 2002) do Instituto de Gestão Informática e Financeira da Saúde (IGIF). As variáveis do Censo de 2001 foram utilizadas para cálculo de alguns indicadores socioeconómicos que foram em seguida utilizados na análise de componentes principais. Foram encontradas quatro componentes e com base nessas componentes foi realizada a análise de *clusters*. Em seguida foram calculadas as taxas de incidência cumulativa (2000 a 2002) das fracturas osteoporóticas de colo do fémur, para cada *cluster*.

Os resultados mostram que o norte e o sul do país apresentam um estatuto socioeconómico distinto e que, principalmente para as mulheres, os factores socioeconómicos podem estar associados às taxas de incidência de fracturas do colo do fémur. Os resultados mostram ainda, evidências de que as condições de vida há 70 anos atrás podem estar relacionadas com a actual incidência de fracturas do colo de fémur.

Em conclusão os factores socioeconómicos podem explicar, em parte, as diferenças geográficas encontradas na incidência das fracturas do colo do fémur.

ABSTRACT

The osteoporosis is a progressive disease and one of the most serious musculoskeletal diseases. It is considered worldwide as a public health problem that affects a great number of persons, mainly over 50 year old, and specially women. Being a silent disease, the first sign when the disease is already installed is, usually, the occurrence of fractures. Consequently it is important to identify which factors contribute to an increase risk of suffering fractures, in order to plan prevention actions.

The pattern of occurrence of osteoporosis in the population can be related with several factors, including socioeconomic factors, such as income, education level, etc and factors related to life style such as diet, physical activity and its variation. These and other aspects have been identified as important determinants of several diseases including osteoporosis.

The aim of this work is to explain the geographic variations of the osteoporotic hip fractures incidence, in Continental Portugal and its association with socioeconomic factors. To achieve the objective an ecologic study has been conducted, correlating the incidence of those fractures with several socioeconomic indicators using as analytical tools the Geographic Information Systems – GIS, Principals Components Analysis and Clusters Analysis. The data used consist of: socio-demographic data - Census 2001, from the National Institute of Statistics – INE, infant mortality data from the beginning and mid XX century also from INE, and hip fracture incidence in persons with 50 years old or more, (2000 to 2002) from the Health Informatics and Financial Management Institute (Instituto de Gestão Informática e Financeira da Saúde – IGIF). The variables of the 2001 Census were used for construction of some socioeconomic indicators that were then used on principal component analysis. Four components have been found and based on them the clusters analysis has been realized. the incidence rates of osteoporotic hip fracture, for each cluster, has been calculated.

The results show that the north and south of the country presents a different socioeconomic status and, mainly among women, the socioeconomic status can be related to the incidence rates of osteoporotic hip fractures. The results show also that the life conditions of 70 years ago can be related to the current hip fracture incidence.

In conclusion the socioeconomic factors can explain, in part, the geographic differences found in the hip fracture incidence.

AGRADECIMENTOS

Na apresentação do presente trabalho exprimo o meu profundo agradecimento a todos os familiares e amigos que me incentivaram e apoiaram na sua realização.

De uma forma muito especial quero apresentar um obrigado muito grande à minha mãe e ao meu pai, que muito me apoiaram para que este trabalho se tornasse realidade.

Em particular manifesto um vivo e profundo agradecimento à Professora Doutora Maria de Fátima Pina pela força, dedicação e orientação dadas ao desenvolvimento deste trabalho, e pelo seu contributo no meu enriquecimento científico e profissional.

Os meus sinceros agradecimentos aos colegas do Instituto de Engenharia Biomédica que sempre estiveram presentes e me ajudaram no desenvolvimento deste trabalho, nomeadamente à Andreia Olheiro e Sandra Alves.

Ao Professor Doutor Fernando Jorge Monteiro deixo aqui o meu agradecimento pelo constante apoio e interesse demonstrados neste trabalho.

ÍNDICE DE TEXTO

1. INTRODUÇÃO	1
1.1 ENQUADRAMENTO	1
1.2 OBJECTIVOS	4
1.3 ESTRUTURA	5
2. ESTADO DA ARTE	6
2.1 EPIDEMIOLOGIA	7
2.1.1 Epidemiologia social	9
2.2 OSTEOPOROSE	12
2.3 EPIDEMIOLOGIA DA OSTEOPOROSE	21
2.3.1 Epidemiologia das Fracturas Vertebrais	26
2.3.2 Epidemiologia das Fracturas do Punho	28
2.3.3 Epidemiologia das Fracturas do Colo do Fémur	26
2.5 SISTEMAS DE INFORMAÇÃO GEOGRÁFICA NA SAÚDE	40
2.5 ANÁLISE DE COMPONENTES PRINCIPAIS	44
2.5.1 Matriz de Correlações	47
2.5.2 Testes para Validação da Análise	48
2.5.3 Extracção das Componentes Principais	49
2.5.4 Selecção do Número de Componentes Principais	54
2.5.5 Métodos de Rotação	57
2.6 ANÁLISE DE <i>CLUSTERS</i>	61
2.6.1 Selecção das Variáveis	63
2.6.2 Medidas de Semelhança	64
2.6.3 Métodos de Aglomeração	66
2.6.4 Selecção do Número de <i>Clusters</i>	71
3. MATERIAIS E MÉTODOS	74
3.1 MATERIAIS	75
3.1.1 Bases de Dados	75
3.1.2 Mapas	79
3.1.3 Software	79

3.2 MÉTODOS	80
3.2.1 Cálculo dos Indicadores demográficos	80
3.2.2 Cálculo da Mortalidade Infantil	86
3.2.3 Análise	87
4. RESULTADOS	93
4.1 RESULTADOS DA ANÁLISE DOS DADOS DEMOGRÁFICOS	94
4.1.1 Análise de Componentes Principais	94
4.1.2 Análise de <i>Clusters</i>	97
4.2 RESULTADOS DA ANÁLISE DOS DADOS DA MORTALIDADE INFANTIL	107
5. DISCUSSÃO	114
5.1 DISCUSSÃO	115
5.2 ANÁLISE DEMOGRÁFICA <i>VERSUS</i> FRACTURAS OSTEOPORÓTICAS DO COLO DO FÉMUR	117
5.3 MORTALIDADE INFANTIL <i>VERSUS</i> FRACTURAS OSTEOPORÓTICAS DO COLO DO FÉMUR	125
5.4 CONCLUSÃO	127
REFERÊNCIAS	129
ANEXO A	135
ANEXO B	140
ANEXO C	144
ANEXO D	151
ANEXO E	155

ÍNDICE DE FIGURAS

2. ESTADO DA ARTE

FIGURA 2.1 VÉRTEBRA ONDE É POSSÍVEL OBSERVAR AS DUAS PARTES DISTINTAS DO OSSO.	13
FIGURA 2.2 (A) OSSO SAUDÁVEL; (B) E (C) OSSO COM OSTEOPOROSE.	15
FIGURA 2.3 DISTRIBUIÇÃO DA DMO EM JOVENS MULHERES SAUDÁVEIS COM IDADES ENTRE OS 30 E 40 ANOS.	18
FIGURA 2.4 (A) FRACTURA DO COLO DO FÉMUR; (B) FRACTURA DAS VÉRTEBRAS; (C) FRACTURA DO PUNHO.	22
FIGURA 2.5 (A) VÉRTEBRA NORMAL; VÉRTEBRAS COM DEFORMIDADES: (B) VÉRTEBRA BICÔNCAVA OU EM BOCA DE PEIXE, (C) VÉRTEBRA EM CUNHA E (D) VÉRTEBRA POR ESMAGAMENTO OU EM BOLACHA.	23
FIGURA 2.6 INCIDÊNCIA CUMULATIVA, DURANTE UM ANO, DE FRACTURAS VERTEBRAIS EM MULHERES POR 100.000, NOS EUA, NO REINO UNIDO E PAÍSES BAIXOS.	24
FIGURA 2.7 INCIDÊNCIA CUMULATIVA, DURANTE UM ANO, DE FRACTURAS VERTEBRAIS EM HOMENS POR 100.000, NOS EUA, NO REINO UNIDO E PAÍSES BAIXOS.	25
FIGURA 2.8 INCIDÊNCIA CUMULATIVA, DURANTE UM ANO, DE FRACTURAS DO PUNHO EM HOMENS E MULHERES POR 100.000 PESSOAS-ANOS, NOS EUA E NO REINO UNIDO.	27
FIGURA 2.9 MAPA DAS TAXAS DE INCIDÊNCIA DE FRACTURAS DO PUNHO (POR 1000 HABITANTES) DOS EUA ENTRE 1986-1990, DE INDIVÍDUOS COM 65-89 ANOS.	27
FIGURA 2.10 FRACTURAS OSTEOPORÓTICAS DO COLO DO FÉMUR (ACIMA 50 ANOS DE IDADE), EM PORTUGAL NO PERÍODO DE 2000 A 2002.	28
FIGURA 2.11 INCIDÊNCIA CUMULATIVA, DURANTE UM ANO, DE FRACTURAS DO COLO DO FÉMUR EM MULHERES POR 100.000 HABITANTES EM VÁRIOS PAÍSES.	30
FIGURA 2.12 INCIDÊNCIA CUMULATIVA, DURANTE UM ANO, DE FRACTURAS DO COLO DO FÉMUR EM HOMENS POR 100.000 HABITANTES EM VÁRIOS PAÍSES.	30
FIGURA 2.13 MAPA DAS TAXAS DE INCIDÊNCIA DE FRACTURAS DO COLO DO FÉMUR (POR 1000 HABITANTES) NOS EUA ENTRE 1986-1990, DE INDIVÍDUOS COM 65-89 ANOS.	31
FIGURA 2.14 DISTRIBUIÇÃO GEOGRÁFICA DAS TAXAS DE INCIDÊNCIA DE FRACTURAS DO FÉMUR	35

FIGURA 2.15 AUMENTO EXPONENCIAL DA INCIDÊNCIA DA FRACTURAS DO COLO DO FÉMUR COM A IDADE, EM AMBOS OS GÉNEROS, NO CENTRO MEDOS DO PORTO.	36
FIGURA 2.16 OCUPAÇÃO DE CAMAS HOSPITALARES EM MULHERES, COM IDADE SUPERIOR A 45 ANOS, DE ACORDO COM O DIAGNÓSTICO NA REGIÃO DE TREND EM INGLATERRA.	38
FIGURA 2.17 EIXOS PRINCIPAIS COM TRÊS VARIÁVEIS.	47
FIGURA 2.18 A PROECÇÃO ORTOGONAL DE PONTOS SOBRE UM SUBESPAÇO.	53
FIGURA 2.19 SELECÇÃO DO NÚMERO DE COMPONENTES PRINCIPAIS PELO MÉTODO SCREE TEST.	55
FIGURA 2.20 DIAGRAMA DE DISPERSÃO ENTRE DUAS COMPONENTES PRINCIPAIS, ONDE É POSSÍVEL OBSERVAR TRÊS <i>CLUSTERS</i>	62
FIGURA 2.21 REPRESENTAÇÃO GRÁFICA DO NÚMERO DE <i>CLUSTERS</i> VERSUS COEFICIENTE DE FUSÃO.	72

4. RESULTADOS

FIGURA 4.1 REPRESENTAÇÃO DAS MÉDIAS DA COMPONENTE 1 PARA CADA <i>CLUSTER</i>	98
FIGURA 4.2 REPRESENTAÇÃO DAS MÉDIAS DA COMPONENTE 2 PARA CADA <i>CLUSTER</i>	99
FIGURA 4.3 REPRESENTAÇÃO DAS MÉDIAS DA COMPONENTE 3 PARA CADA <i>CLUSTER</i>	99
FIGURA 4.4 REPRESENTAÇÃO DAS MÉDIAS DA COMPONENTE 4 PARA CADA <i>CLUSTER</i>	100
FIGURA 4.5 MAPA DOS <i>CLUSTERS</i> DA ANÁLISE DOS DADOS DEMOGRÁFICOS.	103
FIGURA 4.6 REPRESENTAÇÃO DAS TAXAS DE INCIDÊNCIA DAS FRACTURAS OSTEOPORÓTICAS DO COLO DO FÉMUR PARA HOMENS E INTERVALOS DE CONFIANÇA PARA CADA <i>CLUSTER</i>	105
FIGURA 4.7 REPRESENTAÇÃO DAS TAXAS DE INCIDÊNCIA DAS FRACTURAS OSTEOPORÓTICAS DO COLO DO FÉMUR PARA MULHERES E INTERVALOS DE CONFIANÇA PARA CADA <i>CLUSTER</i> . ..	105
FIGURA 4.8 REPRESENTAÇÃO DAS TAXAS DE INCIDÊNCIA DAS FRACTURAS OSTEOPORÓTICAS DO COLO DO FÉMUR PARA AMBOS OS GÉNEROS E INTERVALOS DE CONFIANÇA PARA CADA <i>CLUSTER</i>	106
FIGURA 4.9 REPRESENTAÇÃO DAS CORRELAÇÕES ENTRE A MORTALIDADE INFANTIL EM CADA ANO, POR DISTRITO, E A TAXA DE INCIDÊNCIA DAS FRACTURAS DO COLO DO FÉMUR, TAMBÉM POR DISTRITO, AO LONGO DOS ANOS.	109
FIGURA 4.10 REPRESENTAÇÃO DAS CORRELAÇÕES ENTRE A MORTALIDADE INFANTIL POR ANO NAS DÉCADAS DE 40 E 50, E A COMPONENTE PRINCIPAL 1.	110

FIGURA 4.11 REPRESENTAÇÃO DAS CORRELAÇÕES ENTRE A MORTALIDADE INFANTIL POR ANO NAS DÉCADAS DE 40 E 50, E A COMPONENTE PRINCIPAL 2.	110
FIGURA 4.12 REPRESENTAÇÃO DAS CORRELAÇÕES ENTRE A MORTALIDADE INFANTIL POR ANO NAS DÉCADAS DE 40 E 50, E A COMPONENTE PRINCIPAL 3.	111
FIGURA 4.13 REPRESENTAÇÃO DAS CORRELAÇÕES ENTRE A MORTALIDADE INFANTIL POR ANO NAS DÉCADAS DE 40 E 50, E A COMPONENTE PRINCIPAL 4.	112
FIGURA 4.14 REPRESENTAÇÃO DAS CORRELAÇÕES ENTRE A MORTALIDADE INFANTIL EM CADA ANO, POR DISTRITO, E A TAXA DE INCIDÊNCIA DAS FRACTURAS DO COLO DO FÉMUR, TAMBÉM POR DISTRITO, AO LONGO DOS ANOS, SEM OS DISTRITOS DE PORTO E LISBOA.	113

ÍNDICE DE TABELAS

2. ESTADO DA ARTE

TABELA 2.1 DEFINIÇÕES DA OMS BASEADAS NA DMO.	17
TABELA 2.2 FACTORES DE RISCO ASSOCIADOS A FRACTURAS DO COLO DO FÉMUR, PARA A POPULAÇÃO FEMININA CAUCASIANA, COM MAIS DE 65 ANOS.	29
TABELA 2.3 INCIDÊNCIA POR 100.000 HABITANTES, DAS FRACTURAS DO COLO DO FÉMUR EM POPULAÇÃO COM 50 ANOS OU MAIS DA EUROPA MERIDIONAL (MEDOS).	31
TABELA 2.4 CLASSIFICAÇÃO DA ANÁLISE FACTORIAL, SEGUNDO KAISER, SEGUNDO OS VALORES DO KMO.	48

3. MATERIAIS E MÉTODOS

TABELA 3.1 DICIONÁRIO DE VARIÁVEIS.	83
TABELA 3.2 TABELA DAS VARIÁVEIS ESCOLHIDAS PARA REALIZAR A ANÁLISE.	88
TABELA 3.3 RESULTADO DO TESTE DE BARTLETT E KMO.	88
TABELA 3.4 TOTAL DA VARIÂNCIA EXPLICADA.	89
TABELA 3.5 TOTAL DE CORRELAÇÕES DAS COMPONENTES PRINCIPAIS.	90
TABELA 3.6 MATRIZ DAS COMPONENTES ROTACIONADA.	90

4. RESULTADOS

TABELA 4.1 TABELA DAS VARIÁVEIS ESCOLHIDAS PARA REALIZAR A ANÁLISE, COM AS RESPECTIVAS MÉDIAS E DESVIOS PADRÃO.	94
TABELA 4.2 CORRELAÇÕES E SIGNIFICÂNCIAS RESULTANTES DAS REGRESSÕES ENTRE AS COMPONENTES PRINCIPAIS E A TAXA DE INCIDÊNCIA DE FRACTURAS OSTEOPORÓTICAS DO COLO DO FÉMUR PARA HOMENS.	96
TABELA 4.3 CORRELAÇÕES E SIGNIFICÂNCIAS RESULTANTES DAS REGRESSÕES ENTRE AS COMPONENTES PRINCIPAIS E A TAXA DE INCIDÊNCIA DE FRACTURAS OSTEOPORÓTICAS DO COLO DO FÉMUR PARA MULHERES.	96
TABELA 4.4 CORRELAÇÕES E SIGNIFICÂNCIAS RESULTANTES DAS REGRESSÕES ENTRE AS COMPONENTES PRINCIPAIS E A TAXA DE INCIDÊNCIA DE FRACTURAS OSTEOPORÓTICAS DO COLO DO FÉMUR PARA HOMENS E MULHERES.	96
TABELA 4.5 CORRELAÇÕES E SIGNIFICÂNCIAS RESULTANTES DAS REGRESSÕES ENTRE A VARIÁVEL REFERENTE À PERCENTAGEM DE HOMENS E MULHERES COM 65 ANOS OU MAIS QUE VIVEM SOZINHOS E A TAXA DE INCIDÊNCIA DE FRACTURAS OSTEOPORÓTICAS DO COLO DO FÉMUR, PARA HOMENS E MULHERES.	97

TABELA 4.6 DISTRIBUIÇÃO DA POPULAÇÃO, DOS CONCELHOS E DAS FRACTURAS OSTEOPORÓTICAS DO COLO DO FÉMUR POR <i>CLUSTER</i>	97
TABELA 4.7 DESCRIÇÃO DE CADA <i>CLUSTER</i> COM BASE NAS COMPONENTES PRINCIPAIS.	98
TABELA 4.8 TAXA DE INCIDÊNCIA DAS FRACTURAS OSTEOPORÓTICAS DO COLO DO FÉMUR, PARA HOMENS.	104
TABELA 4.9 TAXA DE INCIDÊNCIA DAS FRACTURAS OSTEOPORÓTICAS DO COLO DO FÉMUR, PARA MULHERES.	104
TABELA 4.10 TAXA DE INCIDÊNCIA DAS FRACTURAS OSTEOPORÓTICAS DO COLO DO FÉMUR, PARA AMBOS OS GÉNEROS.	104
TABELA 4.11 DISTRIBUIÇÃO DOS CASOS DE ÓBITOS E DE NADOS VIVOS, NAS TRÊS DÉCADAS.	107
TABELA 4.12 MORTALIDADE INFANTIL EM CADA DÉCADA, POR DISTRITO (POR MIL NADOS VIVOS). ...	107
TABELA 4.13 CORRELAÇÕES E SIGNIFICÂNCIAS RESULTANTES DAS REGRESSÕES ENTRE A MORTALIDADE INFANTIL POR DISTRITO E A TAXA DE INCIDÊNCIA DAS FRACTURAS DO COLO DO FÉMUR.	108
TABELA 4.14 CORRELAÇÕES E SIGNIFICÂNCIAS RESULTANTES DAS REGRESSÕES ENTRE A MORTALIDADE INFANTIL POR ANO NAS DÉCADAS DE 40 E 50, E A COMPONENTE PRINCIPAL 1.	109
TABELA 4.15 CORRELAÇÕES E SIGNIFICÂNCIAS RESULTANTES DAS REGRESSÕES ENTRE A MORTALIDADE INFANTIL POR ANO NAS DÉCADAS DE 40 E 50, E A COMPONENTE PRINCIPAL 2.	110
TABELA 4.16 CORRELAÇÕES E SIGNIFICÂNCIAS RESULTANTES DAS REGRESSÕES ENTRE A MORTALIDADE INFANTIL POR ANO NAS DÉCADAS DE 40 E 50, E A COMPONENTE PRINCIPAL 3.	111
TABELA 4.17 CORRELAÇÕES E SIGNIFICÂNCIAS RESULTANTES DAS REGRESSÕES ENTRE A MORTALIDADE INFANTIL POR ANO NAS DÉCADAS DE 40 E 50, E A COMPONENTE PRINCIPAL 4.	111
TABELA 4.18 CORRELAÇÕES E SIGNIFICÂNCIAS RESULTANTES DAS REGRESSÕES ENTRE A MORTALIDADE INFANTIL POR DISTRITO E A TAXA DE INCIDÊNCIA DAS FRACTURAS OSTEOPORÓTICAS DO COLO DO FÉMUR, SEM OS DISTRITOS DE PORTO E LISBOA.	112

ABREVIATURAS

ACP	Análise de Componentes Principais
DMO	Densidade Mineral Óssea
DVS	Decomposição em Valores Singulares
EVOS	European Prospective Osteoporosis Study
ICD9-CM	International Classifications of Diseases 9th revision – Clinical modification
IGEO	Instituto Geográfico
IGIF	Instituto de Gestão Informática e Financeira da Saúde
INE	Instituto de Engenharia Biomédica
KMO	Kaiser-Meyer-Olkin
MEDOS	Mediterranean Osteoporosis Study
OMS	Organização Mundial de Saúde
PTH	Paratormona
SIG	Sistema de Informação Geográfica
SNS	Serviço Nacional de Saúde
SPSS	Statistical Package for Social Sciences

1. INTRODUÇÃO

1.1 Enquadramento

Durante muitos anos a osteoporose foi considerada uma consequência inevitável do envelhecimento. No entanto, com o desenvolvimento da ciência e um melhor conhecimento dos mecanismos que conduzem à osteoporose, passou a ser considerada por diversos autores como uma doença, possível de evitar ou retardar (Hunter D, Sambrook P, 2000). Como Kanis J, Delmas P et al. (1997) referem, têm sido realizados desenvolvimentos importantes no campo da osteoporose nos últimos anos e, nos dias de hoje já existe uma quantidade considerável de informação relacionada com o impacto desta doença na saúde pública e consenso no que diz respeito à sua definição.

De acordo com o *Concense Developmente Conference*, em 1991, a osteoporose é uma doença que se caracteriza por anomalias na quantidade e qualidade do tecido ósseo, o que conduz a uma diminuição de força no esqueleto e um aumento da susceptibilidade de sofrer fracturas (Schutte H, 1995). O significado clínico da osteoporose recai, assim, no aparecimento de fracturas. As fracturas mais comuns incluem fracturas vertebrais de compressão, fracturas do rádio distal ou punho e do fémur proximal ou colo do fémur. Estas fracturas são relativamente comuns e o seu tratamento cada vez se torna mais dispendioso e complicado (Court-Brown C, Caesar B, 2006). No entanto quando um esqueleto está osteoporótico outras fracturas podem ocorrer, como por exemplo na pélvis, no úmero proximal, no fémur distal e nas costelas (Kanis J, Delmas P et al., 1997).

Além de ocorrerem mais intensamente em algumas partes do esqueleto, as fracturas osteoporóticas caracterizam-se também pela alta incidência em indivíduos com mais de 50 anos, e por resultarem de impactos de baixa energia.

A osteoporose afecta milhões de pessoas em todo o mundo e o aumento da população idosa em todas as regiões geográficas leva a que as fracturas osteoporóticas sejam vistas como uma epidemia ortopédica e um grave problema de saúde pública.

As mulheres têm um risco maior de sofrer fracturas do que os homens e, em alguns países, a incidência nas mulheres chega a ser três vezes superior à incidência nos homens. Possíveis razões para este evento podem ser: um pico de massa óssea mais baixo quando atingida a maturidade, a aceleração da perda de massa óssea após a

menopausa, uma maior longevidade que os homens e uma maior possibilidade de quedas entre mulheres idosas (Schutte H, 1995).

As mais dramáticas consequências da osteoporose são as fracturas do colo do fémur, porque quase sempre obrigam a uma intervenção cirúrgica e estão associadas a altas taxas de mortalidade e de morbilidade dos doentes, em relação às pessoas da mesma idade não operadas. Por estarem, também fortemente relacionadas com uma baixa densidade mineral óssea, as fracturas do colo do fémur, causadas por traumas de baixo impacto, ocorridas em indivíduos acima de 50 anos são consideradas como um indicador de osteoporose. A incidência de outras fracturas osteoporóticas, nomeadamente das vértebras e do punho, é mais difícil de determinar, pois geralmente são tratadas em ambulatório e não requerem internamento tornando o registo das informações a nível nacional menos eficaz (Laet C, Pols H, 2000).

Os custos sociais da osteoporose são muito elevados e é urgente conhecer a real dimensão do problema em Portugal de maneira a planear acções de intervenção e de prevenção capazes de reduzir as taxas de incidência das fracturas do colo do fémur e melhorar a qualidade de vida das pessoas. Alguns autores, como Hernández JL, Olmos JM et al., (2005), Lodthus C, Osnes E et al., (2001) e Lesic A, Jarebinski M et al., (2006), referem que as acções de prevenção podem estar associadas com a recente tendência de diminuição das taxas de incidências das fracturas do colo do fémur em diversos países europeus, nomeadamente nos países nórdicos.

O processo saúde-doença não é condicionado apenas por factores biológicos e genéticos mas é também resultado de padrões de ocupação do espaço e das condições socioeconómicas e ambientais, que formam a totalidade dos elementos externos e que influenciam as condições de saúde e a qualidade de vida dos indivíduos ou comunidades (Barcelos C, Ramalho W, 2002, citado em Pina MF, 2003^a). Diversos têm sido os estudos que revelam que o ambiente social, determinado pelos grupos populacionais e a sua organização estão associados a doenças e ao risco de morrer, independentemente dos factores de risco individuais (Pina MF, 2003^a).

Os diferentes eventos em saúde, tanto positivos quanto negativos, não ocorrem por acaso. Através dos séculos, observou-se uma relação muito estreita destes eventos com o meio ambiente, as condições sociais e outros determinantes. Por exemplo: alguns eventos tendem a apresentarem-se em grupos populacionais com certas condições de

vida e de trabalho, em certas épocas do ano e, em determinadas condições habitacionais e de migração (Rocha L, 2000). Em termos de ecologia urbana das doenças, as cidades têm sido cada vez mais estudadas. Particularmente em países em desenvolvimento, os moradores de cidades vivem em diferentes condições ambientais como moradia, emprego, estilo de vida, dieta, entre outros. A poluição, superpopulação, stress, pobreza afectam a saúde humana. O espaço, produzido socialmente, exerce pressões económicas e políticas sobre a sociedade, criando condições diferenciadas para a sua utilização por grupos sociais (Ramalho W, Skaba D, 2007).

Um dos principais objectivos da epidemiologia é identificar áreas geográficas e grupos populacionais que apresentam maior risco de adoecer ou morrer prematuramente e que, portanto, precisam de maior atenção, seja preventiva, curativa ou de promoção de saúde. A epidemiologia também permite reconhecer que a frequência, a distribuição e a importância dos diversos factores que influem no aumento de determinados riscos para a saúde não são, necessariamente, os mesmos em todos os grupos populacionais. Permite também identificar grupos que compartilham determinantes de risco similares. O reconhecimento desses grupos facilita a identificação de intervenções sociais e de saúde *ad hoc* para diminuir ou eliminar os determinantes específicos de risco para a saúde (Rocha L, 2000).

De acordo com o estudo de Alves S (2004), no ano de 2020 em Portugal poder-se-á esperar que 10.684 indivíduos sofram fracturas osteoporóticas do colo do fémur e que em 2050 este número tenha aumentado para 11.552. Segundo este mesmo estudo estas fracturas representarão para o país um custo de, aproximadamente, 55.800.000 euros no ano de 2020 e de 60.300.000 euros no ano de 2050.

Desta forma, os problemas relacionados com as fracturas osteoporóticas tendem a agravar-se drasticamente em Portugal, quer em termos sociais quer em termos económicos. Torna-se então urgente reconhecer quais os factores que estão a contribuir para o aumento da incidência destas fracturas de maneira a tomar as medidas preventivas mais apropriadas e consequentemente reduzir o impacto desta epidemia. É neste contexto que este estudo se insere.

Apesar de existirem vários artigos onde são comparadas as diferentes taxas de incidência das fracturas do colo do fémur entre vários países e onde se estuda a forma como estas taxas estão a evoluir ao longo dos anos, muito poucos ou nenhuns foram os

artigos encontrados que evidenciavam estas diferenças geográficas dentro dos vários países e as tentem relacionar com as condições socioeconómicas. No entanto, no seu trabalho Alves S (2004) conclui que em Portugal, existem diferenças geográficas acentuadas e que a diferença entre os concelhos chega a ser superior a três vezes.

Uma vez que em Portugal já foram identificadas estas fortes diferenças geográficas falta explicar porquê estas diferenças ocorrem. Diversos podem ser os factores, entre eles, os que têm sido mais frequentemente indicados na literatura internacional são os estilos de vida, em especial no que diz respeito à actividade física e alimentação, exposição ambiental, factores socioeconómicos, genéticos, entre outros.

Diversos autores apontam a relação entre a incidência ou risco de sofrer fracturas do colo do fémur e factores socioeconómicos, tais como: o estado civil, o rendimento individual ou rendimento familiar, o nível de escolaridade, o tipo de trabalho e algumas condições de habitação (como, o número de pessoas por alojamento).

1.2 Objectivos

O objectivo deste trabalho é explicar as variações geográficas das fracturas do colo do fémur em Portugal através do estudo da associação entre a sua incidência e factores socioeconómicos. As bases de dados utilizadas, dos registos hospitalares não possuem nenhuma informação socioeconómica sobre os doentes internados, e por esta razão a estratégia utilizada foi realizar um estudo ecológico em que serão analisadas variáveis socioeconómicas, a nível concelhio.

Para se alcançar o objectivo proposto foram analisadas diversas variáveis do último Censo realizado em Portugal (2001), de maneira a classificar os concelhos de acordo com condições socioeconómicas homogéneas e em seguida procurar uma explicação para as diferenças geográficas encontradas na incidência das fracturas do colo do fémur com base nos padrões socioeconómicos encontrados.

Este trabalho pode ser considerado inovador quer em objectivos quer em metodologia, uma vez que o estudo das fracturas do colo do fémur num contexto geográfico é uma abordagem nova não só em Portugal como também no resto do mundo.

1.2 Estrutura

O presente trabalho é composto por seis capítulos, sendo o primeiro referente à Introdução. O segundo capítulo aborda os conceitos teóricos das técnicas necessárias para desenvolver o estudo, bem como o estado da arte, no que diz respeito aos estudos existentes sobre osteoporose e fracturas do colo do fémur. É neste capítulo que se apresenta uma abordagem breve à epidemiologia e epidemiologia social, à utilização de sistemas de informação geográfica na saúde, à análise de componentes principais e análise de *clusters*. O terceiro capítulo é dedicado aos materiais e à metodologia utilizada. No quarto capítulo são apresentados os resultados encontrados referentes à análise de componentes principais, à análise de *clusters* e às regressões efectuadas. Por último, no quinto capítulo é apresentada a discussão, onde é feita a análise dos resultados obtidos terminando o capítulo com as conclusões mais relevantes e perspectivas de desenvolvimentos futuros.

Como anexos apresenta-se a bibliografia relevante não citada no texto (**Anexo A**), a tabela de correlações das variáveis retiradas do Censos de 2001 (**Anexo B**) o quadro de aglomeração e o dendograma relativos à análise de *clusters* efectuada (**Anexo C e D**) e, finalmente as tabelas os gráficos correspondentes à mortalidade infantil, por distrito, entre 1930 e 1959 (**Anexo E**).

2. ESTADO DA ARTE

2.1 Epidemiologia

Adoecer não é um fenómeno aleatório e esta verdade tem estimulado a perscrutação científica de repetições elaboráveis como leis e possibilidades de intervenção. Detectar a lógica do equilíbrio - a saúde - e identificar os agentes da sua perturbação - a doença - são os desafios do trabalho epidemiológico (Barros H, 1997).

A saúde pública difere da medicina clínica quer na sua ênfase à prevenção em vez do tratamento quer por se focar em populações em vez de se focar em pacientes individuais (Pearce N, 2005). Ainda segundo este autor, a saúde pública centra-se na prevenção da doença na população humana e a epidemiologia é o seu ramo que procura descobrir as causas de uma doença de forma a tornar possível a sua prevenção.

A palavra *epidemiologia* deriva da composição de três palavras gregas: *epi* que significa “sobre”, *demos* que significa “pessoas” e *logos* que significa “estudo de”. Muitas definições têm sido propostas para o seu significado. Como em Almeida N (1991) menciona, por exemplo, Lilienfeld D (1978) propôs que seria melhor tomar a epidemiologia como uma "abordagem conceptual", um "modo de pensar" e um "método de raciocínio", enquanto que Fox J, Hall C et al. (1970) afirmam que, pelo facto de não se constituir num corpo de conhecimentos, "a epidemiologia tem sido considerada um método e não uma ciência independente". Para Kleinbaum D, Kupper L et al. (1982), também citados em Almeida N (1991), a epidemiologia é simultaneamente "uma síntese de conhecimentos oriundos de diversas disciplinas da saúde", "um ramo aplicado da ciência" e uma "disciplina" (no sentido de uma "metodologia" aplicável a qualquer pesquisa em saúde).

Barros H (1997) e Dicker R (2005) optam pela definição que capta os princípios e o espírito da epidemiologia: “Epidemiologia é o estudo da distribuição e dos determinantes de estados ou acontecimentos relacionados com a saúde em populações específicas, e a aplicação deste estudo ao controlo dos problemas de saúde”.

Como apresentado no dicionário de epidemiologia editado por Last J (2001), o “estudo” inclui a vigilância, observação, verificação de hipóteses, investigação analítica e experimentação; os “estados ou acontecimentos relacionados com a saúde” vão além da posição tradicionalmente central das doenças para incluírem o nível de saúde, as causas de morte, os comportamentos (por exemplo o tabagismo), as respostas a regimes

preventivos e o proporção e utilização dos serviços de saúde – por outras palavras: originalmente a epidemiologia centrava-se nas epidemias de doenças comunicáveis, mas foi estendida a doenças comunicáveis endémicas e doenças infecciosas não comunicáveis, mais recentemente os métodos epidemiológicos têm sido aplicados a doenças crónicas, defeitos de nascimento, saúde de mãe e filho, saúde ocupacional e saúde ambiental (Dicker R, 2005); finalmente, a “aplicação ao controlo dos problemas de saúde” explícita a finalidade da epidemiologia num vasto programa: promover, proteger e restaurar a saúde, ou seja esta disciplina providencia dados que permitem direccionar acções de saúde pública (Barros H, 1997).

Segundo o trabalho de Dicker R (2005) o uso da epidemiologia pode ter os seguintes objectivos:

- Acesso à saúde por parte de populações e comunidades. Ou seja, para objectivar políticas e planear estratégias, os responsáveis pela saúde pública devem determinar se os sistemas de saúde estão disponíveis, acessíveis, se são efectivos e eficientes. Para o fazer, devem responder a perguntas como: Quais são os problemas de saúde actuais e potenciais numa determinada comunidade? Onde estão? Quem é a população sob risco? Quais os problemas que vão melhorar ao longo do tempo? Quais os problemas que aumentam e quais os que têm o potencial maior para aumentar? Como é que estes padrões se relacionam com o nível e a distribuição dos serviços disponíveis?

Com respostas a perguntas como estas, obtidas através dos estudos epidemiológicos, os responsáveis podem tomar decisões informadas que poderão melhorar a saúde da população que servem.

- Decisões individuais. Muitas vezes as pessoas podem não se aperceber que para as decisões que tomam diariamente utilizam informação epidemiológica. Quando decidem parar de fumar, subir as escadas em vez de utilizar o elevador, pedir uma salada em vez de um hambúrguer, ou utilizar um método contraceptivo em vez de outro, as pessoas podem estar a ser influenciadas, consciente ou inconscientemente por estudos epidemiológicos da avaliação do risco. Desde a segunda guerra mundial que os epidemiologistas providenciam informação relacionada com estas decisões. Nos anos 50 do século passado documentaram o aumento do risco de sofrer cancro do pulmão entre os fumadores, no anos 60 e 70 mostraram uma grande variedade de benefícios e riscos associados à utilização de diferentes tipos de anti-contraceptivos e nos anos 80,

os epidemiologistas identificaram o aumento do risco de infecção pelo VIH/Sida entre determinados comportamentos sexuais e comportamentos de toxicodependência. E ainda hoje os epidemiologistas continuam a documentar o papel do exercício físico e de uma dieta adequada no que diz respeito a uma diminuição do risco de sofrer uma doença cardíaca.

Estas e outras centenas de descobertas epidemiológicas são directamente relevantes para as decisões que as pessoas tomam todos os dias, decisões estas, que afectam a sua saúde.

- Completar o quadro clínico. Quando se estuda epidemiologicamente uma doença, os epidemiologistas dependem de laboratórios, médicos e cientistas para obterem o diagnóstico adequado dos pacientes. No entanto, os epidemiologistas contribuem para a compreensão de um determinado acontecimento e para a história natural de uma doença. Por exemplo, foi a epidemiologia que documentou o curso da infecção do VIH desde a sua exposição inicial até ao desenvolvimento de uma grande variedade de síndromes que incluem o síndrome da imunodeficiência adquirida - SIDA.

- Procurar causas. Grande parte da pesquisa epidemiológica é direccionada para a procura de causas e factores que influenciam o risco de adoecer, ou seja, identificação das relações causais de modo a poderem ser tomadas as acções mais adequadas.

Concluindo, a epidemiologia é considerada uma disciplina interactiva e eclética que combina os conhecimentos de outras disciplinas como a estatística, biologia, geografia e sociologia. De forma a atingir os objectivos, os epidemiologistas estudam vários factores como as características demográficas, biológicas, genéticas, sociais, económicas e hábitos pessoais e através deles determinam a associação estatística entre estas características e as doenças, e realizam inferências biológicas destas associações (Lilienfeld D, 1994).

2.1.1 Epidemiologia social

A epidemiologia desenvolveu-se em diversos campos de aplicação, levando à existência de especialidades como: epidemiologia ocupacional, epidemiologia do cancro, fármaco-epidemiologia, epidemiologia social e diversas outras (Zielhuis G, Kiemeney L, 2001).

Segundo Souza E e Grundy E (2004) a epidemiologia social é definida como o ramo da epidemiologia que estuda a distribuição e a influência dos factores sociais na saúde. O campo da epidemiologia social traz o foco de atenção, anteriormente voltado sobretudo para os factores de risco para a saúde, para examinar com mais profundidade, o contexto social em que as doenças ocorrem. Assim, é possível identificar e descrever quais as condições sociais que parecem influenciar o estado de saúde das populações, aspectos estes pouco abordados dentro da epidemiologia tradicional. Da mesma opinião é Kawachi I (2002) que menciona, no seu trabalho, que a variedade de problemas que a epidemiologia social estuda inclui questões como, por exemplo, se o local de trabalho e a sua organização, ou o rendimento e a coesão social, podem afectar a saúde.

As três últimas décadas têm mostrado um interesse crescente em entender como a sociedade e diferentes formas de organização social influenciam a saúde e o bem-estar da população. Autores como Syme SL (1992) e Susser M, Watson W et al. (1985), citados em Zielhuis G, Kiemeney L (2001), referem que a epidemiologia social é “uma distinta e reconhecida especialidade entre a epidemiologia” e que o tem sido desde há 30 ou 40 anos atrás.

Revisões históricas mostram que a relação entre saúde e factores socioeconómicos tem sido documentada por vários séculos desde a Grécia Antiga, Egipto e China e que essa relação, portanto, antecede o estabelecimento da epidemiologia social (Souza E, Grundy E, 2004). Também Krieger N (2001) faz referência a estudiosos dos séculos XVIII e XIX que já se preocupavam com a influência dos determinantes sociais na saúde.

Segundo Souza E, Grundy E 2004, em 1662 Graunt J estudou os efeitos das variações sociais na morbilidade e mortalidade na Inglaterra. Médicos como Vilherme LR, em 1830, e Virchow R, em 1848, apontaram as diferenças entre as classes sociais e as condições de trabalho como factores cruciais na determinação do estado de saúde e de doença, enquanto Durkheim E, em 1897, relacionou a integração social a padrões de mortalidade, mormente o suicídio. Actualmente, vários estudos nas áreas de psicofisiologia e psiconeuroimunologia têm mostrado a influência de factores externos nas respostas fisiológicas capazes de influenciar o prognóstico ou o desenvolvimento de doenças, mesmo aquelas determinada por factores genéticos. Ainda segundo este autor, mais recentemente a hipótese da carga alostática, ou danos do stress induzido, tem

ajudado a traçar os caminhos biológicos que conectam os padrões sociais ao estado de saúde. Outros estudos mostram ainda que a qualidade de vida social é um dos mais poderosos determinantes da saúde das populações, a qual está também estreitamente associada ao grau de equidade na distribuição de renda (Souza E, Grundy E 2004).

Segundo Kawachi I (2002) a epidemiologia tem sido definida como a “ciência da contagem”, e os epidemiologistas estão treinados para contar a ocorrência de uma doença em populações. Em grande parte esta contagem tem como objectivo monitorizar a tendência da ocorrência de uma doença para informar os responsáveis pela saúde. Segundo este autor, os epidemiologistas sociais diferenciam-se dos outros ao nível das razões causais. Em vez da pergunta “porque é que este indivíduo ficou doente?”, os epidemiologistas sociais perguntam-se “porque é que esta sociedade não é saudável?”.

O objectivo da epidemiologia social é conceptualizar, tornar operacional e testar as associações entre os aspectos do ambiente social (famílias, locais de trabalho, regiões onde se vive, e políticas económicas) e a saúde da população (Kawachi I 2002 e Krieger N, 2001).

A observação de que a saúde sofre influência do meio social e físico apela à reflexão para o facto de que em relação às doenças existem modificações epidemiológicas consideráveis (Souza E, Grundy E, 2004). Muitas doenças infecciosas dominaram o cenário e foram erradicadas, outras emergiram e, actualmente, as doenças não infecciosas não comunicáveis correspondem às principais causas de morte em vários países desenvolvidos e muitos países em desenvolvimento ou pobres. Mas as desigualdades sociais continuam a crescer e este é um dos factores de risco para a saúde que tem recebido pouca atenção. Esse padrão clama por uma abordagem epidemiológica mais consistente em que a promoção de saúde tem papel crucial (Souza E, Grundy E, 2004).

2.2 Osteoporose

A osteoporose é uma doença progressiva caracterizada pela diminuição da massa óssea e deterioração da microarquitetura dos tecidos do osso, aumentando a fragilidade dos ossos e subsequentemente o risco de fracturas (Dennison E, Cooper C, 2000; Johnell O, Kanis J, 2004; Johnston C, Slemenda C, 1995). Estas fracturas são uma das causas mais comuns de incapacidade física e um dos maiores contribuidores para os elevados custos hospitalares em várias regiões do mundo (Cummings SR, Melton L, 2002).

A terminologia associada à osteoporose foi desenvolvida no século XIX por patologistas alemães para distinguir osteoporose, osteomalacia e osteíte fibrosa cística. Mais tarde, em 1941, Fuller Albright e os seus colegas definiram patologicamente a osteoporose como sendo uma condição em que existe falta de tecido ósseo mas onde o tecido remanescente está completamente saudável. Esta definição permitiu distinguir a osteoporose de osteomalacia, sendo esta última uma condição que envolve “falhas” na mineralização da matriz óssea, devido, frequentemente, à deficiência de vitamina D ou a distúrbios metabólicos (Wolf A, Dixon A, 1988). Assim, nem sempre a confirmação do diagnóstico da osteoporose foi evidente, tendo sido considerada durante muito tempo como uma consequência a que o envelhecimento conduzia. Este diagnóstico requeria uma história de pelo menos uma fractura de baixo impacto em partes específicas do esqueleto. Foi com a introdução de medições precisas e não-invasivas da massa óssea que se tornou possível fazer um diagnóstico precoce da osteoporose.

O osso é um órgão que dá forma ao corpo, suporta o seu peso, protege os órgãos vitais e facilita a locomoção (Wolf A, Dixon A, 1988). Ou seja, é um sistema mecanicamente optimizado cuja composição e organização reflecte as funções que lhe são exigidas (Marcus R, Feldman D et al., 1996). Para além disto, o osso é também um tecido vivo especial, composto por material orgânico e inorgânico e sendo uma entidade biológica, o tecido ósseo é constituído por material composto por uma matriz extracelular que tem impregnada uma fase inorgânica (Marcus R, Feldman D et al., 1996).

Os ossos são assim compostos de duas camadas (ver **Figura 2.1**): uma camada mais externa e mais dura, denominada de osso cortical (este é um osso mais denso e

compacto) que constitui 80% da massa do esqueleto e envolve uma camada interna e mais esponjosa, denominada de osso trabecular (Marcus R, Feldman D et al., 1996).

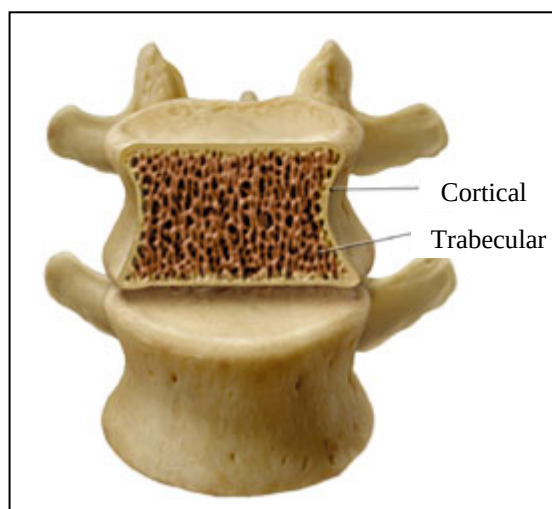


Figura 2.1 Vértebra onde é possível observar as duas partes distintas do osso.

Fonte: < http://www.ama-cmeonline.com/osteo_mgmt/module03/02path/02.htm>.

Num osso saudável, o osso trabecular esponjoso é constituído por milhares de traves fortemente conectadas entre si. Este osso está presente principalmente nas vértebras, crânio, pélvis e porção ultradistal do rádio, enquanto que o osso cortical predomina nos ossos longos, colo do fémur e rádio distal (Marcus R, Feldman D et al., 1996).

O tecido ósseo é muito metabólico e encontra-se em constante processo de formação e remodelação. A remodelação óssea é um processo dinâmico, que depende da acção integrada e acoplada de reabsorção óssea pelos osteoclastos e de formação óssea pelos osteoblastos. Este processo de formação óssea, também chamado de osteogénese, decorre em 3 etapas: a produção de matriz orgânica extracelular ou osteóide, a mineralização da matriz com formação de osso e a remodelação óssea com reabsorção seguida da formação de novo osso.

A actividade celular dos osteoblastos, osteócitos (ou osteoblastos maduros) e osteoclastos é fundamental neste processo. Os osteoblastos sintetizam os precursores do colagénio I, que constituem 90 a 95% da matriz orgânica do osso e os precursores moleculares da matriz óssea e regulam a sua mineralização. Algumas substâncias como hormonas e factores de crescimento e outros estímulos como a actividade física, exercem efeitos no osso, actuando através dos osteoblastos. À medida que progride o processo de formação óssea, os osteoblastos preenchem as lacunas de reabsorção

produzidas pelos osteoclastos, produzem osteóide e passam a denominar-se osteócitos (Canhão H, Fonseca J et al., 2005).

O constituinte mineral mais importante do osso é a hidroxiapatite, um composto formado por cálcio e fosfato, que constitui cerca de um quarto do volume e mais de metade da massa do osso adulto normal. A vitamina D e a paratormona (PTH) são importantes mediadores da regulação do cálcio e a deficiência da primeira ou excesso da segunda podem conduzir a uma redução mineral (Canhão H, Fonseca J et al., 2005).

Idealmente o tecido ósseo deve possuir a rigidez necessária para suportar a carga do corpo, a flexibilidade suficiente para absorver impactos sem fracturar e a leveza adequada para permitir movimentos rápidos. Em grande parte, estas características são moduladas pela quantidade de cristais de hidroxiapatite. Se o osso sofre uma desmineralização, como em algumas situações em que ocorre o aumento da reabsorção óssea, torna-se demasiado flexível, flexiona demasiado durante a carga e fractura ou, se fica excessivamente mineralizado, flexiona pouco durante a carga e também fractura (Canhão H, Fonseca J et al., 2005).

Esta remodelação, que resulta do balanço entre a formação e a reabsorção ósseas, decorre ao longo de toda a vida e é uma função essencial do osso que assegura o equilíbrio do metabolismo do cálcio e do fósforo e também a reparação de microdanos do osso (Canhão H, Fonseca J et al., 2005).

Assim, a nível fisiológico o osso é continuamente depositado por osteoblastos e absorvido nos locais onde os osteoclastos estão activos (Marcus R, Feldman D et al., 1996). Após a degradação do osso da qual resulta a formação de uma lacuna de reabsorção, os osteoclastos destacam-se e são substituídos por osteoblastos, que nos jovens restauram completamente o osso que foi previamente reabsorvido. Ou seja, normalmente, a não ser nos ossos em crescimento, há equilíbrio entre deposição e absorção óssea. Assim que o crescimento fica completo, a continuação de formação de osso saudável depende deste processo de deposição/absorção. Nos idosos a quantidade de osso depositada pelos osteoblastos nas lacunas de reabsorção é menor do que a removida pelos osteoclastos (Canhão H, Fonseca J et al., 2005). Na osteoporose existe desproporção entre as actividades osteoblástica e osteoclástica, com predomínio da última (Marcus R, Feldman D et al., 1996). Como o osso trabecular apresenta um maior metabolismo é também o mais susceptível a alterações da massa óssea.

Sendo assim, na osteoporose o osso cortical afina-se gradualmente e os orifícios do osso trabecular tornam-se cada vez maiores e irregulares (ver **Figura 2.2**). Quando a estrutura interna do osso está comprometida, um traumatismo resultante de uma pequena queda ou mesmo o peso normal do corpo pode causar fracturas.

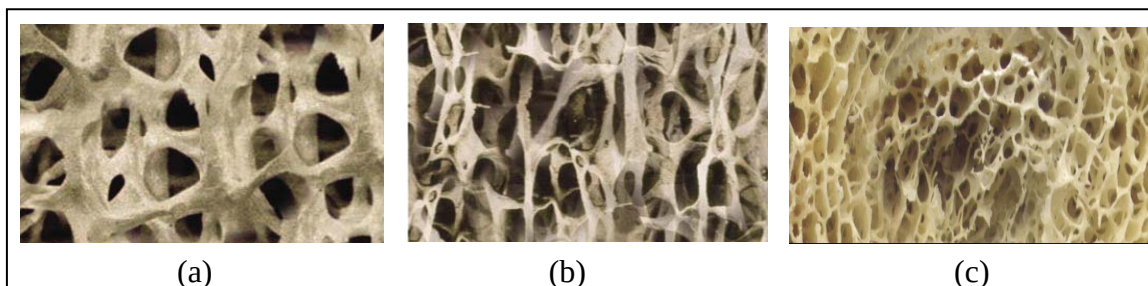


Figura 2.2 (a) osso saudável; (b) e (c) osso com osteoporose.

Fonte: Hashefi M, Burke M et al., (2005).

A osteoporose pode ser considerada primária (ou idiopática) ou secundária.

A forma primária é classificada em tipo I e tipo II:

- Na osteoporose primária do tipo I, também conhecida por tipo pós-menopausa, há uma rápida perda óssea, directamente resultante da falta de estrogénio endógeno e ocorre na mulher que entrou recentemente na menopausa. Predominantemente, atinge o osso trabecular e é associada a fracturas das vértebras e do punho.
- A osteoporose primária do tipo II, ou senil, está relacionada com o envelhecimento, tanto nos homens como nas mulheres e aparece por deficiência de cálcio e vitamina D, alterações na absorção intestinal, diminuição da formação óssea, entre outros.

A osteoporose secundária é decorrente de processos inflamatórios, como a artrite reumatóide, alterações endócrinas, como hipertireoidismo, mieloma múltiplo e uso de drogas (como heparina, álcool, vitamina A e glicocorticóides) (Feldstein AC, Nichols G et al., 2005; Lane N, 2006). Ou seja, refere-se à perda óssea resultante de distúrbios clínicos específicos (Marcus R, Feldman D et al., 1996).

A massa óssea resulta do balanço entre o pico de massa óssea obtido no fim da maturação do esqueleto e a taxa de perda óssea posterior.

Esta massa óssea é um dos maiores indicadores da osteoporose a longo prazo e, é adquirida durante o período de crescimento e desenvolvimento do esqueleto (Yoshimura T, Tohya T et al., 2005). Do total de massa óssea, 90% é adquirida até aos 18 anos e os restantes 10% até os 30 anos, atingindo-se um pico ósseo nesta idade,

normalmente maior nos homens do que nas mulheres. Este pico é o maior determinante do risco de fracturas osteoporóticas ao longo da vida.

A partir dos 30 anos irá haver sempre perda de massa óssea (aproximadamente 0,3% ao ano), no entanto, numa velocidade maior ou menor dependendo de como a sua gestão é feita. Logo, em adultos, a massa óssea presente em qualquer fase da vida depende não só da quantidade adquirida até à maturidade, como também da perda resultante do envelhecimento. Esta perda, que começa por inicialmente ser lenta e semelhante entre homens e mulheres, mantém-se sempre constante nos homens e torna-se mais acentuada nas mulheres, nos primeiros dez anos após a menopausa. Nesta altura a perda pode chegar até 3% de massa óssea por ano e afecta maioritariamente as mulheres que têm vidas mais sedentárias. (Marcus R, Feldman D et al., 1996). No entanto, apesar desta perda ser inevitável, nem sempre é tão elevada que permite classificar a condição que lhe é associada de osteoporose; quando tal acontece atribui-se o nome de osteopenia.

Medir a densidade óssea permite estabelecer ou confirmar o diagnóstico de osteoporose e prever risco de futuras fracturas (Wolf A, Dixon A, 1988). A densidade mineral óssea (DMO) de doentes com fracturas osteoporóticas é, em geral, mais baixa do que em indivíduos saudáveis e já existem vários estudos que indicam de que forma o risco de fracturas osteoporóticas aumenta continuamente à medida que a DMO diminui (nomeadamente, nas fracturas do colo do fémur). De acordo com esta premissa foi proposto, por um grupo de investigadores, uma definição de osteoporose baseada exclusivamente nos valores da massa óssea (Marcus R, Feldman D et al., 1996).

A DMO é expressa como a relação entre duas normas: a DMO esperada para o paciente conforme a idade e o género (*Z-score*) ou para jovens adultos, considerados saudáveis, do mesmo género (*T-score*). Os resultados dos pacientes são expressos com base no número de desvios-padrão em relação aos resultados considerados normais. Os *Z-scores* não são utilizados para diagnosticar a osteoporose devido à redução da massa óssea que usualmente ocorre com o aumento da idade (Hashefi M, Burke M et al., 2005).

A Organização Mundial de Saúde (OMS) definiu a osteoporose em função do número de desvios padrão em relação à média da DMO de mulheres jovens, consideradas saudáveis. A presença de um valor igual ou superior a 2,5 desvios-padrão numa mulher é, geralmente, interpretado como um diagnóstico de osteoporose. A osteoporose severa

(ou estabelecida) tem o mesmo limite de desvios-padrão da osteoporose, mas é associada à presença de uma ou mais fracturas, normalmente, fracturas no punho, nas vértebras ou no colo do fémur. Na osteopenia os valores da DMO situam-se entre 1 e 2,5 desvios-padrão (Wolf A, Dixon A, 1988) (ver **Tabela 2.1**).

Tabela 2.1 Definições da OMS baseadas na DMO.

Normal	Até 1 desvio-padrão em relação a uma média de jovens adultas saudáveis ($T\text{-score} > -1,0$)
Osteopenia	Entre 1 e 2,5 desvios-padrão em relação uma média de jovens adultas saudáveis ($-1,0 < T\text{-score} < -2,5$)
Osteoporose	Superior a 2,5 desvios-padrão em relação uma média de jovens adultas saudáveis ($T\text{-score} < 2,5$)
Osteoporose severa ou estabelecida	Osteoporose acrescida de uma ou mais fracturas.

Baseado em: Hashefi M, Burke M et al. (2005).

A DMO, de uma população feminina jovem e saudável, segue uma distribuição normal e avalia-se que 15% das mulheres terá um $T\text{-score}$ inferior a -1 desvio-padrão e, portanto, sofrerão de osteopenia e, aproximadamente, 0,6% terão já osteoporose (Kanis J, Delmas P et al., 1997) (ver **Figura 2.3**). Segundo Schutte H (1995), a proporção de mulheres, com idade de 50 anos ou mais, com uma baixa massa óssea varia entre 30 e 45%.

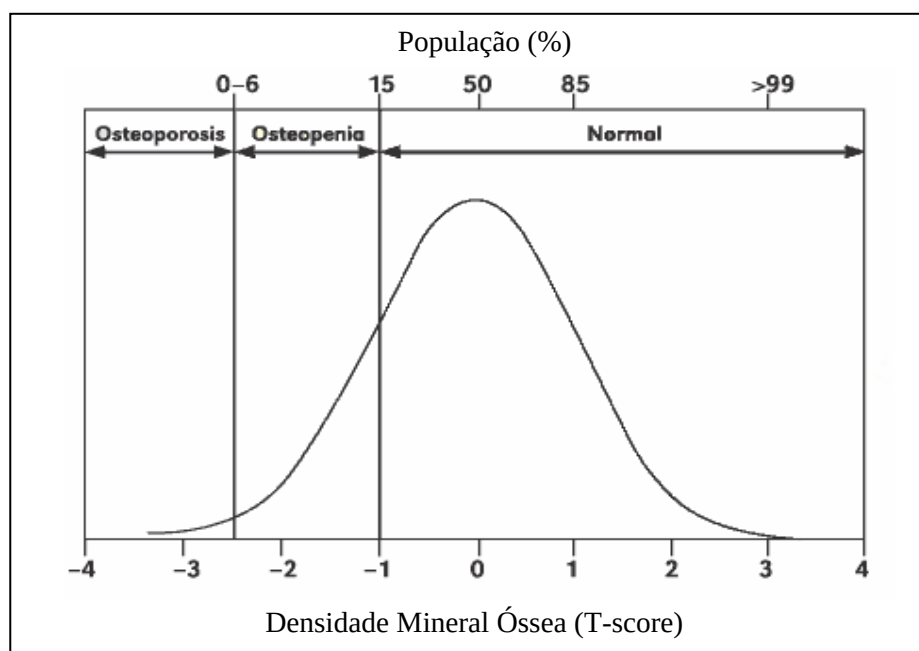


Figura 2.3 Distribuição da DMO em jovens mulheres saudáveis com idades entre os 30 e 40 anos.

Fonte: Kanis J, Johnell O, (2000).

As principais causas, apontadas por vários autores, para o aumento da incidência da osteoporose são o envelhecimento da população e o aumento da esperança média de vida. O facto da perda de massa óssea se tornar mais acentuada nas mulheres nos anos seguintes à menopausa leva a que estas tenham um maior risco de osteoporose e consequentemente um maior risco de vir a sofrer fracturas.

Como referido anteriormente, as características do tecido ósseo e as suas propriedades mecânicas alteram-se com a idade porque os mecanismos de construção (síntese de matriz óssea e mineralização subsequente) e de remodelação (o balanço entre a reabsorção e a formação óssea) sofrem uma progressiva degradação. A remodelação é uma função normal do osso e assegura o equilíbrio do metabolismo do cálcio e do fósforo e também a reparação de microdanos do osso. Durante o envelhecimento é depositado menos osso do que aquele que é removido, sofrendo este processo uma marcada aceleração após a menopausa, em que ainda mais osso é removido, a maior velocidade, de um tecido ósseo que já está a perder a sua arquitectura normal. Mas para além do balanço ósseo negativo, que se gera progressivamente com o envelhecimento e com a menopausa, o osso mais antigo, mais distante das superfícies de remodelação e mais mineralizado, acumula progressivamente pequenas lesões que não são reparadas,

enquanto o osso mais superficial, é substituído por um osso jovem, pouco mineralizado. O resultado de todas estas alterações é a redução da rigidez. Este efeito é parcialmente compensado através da aposição de osso no perióstio (membrana muito vascularizada, fibrosa e resistente, que envolve por completo os ossos, excepto nas articulações), que é mais significativo no homem do que na mulher. Além disso, no homem, o facto do esqueleto ser globalmente maior, permite tolerar uma carga superior, adiando o risco de sofrer fracturas. A relevância dos estrogénios no controle da remodelação óssea estende-se também ao homem. A redução da massa óssea no homem está essencialmente associada aos níveis de estrogénios e não aos de testosterona. Embora, isoladamente, a testosterona tenha um efeito anabólico sobre o osso, o efeito combinado dos estrogénios sobre a reabsorção e a formação do osso é mais determinante. Devido a este efeito modesto da testosterona, a perda de osso trabecular, com a idade, no homem, progride de uma forma linear, com o aparecimento de trabéculas mais delgadas. No entanto, a conectividade fica mais preservada do que na mulher, onde a perda de osso é mais determinada pela remoção de osso em cada unidade de remodelação (Canhão H, Fonseca J et al., 2005).

Estas razões justificam, de certa forma, a maior incidência de fracturas encontrada nas mulheres comparada com a encontrada nos homens, quando combinadas também com o facto das mulheres terem, em média, uma maior esperança média de vida. Vários estudos mostram que as mulheres representam 66-78% dos pacientes com fracturas do colo do fémur (Lesic A, Jarebinski M, 2006). As diferenças encontradas nas taxas de incidência por género variam, sendo na Europa duas a três vezes maior no género feminino do que no género masculino. No entanto existem algumas variações, por exemplo na Islândia encontrou-se um rácio homem:mulher de 1:4,2, dos mais altos na Europa (Lips P, 1997) ou em Malta onde o rácio é 1:(1,6-1,96), entre os mais baixos da Europa (Baron YM, Brincat M et al., 1994). Apesar de haverem mais mulheres com osteoporose ou com risco de sofrer uma fractura, são muitos os estudos que fazem referência ao facto de ser nos homens que as consequências são mais graves ocorrem.

O metabolismo ósseo, referido anteriormente, é um processo complexo que depende de um elevado número de factores, desde factores genéticos a factores externos como a dieta e a actividade física. A massa óssea obtida durante a infância e adolescência é talvez o mais importante determinante da saúde dos ossos ao longo da vida. Ao se

alcançar um bom nível de massa óssea mais cedo pode-se reduzir o impacto da perda de osso durante o processo de envelhecimento. Esta optimização da massa óssea pode ser mais facilmente conseguida com uma alimentação rica em cálcio e com exercício físico, pois há uma forte evidência que actividades físicas na infância e adolescência contribuem para aumentar o pico de massa óssea.

Desta forma, os hábitos alimentares são particularmente importantes: a fraca ingestão de leite e seus derivados e algumas verduras (principais fornecedores de cálcio) é um factor de risco. De igual modo, o consumo excessivo de álcool, chá, café e tabaco aumentam o risco de osteoporose. Certas condições médicas associadas, como artrite reumatóide ou o uso prolongado de cortisona, podem também aumentar este risco.

São vários os factores que contribuem para o aparecimento e prevalência da osteoporose e têm sido muitos os estudos feitos para os identificar e os relacionar entre si. Alguns dos factores mais referidos são: factores hormonais, como a concentração de estrogénios e androgénios; a pouca actividade física; baixo consumo de cálcio e vitamina D, o que conduz a uma diminuição da densidade óssea e ao aumento do risco de quedas e de fragilidade (Diamond T, Eisman J et al., 2005); o fumo de tabaco, que reduz a absorção intestinal de cálcio (Lane N, 2006); o álcool e a cafeína; a raça, os números indicam que, por exemplo, nos indivíduos brancos existe uma maior taxa de incidência desta doença do que nos indivíduos negros; o peso e a altura; e a história pessoal ou familiar relativamente a fracturas.

2.3 Epidemiologia da Osteoporose

Como mencionado anteriormente, a epidemiologia é o estudo de doenças em populações, em oposição ao estudo das doenças nos indivíduos e as suas principais preocupações são relacionadas com a prevalência e incidência das doenças, com a severidade das mesmas (no sentido dos distúrbios causados na comunidade) e com a responsabilidade da comunidade (no sentido da procura de formas de prevenção e tratamento). A epidemiologia da osteoporose levanta questões especiais, uma vez que a doença só por si não acarreta muitas “queixas”. As pessoas que sofrem de osteoporose não procuram ajuda médica até alguma consequência aparecer, normalmente, fracturas, ou a menos habitual perda de peso e dores nas costas, pois, tanto a perda de massa óssea como a deterioração da microarquitECTURA óssea são silenciosas (Wolf A, Dixon A, 1988).

Actualmente, a osteoporose é um problema que afecta mais de 10 milhões de pessoas nos Estados Unidos da América, estando previsto que, aproximadamente, 14 milhões de mulheres venham a sofrer osteoporose no ano 2020 (Lane N, 2006). Mundialmente, este problema afecta aproximadamente 200 milhões de adultos com mais de 50 anos (Lane N, 2006).

A consequência clínica da osteoporose são as fracturas (Kanis J, Delmas P et al., 1997). As fracturas típicas ocorrem no colo do fémur, vértebras e no punho (ver **Figura 2.4**), mas quando o esqueleto sofre de osteoporose outras fracturas também ocorrem mais frequentemente (Kanis J, Delmas P et al., 1997). O risco de ter uma fractura osteoporótica ao longo da vida é muito alto, varia entre 40% e 50% nas mulheres e é aproximadamente 25% nos homens (Johnell O, Kanis J, 2004). E, segundo Schutte H (1995), o risco para mulheres na menopausa, varia entre 30 e 40%.

Para além de consequências físicas, como dor, perda de peso, deformidade e mortalidade, as fracturas osteoporóticas têm também consequências psicológicas como depressão, perda de auto-estima, medo, raiva e restrição das relações interpessoais (Lane N, 2006).

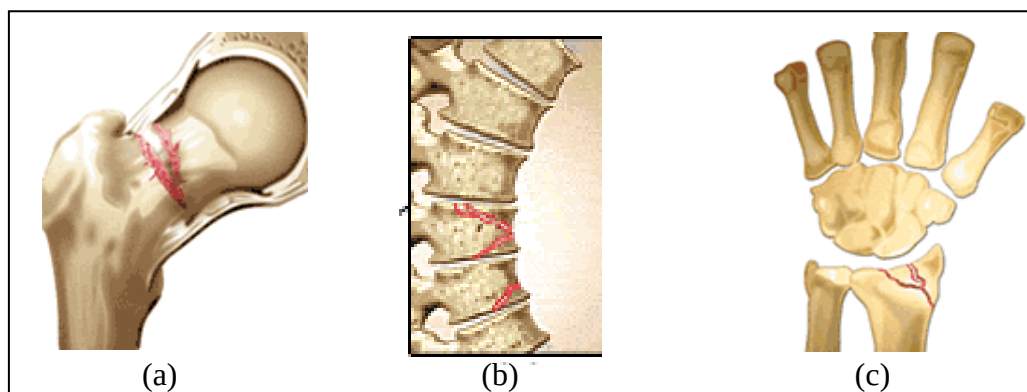


Figura 2.4 (a) fractura do colo do fémur; (b) fractura das vértebras; (c) fractura do punho.

Fonte: < <http://www.imot.com.br/osteoporose.html>>.

As fracturas osteoporóticas têm elevados custos para a sociedade, quer directos, quer indirectos, uma vez que qualquer uma destas fracturas implica algum tempo de imobilização, perda da qualidade de vida, alguma incapacidade e, no caso das fracturas de colo do fémur, elevada mortalidade nos primeiros três meses após a fractura.

Relativamente aos custos que esta doença acarreta é difícil exprimir-se o valor exacto uma vez que englobam cuidados hospitalares, perda de dias de trabalho, custo em instituições que prestam cuidados de saúde, que variam entre países e entre as regiões do próprio país. No entanto, está estimado que para os EUA o custo das fracturas osteoporóticas varia entre 7 a 10 biliões de dólares, para uma população de 250 milhões de pessoas; em Inglaterra e Gales está estimado o valor de 614 milhões de libras por ano, para uma população de 52 milhões de pessoas; e para a França aproximadamente 3,7 biliões de francos para uma população de perto de 50 milhões de pessoas (Schutte H, 1995).

2.3.1 Epidemiologia das Fracturas Vertebrais

As fracturas das vértebras apesar de terem sido reconhecidas desde a antiguidade, só em 1941 começaram a ser consideradas como um dos sinais de osteoporose, depois de provado que a maioria destas fracturas ocorria espontaneamente ou após traumatismos mínimos (Dias A, Ventura S et al., 1995). Estas fracturas são das mais difíceis de identificar e embora sejam das menos estudadas têm consequências graves a nível funcional e estético.

Apesar de serem umas das fracturas osteoporóticas que ocorrem com mais frequência (Finnern H, Sykes D, 2003) e uma das mais adversas consequências da osteoporose, existem poucos dados relativos à incidência destas fracturas, tanto em homens como em mulheres (Felsenberg D, Silman AJ et al., 2002). As fracturas vertebrais constituem um dos sinais clássicos da osteoporose e são consideradas cada vez mais um problema de saúde pública por terem como consequências: dores (agudas ou crónicas) nas costas, incapacidade, deformidades, perda de peso, aumento de morbilidade, mortalidade e claras alterações na qualidade de vida (Felsenberg D, Silman AJ et al., 2002 e Dias A, 2000). Enquanto que nas fracturas do colo do fémur a mortalidade é particularmente elevada nos primeiros meses após a fractura, nas fracturas das vértebras sintomáticas a mortalidade aumenta progressivamente com o decorrer dos anos.

À medida que a osteoporose se desenvolve, vai existir perda de densidade óssea no centro da vértebra, comparando com o osso cortical que a envolve. Estas vértebras sofrem micro fracturas e ficam deformadas (ver **Figura 2.5**). Este acontecimento nem sempre envolve dores ou algum tipo de trauma e estas deformidades são detectadas radiologicamente. O problema reside em não existir consenso relativamente à definição radiológica da fractura da vértebra, o que torna difícil a sua caracterização epidemiológica (Wolf A, Dixon A, 1988; Dias A, 2000).

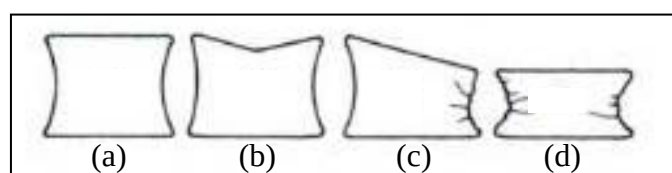


Figura 2.5 (a) vértebra normal; vértebras com deformidades: (b) vértebra bicôncava ou em boca de peixe, (c) vértebra em cunha e (d) vértebra por esmagamento ou em bolacha.

Fonte: Wolf A, Dixon A, (1988).

Em contraste com as fracturas do colo do fémur existem poucos dados relativos à incidência das fracturas das vértebras. Este facto pode ser justificado porque as pessoas afectadas com estas fracturas não recebem tratamento hospitalar e por nem todas estas fracturas serem sintomáticas (Felsenberg D, Silman AJ et al., 2002; Finnern H, Sykes D, 2003). No entanto as fracturas das vértebras são um forte indicador de futuras fracturas noutros locais do esqueleto (Finnern H, Sykes D, 2003 e Schutte H, 1995).

Como estas deformidades não são diagnosticadas até se tornarem casos clínicos, a exacta dimensão desta epidemia não é conhecida. No entanto, há estimativas que mostram que a incidência em mulheres brancas americanas, com 50 anos ou mais, é de 18/1000 (Schutte H, 1995). Na Europa, Felsenberg D, Silman AJ et al. (2002), verificaram que a incidência destas fracturas aumenta com a idade em ambos os géneros, mas que é maior nas mulheres do que nos homens em todas as idades (ver **Figuras 2.6 e 2.7**). E tal como em outras fracturas osteoporóticas encontraram uma maior taxa de incidência na Suécia e nos países Escandinavos do que em outras regiões da Europa. Apesar das diferenças geográficas encontradas, estas são menos marcadas do que as encontradas ao nível das fracturas do colo do fémur. Quanto ao futuro, Felsenberg D, Silman AJ et al. (2002) prevêem que, devido às mudanças demográficas, ou seja, ao envelhecimento da população, a incidência das fracturas das vértebras vai continuar a aumentar, mantendo-se um importante problema de saúde pública.

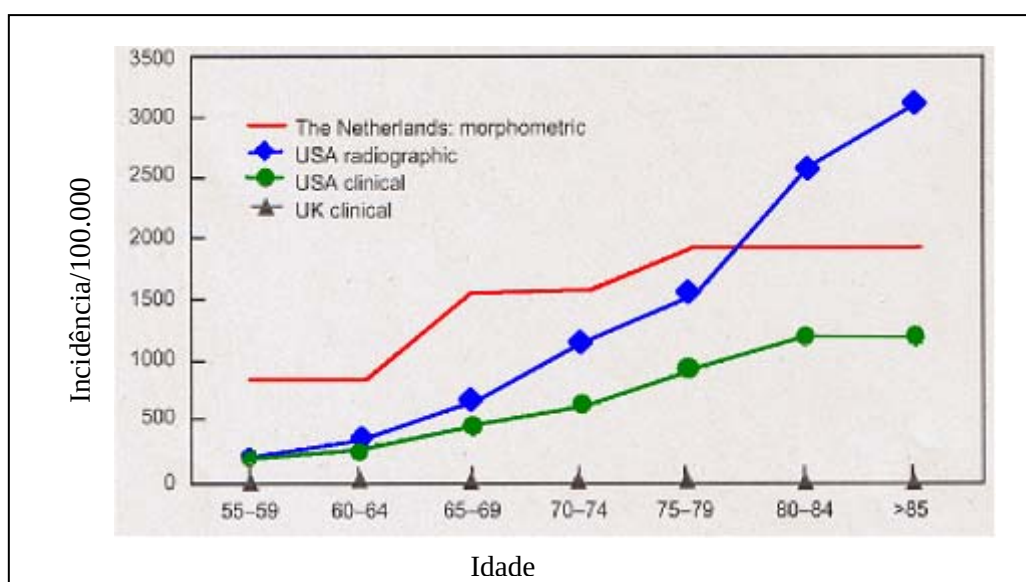


Figura 2.6 Incidência cumulativa, durante um ano, de fracturas vertebrais em mulheres, por 100.000, nos EUA, no Reino Unido e Países Baixos.

Fonte: Laet C, Pols H, (2000).

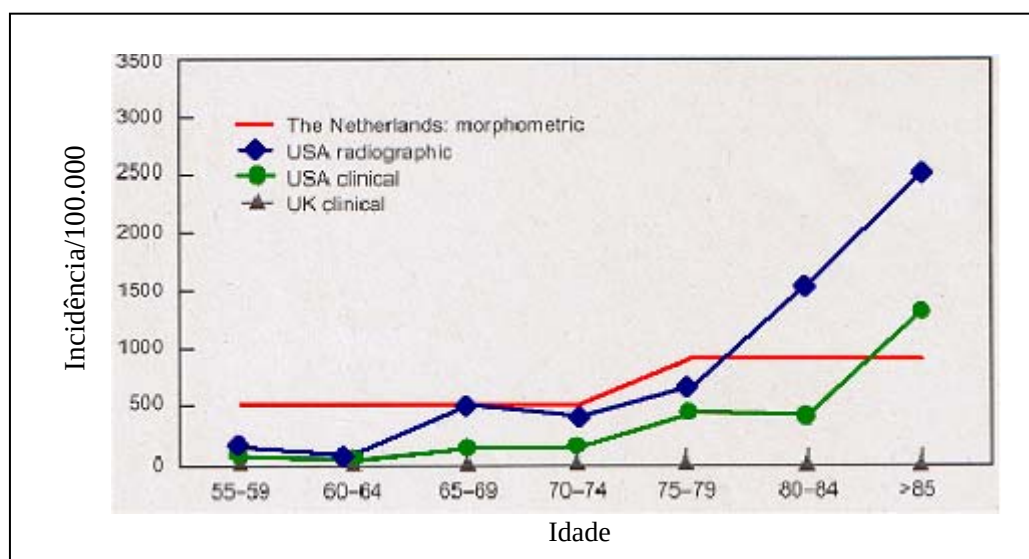


Figura 2.7 Incidência cumulativa, durante um ano, de fracturas vertebrais em homens por 100.000, nos EUA, no Reino Unido e Países Baixos.

Fonte: Laet C, Pols H, (2000).

Segundo Dias A (2000), na amostra populacional portuguesa que participou no *European Prospective Osteoporosis Study* (EVOS) a prevalência das deformações vertebrais mostrou-se ligeiramente acima da média europeia e extrapolando estes valores para a população portuguesa com idade superior a 50 anos, estima-se que pelo menos 360.000 indivíduos apresentem actualmente estas deformações.

Schutte H (1995) refere que apenas um terço das fracturas das vértebras chegam ao conhecimento dos médicos e que apenas entre 2 a 8% necessita de admissão hospitalar. Este autor realça ainda que as medidas feitas em radiografias de mulheres com deformidades vertebrais, que procuraram ajuda médica, revelam que 80% têm valores superiores a 4 desvios-padrão em relação à média das mulheres jovens, consideradas saudáveis.

2.3.2 Epidemiologia das Fracturas do Punho

As fracturas do punho ou fracturas da extremidade distal do antebraço (ou do tipo *Colles*) são normalmente causadas por quedas sobre um braço estendido e também se encontram relacionadas com a osteoporose. Tipicamente estas fracturas ocorrem maioritariamente em mulheres sendo, segundo Dias A (2000), o rácio homem:mulher igual a 1:4 e, quando comparadas com as fracturas do colo do fémur, ocorrem em indivíduos mais jovens (Melton L, Achenbach S et al., 2002). Apesar de ocorrerem nas mulheres antes e após a menopausa, a DMO está habitualmente reduzida nas mulheres pré-menopáusicas, sugerindo a importância do pico de massa óssea na determinação da susceptibilidade para estas fracturas (Dias A, 2000).

Apesar das fracturas do punho causarem muito menos morbilidade do que as fracturas do colo do fémur as suas consequências são muitas vezes subestimadas (Dias A, 2000). Apenas uma pequena percentagem das fracturas do punho requer hospitalizações e como consequência poucos estudos epidemiológicos têm sido feitos sobre a etiologia destas fracturas, principalmente nos homens (Melton L, Achenbach S et al., 2002). Os estudos existentes, no entanto, sugerem que a incidência destas fracturas tem vindo a aumentar progressivamente ao longo do tempo (Dias A, 2000).

Também nas taxas de incidência das fracturas do punho são encontradas diferenças geográficas. Como Dias A (2000) refere no seu estudo, a incidências destas fracturas apresenta, geograficamente, um marcado paralelismo com as fracturas do colo do fémur. No nível Europeu e semelhante a outras fracturas osteoporóticas, os países do norte da Europa, nomeadamente os países Escandinavos, são os que têm maiores taxas quando comparados com os outros países (EPOS, 2002). No que respeita aos EUA existem estudos que mostram diferenças, principalmente em longitude, nas taxas de incidência das fracturas, sendo estas mais altas nos estados do este do que nos estados do oeste (ver **Figura 2.8** e **2.9**). No que diz respeito às diferenças raciais estas fracturas têm uma incidência mais reduzida nas populações de raça negra (Dias A, 2000).

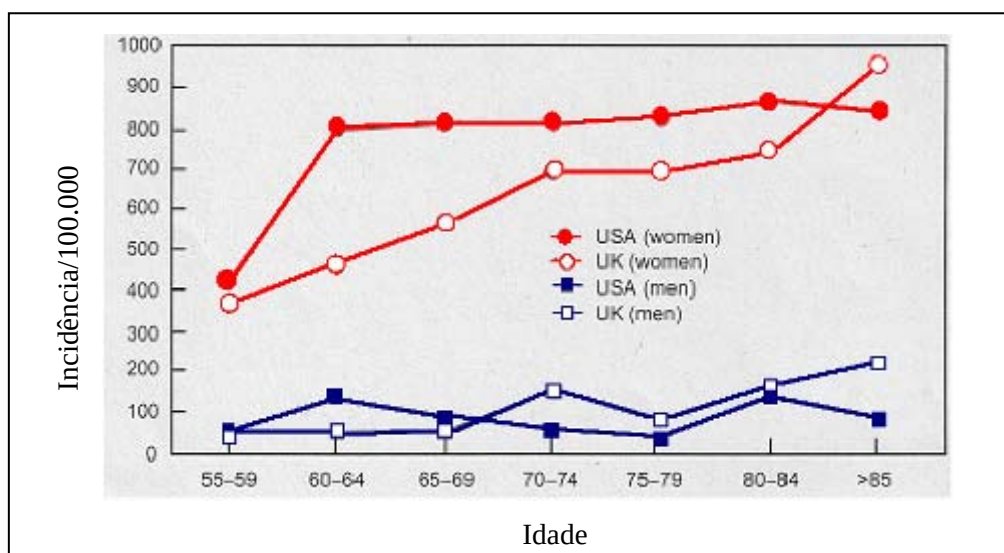


Figura 2.8 Incidência cumulativa, durante um ano, de fracturas do punho em homens e mulheres por 100.000 pessoas-anos, nos EUA e no Reino Unido.

Fonte: Laet C, Pols H, (2000).

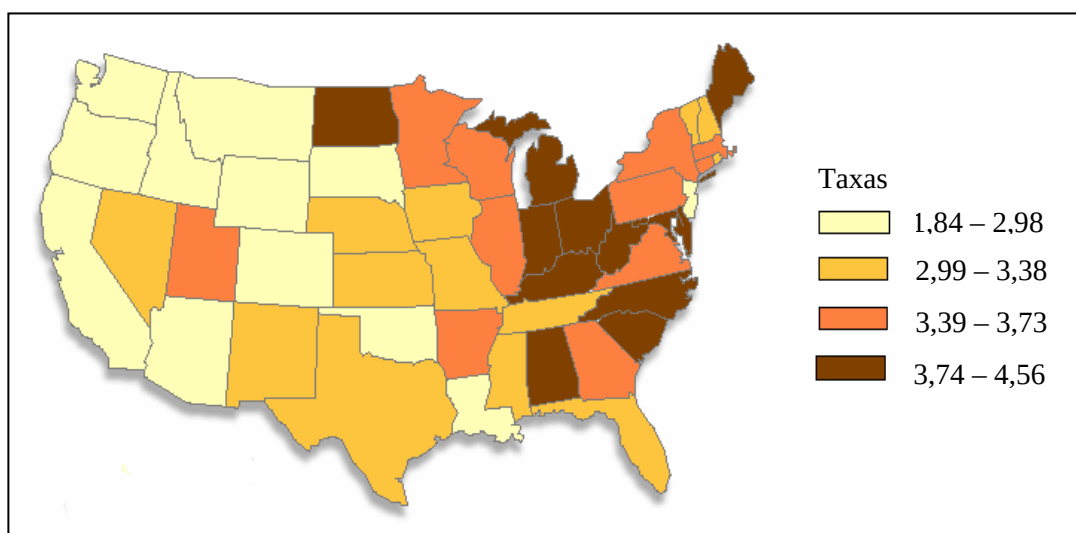


Figura 2.9 Mapa das taxas de fracturas do punho (por 1000 habitantes) dos EUA entre 1986-1990, de indivíduos com 65-89 anos.

Fonte: Karagas MR, Baron JA et al., (1996).

2.3.3 Epidemiologia das Fracturas do Colo do Fémur

As fracturas osteoporóticas que ocorrem nas vértebras e no punho estão associadas a um aumento acentuado da morbilidade, mas as fracturas que conduzem a consequências clínicas mais graves são as fracturas do colo do fémur, que estão associadas a um aumento de mortalidade, particularmente em indivíduos idosos (Kanis J, Delmas P et al., 1997). Como já referido, as fracturas do colo do fémur ocorridas em indivíduos com mais de 50 anos e causadas por baixo impacto são, frequentemente, usadas como um indicador de osteoporose (Lane N, 2006 e Soveid M, Serat A et al., 2005).

Segundo Marcus R, Feldman D et al. (1996), o risco de sofrer uma fractura osteoporótica do colo do fémur é cerca de 15% para as mulheres caucasianas e cerca de 6% para os homens caucasianos ao longo da vida.

O aumento das taxas de incidência das fracturas do colo do fémur com a idade está fortemente associado ao aumento da incidência da osteoporose. Em muitos países, sobretudo de população caucasiana (Alemanha, Suécia, Suíça, Grécia, Servia e Montenegro, Portugal, entre outros) verifica-se um aumento exponencial das taxas com a idade em ambos os géneros, em contraste com as populações asiáticas, onde se verifica também um crescimento mas de forma linear (Jéquier V, Burnand B et al., 1995) (ver **Figura 2.10**).

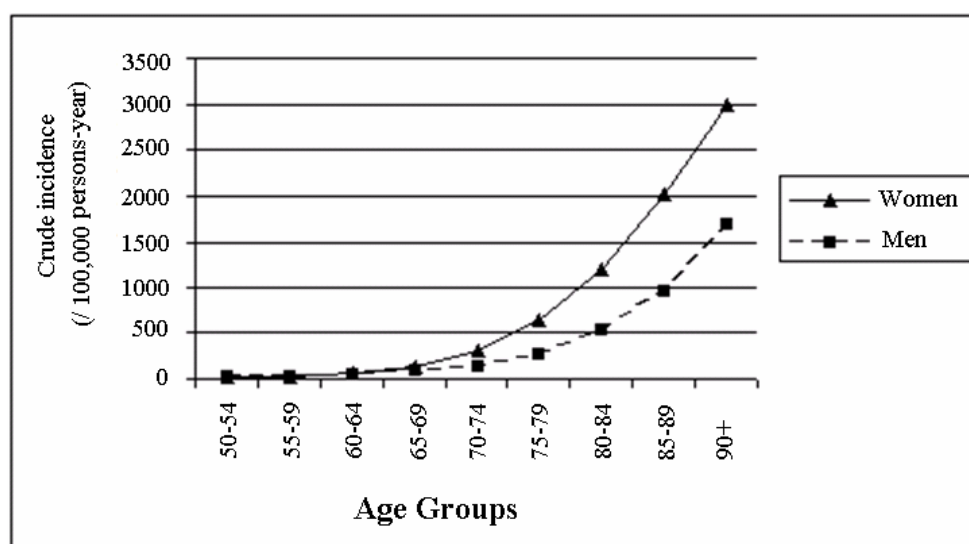


Figura 2.10 Fracturas osteoporóticas do colo do fémur (acima 50 anos de idade), em Portugal no período de 2000 a 2002.

Fonte: Pina MF, Alves S et al., (in Press).

As fracturas do colo do fémur na população mais envelhecida estão associadas à diminuição da DMO e ao aumento da tendência para quedas, por sua vez relacionadas à pouca força muscular, pior visão, menor equilíbrio, uso de múltiplos medicamentos e psicotrópicos (Kannus P, Niemi S et al., 1999) (ver **Tabela 2.2**).

Tabela 2.2 Factores de risco associados a fracturas do colo do fémur, para a população feminina caucasiana, com mais 65 anos.

Variável	Percentagem do aumento do risco
Idade (por 5 anos)	40%
História de fractura por parte da mãe	80%
Peso na idade de 25 anos	30%
História de hipertiroidismo	60%
Uso corrente de benzodiazepinas	70%
Uso corrente de drogas anticonvulsivas	100%
Incapacidade de se levantar de uma cadeira	70%
Pouca percepção de profundidade (na visão)	40%
Pulsação superior a 80 batimentos/s numa situação de descanso	70%
Qualquer fractura a partir dos 50 anos	50%

Observação: Os números indicam quão maior é a possibilidade de se sofrer uma fracturas quando estes factores estão presentes.

Fonte: Baseado em Cummings SR et al., (1995).

Como Laet C, Pols H (2000) mencionam, estas fracturas ocorrem em todo o mundo, mas com uma maior taxa de incidência em países ocidentais, nomeadamente na Europa e nos EUA. Foi estimado que o número total de fracturas do colo do fémur a nível mundial, em 1990, era 1,7 milhões, destes 560.000 na Europa e 360.000 na América do Norte (Laet C, Pols H, 2000).

No entanto, tanto na Europa como nos EUA existem diferenças nas taxas de incidência destas fracturas entre os vários países ou estados. Existem estudos que mostram que, por exemplo, a nível europeu há diferenças entre os países do norte e do sul, chegando mesmo a existir uma variação na incidência das fracturas osteoporóticas de até 7 vezes entre os diversos países. Em relação ao EUA as taxas mais altas são encontradas nos estados do sul quando comparados com os estados do norte (ver **Figuras 2.11 a 2.13**).

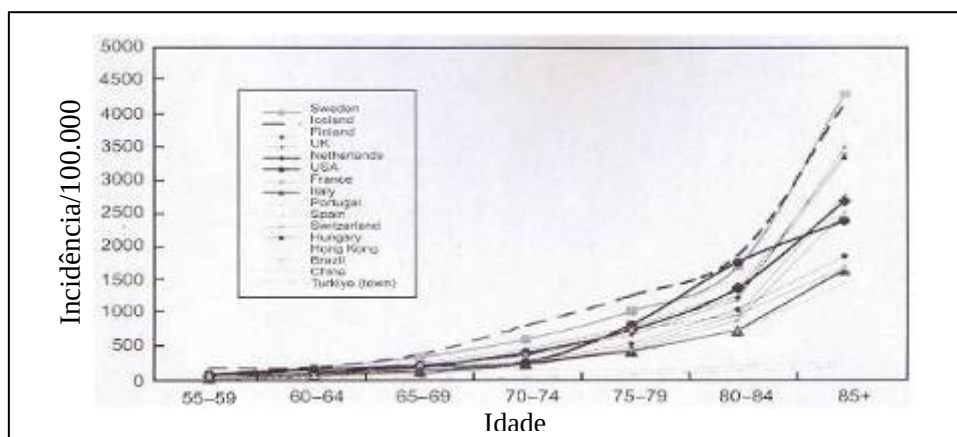


Figura 2.11 Incidência cumulativa, durante um ano, de fracturas do colo do fémur em mulheres por 100.000 habitantes em vários países.

Fonte: Laet C, Pols H, (2000).

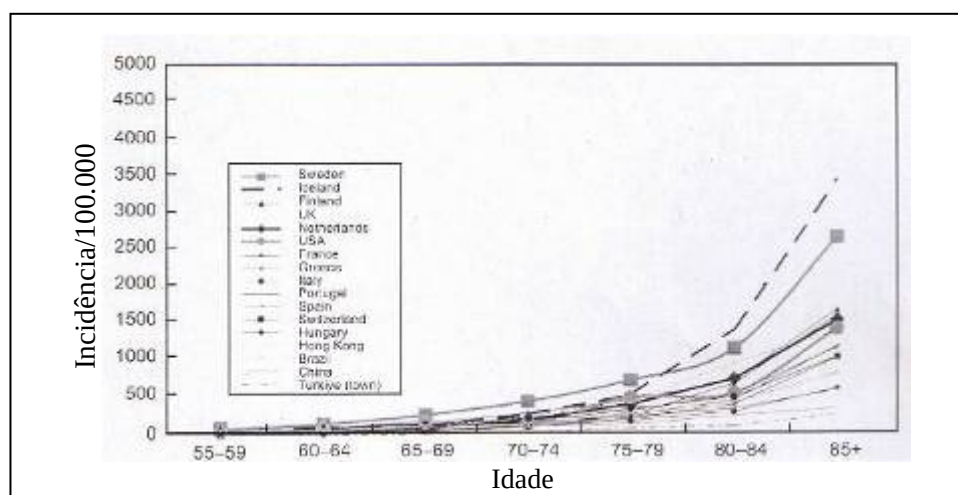


Figura 2.12 Incidência cumulativa, durante um ano, de fracturas do colo do fémur em homens por 100.000 habitantes em vários países.

Fonte: Laet C, Pols H, (2000).

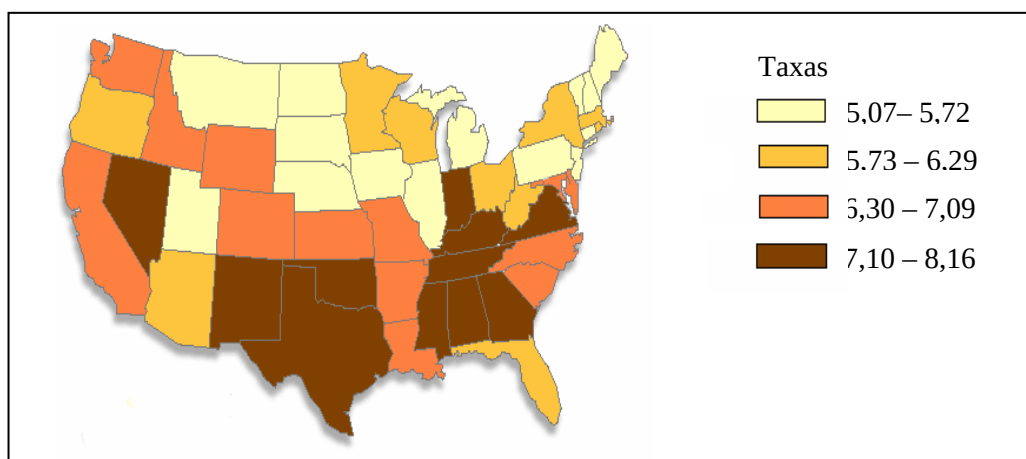


Figura 2.13 Mapa das taxas de incidência de fracturas do colo do fémur (por 1000 habitantes) nos EUA entre 1986-1990, de indivíduos com 65-89 anos.

Fonte: Karagas MR, Baron JA et al., (1996).

São vários os estudos, como o *Mediterranean Osteoporosis Study* (MEDOS), que referem a existência de diferenças entre os países do norte e do sul da Europa, em relação às taxas de incidência das fracturas osteoporóticas, tendo a Escandinávia a maior taxa de incidência e os países mediterrânicos as menores taxas. No MEDOS participaram indivíduos com idade igual ou superior a 50 anos de 14 centros de 6 países da Europa Meridional e os resultados relativos à incidência das fracturas do colo do fémur podem ser vistos na **Tabela 2.3**.

Tabela 2.3 Incidência por 100.000 habitantes, das fracturas do colo do fémur em populações com 50 anos ou mais da Europa Meridional (MEDOS).

Centro*	Homens	Mulheres	Rácio Homem:Mulher
Porto	114	259	1:2,28
Sevilha	98	370	1:3,79
Toulouse	108	366	1:3,39
Parma	105	257	1:2,45
Creta	179	347	1:1,94
Istambul	26	34	1:1,28
Turquia Rural	62	23	1:0,37

*Representados apenas alguns centros MEDOS.

Fonte: Dias A, (2000).

A Suécia e a Noruega são os países em que são encontradas as mais altas taxas de incidência de fracturas do colo do fémur, tanto para as mulheres como para os homens, sendo respectivamente (por 100.000 habitantes), 1380 e 690 em Malmo, Suécia; 1250 e

620 em Hassleholm-Kristianstad, Suécia (Billsten M, Sernbo I et al., 2000) e 1180 e 440 em Oslo - Noruega (Lodthus C, Osnes E et al., 2001). Em Kaunas, Lituânia, foi encontrada uma elevada taxa de incidência das fracturas de colo do fémur (por 100.000 habitantes) para os homens, sendo 470, comparativamente a 676 encontrada nas mulheres (Lashas L, Masanauskaite D et al., 1996). As taxas mais baixas foram encontradas na Sérvia e Montenegro, respectivamente para homens e mulheres (por 100.000 habitantes) 35,5 e 62,2 (Lesic A, Jarebinski M et al., 2006).

Comparando as taxas de vários países da Europa com a taxa dos EUA é possível perceber que, por exemplo, a Suécia tem uma taxa superior aos EUA, a Finlândia tem uma taxa idêntica para as mulheres mas ligeiramente superior para os homens, em países como Reino Unido, Países Baixos e Alemanha as taxas são similares e que nos países do sul, nomeadamente Portugal e Itália, a incidência é inferior quando comparada com a do EUA (Laet C, Pols H, 2000).

Relativamente aos EUA, para além das diferenças já referidas nas taxas de incidência das fracturas osteoporóticas entre as várias regiões, existem grandes diferenças entre estas taxas nas populações de diferentes raças. Alguns estudos mostram que, por exemplo, a população caucasiana tem taxas de incidência de fracturas de colo do fémur mais elevadas quando comparadas com a população negra ou que na população hispânica entre os 50 e 59 anos as taxas de incidência são mais elevadas nos homens (Karagas MR, Baron J et al., 1996).

Menos informação existe sobre os outros continentes, no entanto, tanto na América do Sul como na Ásia a taxa de incidência das fracturas do colo do fémur é menor do que na América do Norte e na Europa (Laet C, Pols H, 2000).

Em África, a osteoporose é conhecida como sendo uma doença rara e como tal nenhuma medida preventiva ou curativa tem sido tomada. Num estudo realizado em Camarões, Zebaze R, Seeman E (2003) encontraram que, numa população com idades iguais ou superiores a 65 anos, a taxa ajustada de incidência das fracturas do colo do fémur é 24,4 nas mulheres e 20,7 nos homens (por 100.000 habitantes). Além disso mostraram que, considerando apenas as fracturas de baixo impacto, a idade média é de $72,6 \pm 9,4$ nas mulheres e de $66,2 \pm 5,8$ nos homens. No entanto cabe realçar que menos de 3% da população de Camarões tem mais de 65 anos e a esperança média de vida naquele país é de apenas 57 anos.

Em relação à Ásia prevê-se que até ao ano 2050, 50% de todas as fracturas do colo do fémur no mundo ocorram neste continente. Estudos recentes, que focam a existência de variações geográficas acentuadas na incidência das fracturas de colo do fémur, demonstram que as taxas de incidência em países asiáticos são mais baixas quando comparadas com as encontradas na Europa ou nos EUA. Num estudo realizado em Shenyang (República da China), Yan L, Zhou B et al. (1999) encontraram que, para a população com idades iguais ou superiores a 50 anos, a taxa ajustada de incidência das fracturas do colo do fémur é 67 nas mulheres e 81 nos homens (por 100.000 habitantes). Encontram também um rácio homem:mulher de 1,21:1 e idades médias de $66,9 \pm 9,0$ anos nas mulheres e de $67,5 \pm 9,8$ anos nos homens.

Como a osteoporose está relacionada com o envelhecimento da população e com vários factores de risco, tem-se verificado um aumento nas taxas de incidência e espera-se um aumento especialmente acentuado nos países em desenvolvimento à medida que a longevidade da população aumentar (Lane N, 2006). Apesar de já existirem referências a estudos recentes que revelam uma redução nas taxas de incidência (nomeadamente, no Canadá e na Suécia), na maioria dos países ocidentais verificou-se durante a segunda metade do século um aumento nas taxas de incidência da osteoporose.

Dias A (2000) chega mesmo a referir que mais importante que as diferenças populacionais encontradas na incidência das fracturas de colo do fémur, será talvez o facto do número de fracturas osteoporóticas do colo do fémur estar a aumentar em praticamente todo o mundo, mais acentuada e rapidamente nos países com um acelerado ritmo de desenvolvimento, sobretudo em países asiáticos e da América Latina. Este aumento varia, por ano, entre 0,5% e 3,3% (Reginster JY, Gillet P et al., 2001) em países como a Finlândia, Países Baixos, Noruega, Suécia e Reino Unido. No entanto, na Espanha, onde as taxas são mais baixas do que as encontradas no Norte da Europa, verificou-se que em 14 anos (entre 1988 e 2002) houve um aumento de 4% ao ano (Hernández JL, Olmos JM et al., 2005).

Apesar das altas taxas de incidência de fracturas do colo do fémur encontradas nos países do Norte da Europa (Noruega, Suécia ou Finlândia), estes países têm apresentado uma tendência de decréscimo na ocorrência destas fracturas, talvez devido à prevenção e tratamento da osteoporose bem sucedidos (Parkkari J, Kannus P et al., 1994).

Esta comparação das taxas de incidência em diferentes países deve ser feita de forma cuidadosa, devido à diferença entre métodos e dados. Os estudos realizados tanto sobre osteoporose como sobre as fracturas osteoporóticas variam entre retrospectivos e prospectivos, sendo que as durações dos estudos retrospectivos variam de 1, 2, 5 até 12 e 14 anos. Além disso os dados obtidos são recolhidos de fontes diversificadas, desde registos hospitalares e registos radiológicos a simples questionários (existindo estudos em que, por exemplo, apenas 6% da população está representada (Van Stan TP, Dennison EM et al., 2001), o que por si só não é necessariamente um problema, se se tratar uma amostra representativa). A quase totalidade dos estudos incide nas fracturas de colo do fémur tratadas em hospitais. As repetições nas contagens/registos dos pacientes (que podem mudar de hospital ou mudar de serviço dentro do mesmo hospital) também são uma diferença a considerar, sendo que nem todos os autores excluem as contagens múltiplas, caso existam. Outras diferenças significativas são: o corte da idade em que se inicia o estudo, este corte nem sempre existe e quando tal acontece varia entre os 20 anos e os 65 anos (Billsten M, Sernbo I et al., 2000), sendo, no entanto, na maioria dos estudos os 50 anos; a região geográfica a que o estudo se refere, por exemplo se a territórios específicos (como uma cidade), se regiões e países, ou vários países; e por fim, se os autores excluem as fracturas resultantes de condições patológicas.

Em Portugal, apesar da taxa de incidência de fracturas osteoporóticas do colo do fémur ser menos da metade da taxa de incidência apresentada pelos EUA (Laet C, Pols H, 2000) espera-se um aumento acentuado no futuro. Dias A (2000) refere no seu trabalho, que em Portugal foram realizados quatro estudos epidemiológicos sobre a incidência de fracturas do colo do fémur entre 1981 e 1988 (dois no Porto, um em Lisboa e um na Ilha da Madeira) e, que a incidência destas fracturas varia entre 125 e 297 (por 100.000 habitantes) no género feminino e entre 81 e 136 (também, por 100.000 habitantes) no género masculino.

Com relação a diferenças geográficas em Portugal, Araújo D, Barros H et al. (1995) afirmam que a incidência das fracturas é maior no interior e no sul e menor no norte e no litoral. Embora não esteja claro no artigo qual o método de medição da osteoporose, quais as fracturas utilizadas e qual o padrão geográfico. Num outro trabalho, Alves S (2004) obteve resultados que mostram que as regiões com as mais altas taxas de

incidência, são: Algarve, Lisboa e arredores, Trás-os-Montes e alguns concelhos do Centro. O Norte, Alentejo e interior da região Centro têm as taxas de incidência mais baixas (ver **Figura 2.14**). Em Portugal como o Sistema de Saúde é público e universal, quase a totalidade de casos de fracturas do colo do fémur são tratados em hospitais públicos e desta maneira a base de dados utilizada no trabalho de Alves S (2004) tem uma abrangência nacional.

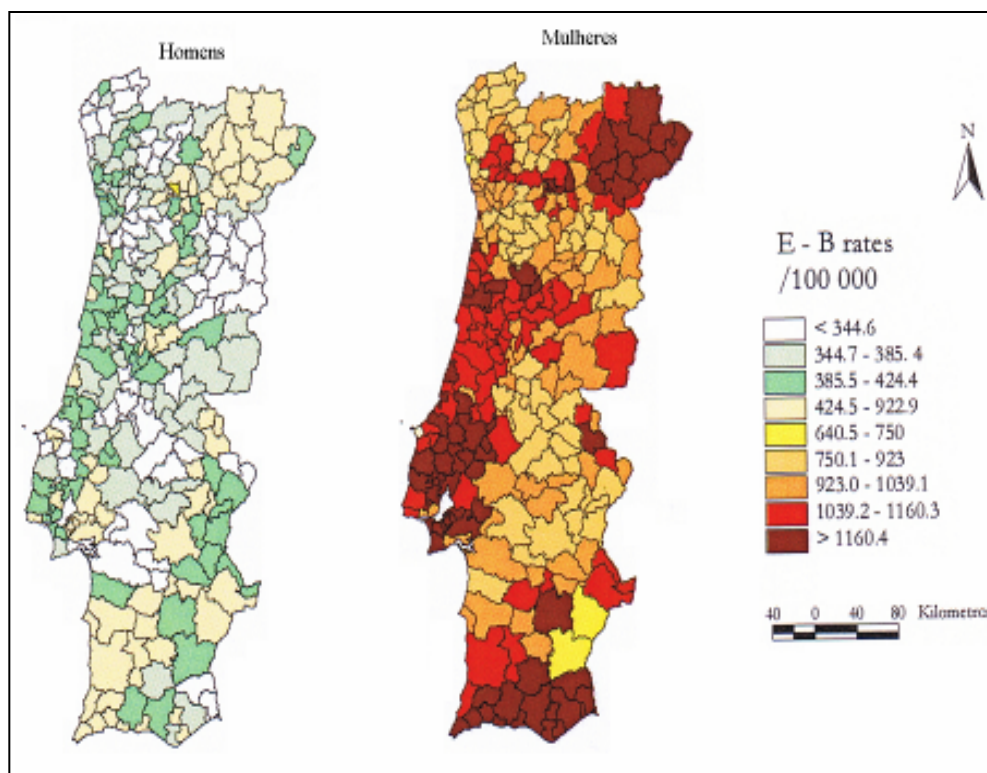


Figura 2.14 Distribuição geográfica das taxas de incidência de fracturas do colo do fémur.

Fonte: Alves S, (2004).

Ainda segundo Dias A (2000) no centro MEDOS do Porto, a incidência das fracturas encontrada foi relativamente elevada e, como seria de esperar, aumentava exponencialmente com a idade (ver **Figura 2.15**).

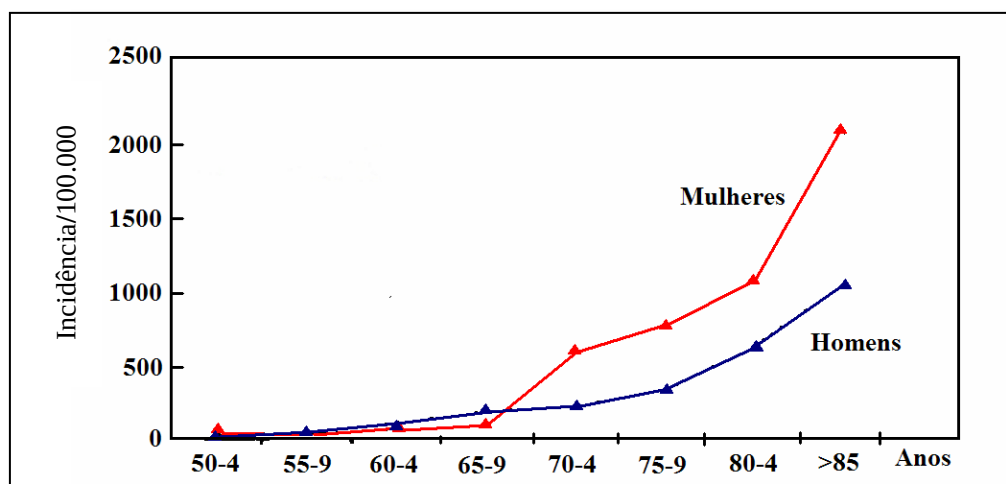


Figura 2.15 Aumento exponencial da incidência da fracturas do colo do fémur com a idade, em ambos os géneros, no centro MEDOS do Porto.

Fonte: Dias A, (2000).

As diferenças geográficas encontradas existem, possivelmente, porque a área geográfica representa uma combinação única de factores de risco (sejam eles factores socioeconómicos, genéticos, estilos de vida ou dieta alimentares), podendo ou não estas diferenças coincidirem com fronteiras administrativas.

Estas diferenças geográficas podem ainda, segundo alguns autores, estar relacionadas com o clima, pelo que as baixas temperaturas e o gelo podem contribuir para as altas taxas de incidência encontradas nos países do Norte da Europa. Apesar do aumento da ocorrência das fracturas no Inverno já ter sido várias vezes descrito (talvez devido à pouca luz solar ou à existência de gelo, que facilita as quedas), alguns estudos recentes mostram que não existem grandes variações das taxas de incidência com as condições climáticas, principalmente entre as mulheres (Giversen I, 2006 e Lodthus C, Osnes E et al., 2001). Um facto que permite explicar esta inexistência de correlação entre o clima, ou as temperaturas mais baixas e as taxas de incidência das fracturas do colo do fémur, é que na maioria dos casos as fracturas ocorrem dentro de casa (o que pode indicar a limitada actividade física dos pacientes). Lonnroos E, Kautiainen H et al. (2006), num estudo realizado na Finlândia, encontraram que as quedas de baixo impacto, numa superfície plana em casa ou no jardim, foram a causa das fracturas do colo do fémur em 65,6% dos pacientes hospitalizados. Estes autores afirmam ainda que, analisando as séries temporais das taxas de incidência, talvez seja possível justificar as

flutuações sazonais existentes, no entanto salientam que estas flutuações são mais comuns em países de climas quentes ou de clima moderado.

Para além das diferenças entre géneros e as diferenças sazonais já referidas, existem outros estudos que, por exemplo, evidenciam as diferenças entre as taxas das fracturas osteoporóticas em meios urbanos e em meios rurais, sendo as maiores taxas encontradas em meios urbanos (Billsten M, Sernbo I et al., 2000; Soveid M, Serat A et al., 2005). Estes estudos mostram: que nos meios rurais o estilo de vida envolve mais actividade física o que promove a força muscular e uma postura mais estável; que existe uma grande diferença entre as dietas destes dois meios pelo que em meios urbanos há uma maior deficiência de vitamina D; e que em meios urbanos há uma maior coexistência de problemas de saúde e maior uso frequente de drogas. Outro factor relacionado com estas diferenças encontradas é que nos meios urbanos existem mais mulheres a viverem sozinhas.

Alguns estudos relacionaram factores sociais, como o casamento, e factores socioeconómicos com o risco de sofrer fracturas osteoporóticas, tendo sido concluído que, tanto o baixo nível de escolaridade como o não estar casado está associado ao aumento do risco de vir a sofrer uma fractura do colo do fémur. Num estudo realizado na Suécia, Farahmand B, Persson PG et al. (2000) mostraram que o risco de sofrer estas fracturas era menor em mulheres que tinham estado empregadas nas duas décadas anteriores e que tinham maiores rendimentos. Este estudo também chega a conclusões semelhantes a outras já referidas anteriormente, que as fracturas do colo do fémur, nas mulheres, são menos comuns em meios rurais (Sawka A, Thabane L et al., 2005).

Os factores socioeconómicos, nomeadamente a industrialização e urbanização e a ocidentalização das sociedades estão relacionados com o aumento da incidência da osteoporose. Num estudo realizado na Alemanha, Wildner M, Casper W et al. (1999) concluíram que na Alemanha do Leste (antiga RDA), apesar das taxas de incidência da fractura do colo do fémur serem cerca de 10% mais baixas nas mulheres e 20% mais baixas nos homens em comparação com a Alemanha Ocidental (antiga RFA), a incidência duplicou durante o período de 1971 e 1996. Apesar da causa para este grande aumento, quando comparado com o da Alemanha Ocidental, não ter sido identificada, os autores levantaram a hipótese de ser devido à influência significativa do estilo de

vida ocidental (ou seja, mudança de meios de transporte, habitações, disponibilidade de comida e outros bens).

Tal como outras doenças crónicas, a osteoporose é um problema social significativo (Crabtree N, Lunt M et al., 2000) cujas consequências clínicas, custos sociais e económicos evidenciam a necessidade de prevenção e intervenção desde o diagnóstico (Lane N, 2006). Como já referido anteriormente, as fracturas do colo do fémur são uma das mais sérias consequências da osteoporose, conduzindo a hospitalizações prolongadas e quase invariavelmente à necessidade de uma intervenção cirúrgica (tendo como consequência, por exemplo a alta ocupação de camas ortopédicas - na Grécia, Paspatis I, Galanos A et al. (1998) encontraram que a ocupação de 24% das camas se devia a fracturas osteoporóticas) e sendo a sua prevenção tão cara, torna-se extremamente importante realizar-se estudos epidemiológicos de modo a identificar as populações de alto risco (Soveid M, Serat AR et al., 2005) e a frequência e distribuição da doença (Orimo H, Hoshimoto T et al., 2000). Também segundo Kanis J, Delmas P et al. (1997) estas fracturas justificam a ocupação de mais de 20% ortopédicas no Reino Unido, na Escandinávia e em vários outros países.

Nas mulheres, após os 45 anos de idade, as fracturas do colo do fémur são responsáveis pela maior percentagem de ocupação de camas hospitalares, quando comparada com outras doenças como diabetes e cancro da mama (ver **Figura 2.16**).

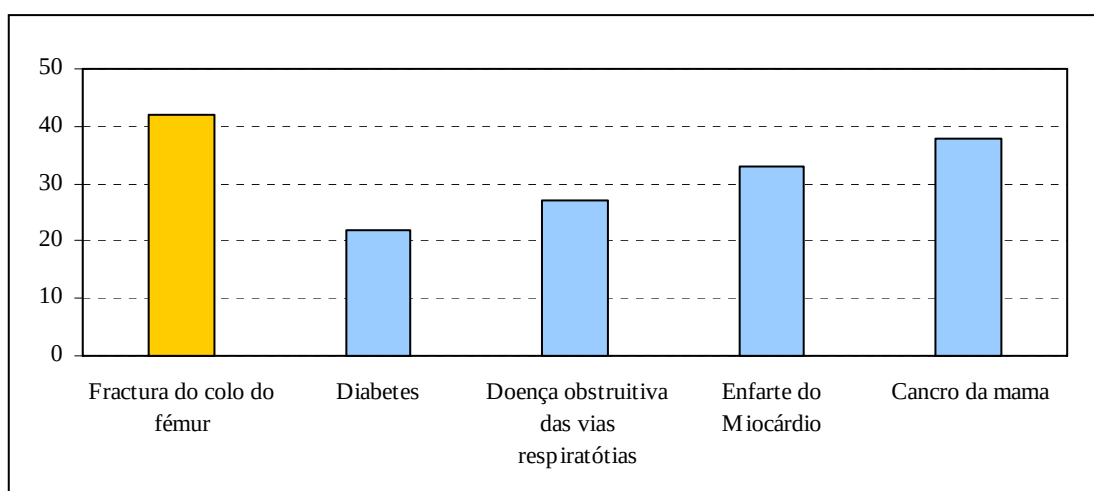


Figura 2.16 Ocupação de camas hospitalares em mulheres, com idade superior a 45 anos, de acordo com o diagnóstico na região de Trend em Inglaterra.

Fonte: Kanis J, Delmas P et al., (1997).

Ao se estudar espacialmente a osteoporose é possível mapearem-se indicadores como as taxas de incidência das fracturas do colo do fémur, que variam acentuadamente a nível mundial, de região para região (Cummings SR, Melton L, 2002 e Sawka A, Thabane L et al., 2005), programar actividades de intervenção, prevenção e controle (Pina MF, 2003^a), capazes de reduzir as taxas de incidência das fracturas e melhorar a qualidade de vida das pessoas, bem como identificar características semelhantes (como condições socioeconómicas ou ambientais) que possam ter influência na taxa de incidência da doença e implementar planos de educação nessas regiões.

2.4 Sistemas de Informação Geográfica na Saúde

Analisar a saúde de grupos populacionais considerando a sua localização espacial e temporal, a sua interacção com o ambiente, com a distribuição dos recursos de saúde e com outros grupos traz uma importante compreensão do processo de saúde e doença nas populações (Carvalho MS, Nobre FF, 2001).

De acordo com Carvalho MS e Pina MF et al. (2000) uma das maneiras de se conhecer mais detalhadamente as condições de saúde da população é através de mapas que permitem observar a distribuição espacial de situações de risco e de problemas de saúde. A abordagem espacial permite a integração de dados demográficos, socioeconómicos e ambientais promovendo o interrelacionamento das informações de diversas origens.

A associação entre a saúde e a geografia é antiga. Em 1768, o médico escocês James Lind publicou um livro chamado “*An Essay on Diseases Incidental to Ruopians and Hot Climates*” no qual procurou explicações para distribuição de doenças, chegando inclusive a determinar áreas geográficas específicas. Desde então diversos trabalhos foram escritos na geografia médica, descrevendo as variações geográficas na distribuição das doenças (Carvalho MS, Pina MF et al., 2000). Um dos mais famosos utilizadores da geografia médica foi John Snow (sec. XIX), um médico inglês que ao investigar a relação entre a propagação da cólera em Londres e o escoamento superficial das águas contaminadas, abriu caminho para a relação entre as aplicações geográficas na investigação dos problemas da saúde (Pina MF, 2003^a). Já no final do século XX, com o crescente interesse dos epidemiologistas pela análise espacial dos dados de saúde e o desenvolvimento das geotecnologias de informação surge um campo da epidemiologia, denominada epidemiologia espacial, cujo objectivo é, através do mapeamento, descrever, quantificar e explicar as variações geográficas das doenças, determinar a associação entre a incidência de eventos observados e potenciais factores de risco e identificar aglomerados espaciais de doenças (Elliot P, Wakefield et al., 2000).

Aplicadas a questões de saúde pública as geotecnologias e em especial os Sistemas de Informação Geográfica - SIG, permitem o mapeamento de doenças, a avaliação de riscos, o planeamento de acções de saúde e a avaliação de redes de saúde (Resendes A, Barcellos C et al., 2006).

Segundo Vine et al (1997), citado em Rocha L (2000), um Sistema de Informação Geográfica é uma ferramenta computacional poderosa que possibilita a análise de grandes quantidades de dados dentro de um contexto geográfico. Segundo outros autores, um SIG é um conjunto organizado de equipamentos, programas e dados geográficos e individuais, planeados para recolher, armazenar, actualizar, manipular, analisar e apresentar diferentes formas de informação geograficamente referenciada, operando de maneira integrada. Uma outra definição descreve um SIG como um sistema computacional que armazena e relaciona (mapas e gráficos) para permitir a apresentação de um leque de processos, de informação, produção de mapas, análise e modelação de dados. Uma definição mais operacional descreve o SIG como um sistema de apoio para a tomada de decisões, envolvendo a integração de dados referenciados espacialmente para a solução de um problema (Rocha L, 2000).

A utilização dos SIGs possibilita realizar análises espaciais complexas, pois permite a integração de dados de diversas fontes, manipulação de grande volume de dados e recuperação rápida de informações armazenadas (Resendes A, Barcellos C et al., 2006). Os SIG assumem uma importância cada vez maior em diversas áreas, da cartografia a estudos de impacto ambiental ou da prospecção de recursos ao marketing. Estes estudos têm experimentado avanços importantes motivados pelas novas tecnologias de informação (Prado A, Baranauskas M et al., 2000), sendo que a rápida proliferação dos computadores e dos sistemas de informação mudaram também a forma de organização dos cuidados de saúde e de compreensão das necessidades relacionadas com a saúde (Ricketts T, 2003).

Um SIG possui quatro funções básicas:

- Recolha de dados: envolve a aquisição, importação, validação e edição dos dados.
- Gestão de bases de dados: envolve o armazenamento dos dados de forma estruturada, de modo a possibilitar e facilitar a realização de análises.
- Consulta e análise: uma função que pode ser considerada como a principal de um SIG, pois possibilita operações de extracção e geração de novas informações sobre o espaço geográfico, a partir de critérios especificados pelo próprio utilizador.
- Visualização e apresentação cartográfica: relacionada com a capacidade de utilizar as diversas camadas de dados e exibir este resultado através de mapas e gráficos com boa qualidade gráfica.

Um Sistema de Informação, para ser considerado geográfico, necessita possuir uma referência da localização de onde está a informação, ou seja, a informação precisa ter uma posição geográfica. Essa posição geográfica pode ser identificada através de um par de coordenadas, ou um geocódigo.

A introdução das geotecnologias de informação na área de saúde, apesar de não serem uma solução completa para a compreensão da distribuição das doenças e de problemas de saúde pública, são um importante caminho para o esclarecimento de como os indivíduos interagem com o ambiente (Ricketts T, 2003).

Segundo a OMS, num sentido mais abrangente, a saúde é o resultante das condições de alimentação, habitação, educação, renda, meio ambiente, transporte, trabalho, emprego, lazer, liberdade, acesso e posse de terra e acesso aos serviços de saúde (Barcelos C, Ramalho W, 2002). Na área da Saúde os SIGs permitem avaliar riscos, programar actividades de prevenção e controle de doenças, estudar a acessibilidade da população aos serviços de saúde, identificar áreas próximas fontes de contaminação e determinar factores ambientais adversos em locais onde há concentração de consequências prejudiciais à saúde (Carvalho MS, Pina MF et al., 2003).

Segundo Carvalho MS, Nobre FF (2001) as principais áreas de aplicação das geotecnologias de informação dividem-se basicamente em três grupos: mapeamento de doenças - onde se procura mapear a distribuição espacial subjacente da doença; estudos ecológicos - cujo objectivo é encontrar a associação entre a incidência observada e potenciais factores de risco, medidos por grupos espacialmente delimitados; e a análise de ocorrência de *cluster* de doenças – de forma a visar diferenças significativas no padrão de distribuição espacial de doenças, e/ou relacionar este padrão com potenciais fontes ambientais de risco.

Tradicionalmente, têm sido utilizados mapas para se descrever a distribuição espacial de fenómenos relacionados com a saúde. Mais recentemente, a utilização de mapas tem contribuído para orientar a formulação de hipóteses sobre a etiologia das doenças, considerando as variações espaciais de factores ambientais e sociais. Pela sua capacidade de integrar num único ambiente computacional dados ambientais, socioeconómicos e de saúde, os SIGs oferecem novas e importantes oportunidades para a descrição e análise das relações entre atributos dos indivíduos e a distribuição de eventos de saúde no espaço geográfico (Rocha L, 2000), permitindo o seu

interrelacionamento para o desenvolvimento de diversas análises geográficas (Pina MF, 2003^a).

Segundo a Organização Pan-Americana da Saúde (1996), entre as mais importantes aplicações dos SIGs na saúde, encontram-se as seguintes: a descrição espacial de um evento de saúde; a identificação de riscos ambientais e ocupacionais; a análise da situação da saúde numa área geográfica; a análise de padrões ou diferenças na situação de saúde em diversos níveis de agregação; a identificação de grupos de alto risco à saúde e áreas críticas; a vigilância e monitoramento da saúde pública; a geração de hipóteses de pesquisa operacional; o planeamento e a programação de actividades em saúde; e a avaliação de intervenções de saúde (Rocha L, 2000).

De acordo com estas aplicações estão autores como Resendes A, Barcellos C et al. (2006) e Pina MF (2003^a). Estes autores referem nos seus estudos que os SIGs permitem responder a algumas perguntas fundamentais da epidemiologia como:

- A distribuição dos casos de uma doença forma um padrão no espaço?
- Existe alguma associação da doença com possíveis fontes de contaminação ou meios de disseminação?
- Existe a evidência de uma determinada doença ser transmitida de indivíduo a indivíduo, ou através de uma fonte comum?

A adaptação desta técnica às áreas relacionadas com a saúde pública é de grande valor não apenas por oferecer a possibilidade de desenvolver diversas tarefas fundamentais com maior rapidez e economia, mas também por oferecer aos pesquisadores novas alternativas metodológicas, mais confiáveis e cientificamente validadas para o tratamento de informações espaciais (Pina MF, 1998). Por outro lado, a integração dos métodos e técnicas de epidemiologia e os SIGs permite visualizar padrões epidemiológicos de eventos e processos de saúde, bem como reconhecer a importância dos factores que os determinam facilitando a tomada de decisões sobre possíveis acções de saúde pública (Rocha L, 2000).

Alguns estudos integrando os SIGs e as técnicas de análise estatística espacial já foram realizados na área da saúde em Portugal, dos quais se destaca o estudo de Aves S (2004) na análise geográfica das taxas de incidência de fracturas osteoporóticas do colo do fémur.

2.5 Análise de Componentes Principais

A análise multivariada é um ramo da estatística que tem registado nas duas últimas décadas um desenvolvimento e um esforço de investigação notáveis, provocados essencialmente pela explosão de meios computacionais que tornaram possível a análise de conjuntos de dados cada vez mais volumosos e complexos. Consequentemente, assiste-se a uma grande procura deste tipo de métodos por parte dos utilizadores, nas mais variadas áreas do conhecimento.

A Análise de Componentes Principais (ACP) é uma técnica estatística de análise multivariada cujo objectivo principal é a obtenção de um pequeno número de combinações lineares (componentes principais), a partir de um conjunto de variáveis, que retenham o máximo possível da informação contida nas variáveis originais. Ou seja, é uma técnica que transforma linearmente um conjunto original de variáveis num conjunto substancialmente menor de variáveis, não correlacionadas, que contêm a maior parte da informação do conjunto original (Marques J, Marques M, 2005). Um conjunto menor de variáveis não correlacionadas é muito mais fácil de compreender e de utilizar em futuras análises do que um conjunto grande de variáveis correlacionadas (Dunteman G, 1989). Outro objectivo da ACP é analisar quais as variáveis ou quais conjuntos de variáveis que explicam a maior parte da variabilidade total, revelando que tipo de relacionamento existe entre eles.

Scremin M (2003) refere no seu trabalho que o autor Crivisqui E (1999) aponta como objectivos da Análise de Componentes Principais:

- Avaliar a semelhança entre os casos por meio dos atributos considerados, explorando questões do tipo: “Existem grupos de casos semelhantes?” e “Observa-se uma tipologia de casos?”. Isso pode ser avaliado por uma métrica de distância (euclidiana, *Manhatan*, etc);
- Avaliar a relação existente entre as características consideradas, explorando questões do tipo: “Existem grupos de variáveis correlacionadas entre si?” e “Observa-se uma tipologia de variáveis?”. Esta avaliação pode ser realizada pelo exame das correlações entre as variáveis.

O objectivo da ACP é similar ao da análise factorial no sentido em que ambas as técnicas tentam explicar parte da variação de um conjunto de variáveis a partir de um

número menor de variáveis subjacentes (Dunteman G, 1989). Sumariamente, a principal diferença entre as duas técnicas é a de que a ACP parte da ausência de um modelo estatístico subjacente na divisão das variáveis observadas e foca a explicação da variância total das variáveis observadas baseando-se nas propriedades da variância máxima das componentes principais. A análise factorial, por outro lado, parte de um modelo estatístico prévio que divide a variância total (Fleck M, Bourdel M, 1998). Esta técnica é também semelhante a outros métodos da análise multivariada como a análise discriminante e análise da correlação canónica, por estas também envolverem combinações lineares de variáveis correlacionadas (Dunteman G, 1989).

Nos modelos de regressão múltipla é muito comum a ocorrência de variáveis independentes altamente correlacionadas, resultando em coeficientes de regressão estimados com baixa precisão. Por esta razão, é frequente procurar formas de reorganizar ou sumariar os dados para que, com a menor perda de informação possível, a dimensão do problema seja reduzida (Press J, 1972). Entre as várias alternativas que existem para reduzir a “dimensionalidade” do modelo, uma delas consiste na utilização de componentes principais.

De salientar que a redução de “dimensionalidade” introduzida pela ACP não significa a redução no número de variáveis originais com que se trabalha. Mesmo que se venha a reter apenas q componentes principais (uma vez que não trás nenhuma vantagem reter todas as componentes principais, pois ter-se-á tantas componentes quantas variáveis e, isso, não simplifica o problema) precisamos da totalidade das p variáveis originais para definir essas q componentes principais. No fundo a redução de “dimensionalidade” é uma falsa redução.

Como nos modelos de regressão, cujo propósito é a explicação da variável dependente, devem-se reter aquelas componentes principais que têm altas correlações com a variável dependente (Marques J, Marques M, 2005). Isto é, quando uma variável (comparada com outras) não varia muito nas repetidas “experiências” pode ser considerada como sendo, aproximadamente, estável. Libertando as variáveis com maiores correlações entre si e centrando a atenção nas restantes variáveis, o fenómeno de interesse pode ser estudado mais convenientemente num subespaço de dimensão menor do que o definido pelo conjunto completo das variáveis. Apesar de alguma informação sobre a relação entre as variáveis ser perdida por este procedimento, na

maior parte dos estudos esta perda não é significativa tornando-se vantajoso utilizar esta análise (Press J, 1972). Isto pode ser conseguido transformando os dados originais num novo conjunto de variáveis, as componentes principais, que são constituídas por combinações lineares das variáveis originais, que não estão ou estão pouco correlacionadas, e que estão ordenadas de modo que a primeira componente explica a maior proporção da variância total (fazendo com que as correlações do primeiro factor com as variáveis sejam maiores), a segunda retem o máximo de variabilidade restante (fazendo com que as correlações deste factor com as variáveis sejam as segundas maiores), e assim sucessivamente (Jolliffe I, 1986).

A análise factorial teve início com os trabalhos de Karl Pearson em 1901 e Charles Spearman em 1904, durante um estudo para a criação de um índice de inteligência. Em 1947, Thurstone L retomou a ideia dos precursores e desenvolveu a análise factorial. Desde então, ela tem sido amplamente empregada em diferentes áreas da ciência (Reis E, 2001).

Pearson desenvolveu as fundações matemáticas da análise de componentes principais (Burstyn I, 2004), considerando casos de duas e três dimensões (Krzanowski W, Marriot F, 1994). No entanto, o cálculo dos valores e vectores próprios de grandes matrizes, para dimensões maiores, desencorajou o desenvolvimento desta ideia. Segundo Krzanowski W e Marriot F (1994), foi Hotelling H (1933) que inicialmente desenvolveu o conceito das componentes principais no contexto de análise de factores e que primeiro descreveu um método iterativo para encontrar as componentes principais num caso geral.

A ACP é por vezes precedente a procedimentos da análise factorial, podendo também ser utilizada para seleccionar um subconjunto de variáveis de um conjunto maior, ou ser útil na exposição de dados multivariados em gráficos, de modo que, por exemplo, possam ser detectadas observações atípicas (Dunteman G, 1989).

Geometricamente, as combinações lineares representam a selecção de um novo sistema ortogonal de coordenadas, obtido por rotação do sistema ortogonal original com $X_1, X_2, \dots, X_p$ como eixos (ver **Figura 2.17**). O objectivo da ACP pode ser expresso como a procura do subespaço de dimensão m no qual a representação dos dados seja o mais “fiel” possível (Cadima J, 2002). Os novos eixos, $Y_1, Y_2, \dots, Y_p$, vão representar as direcções com variabilidade máxima, permitindo uma interpretação mais simples da

estrutura da matriz de covariância (Marques J, Marques M, 2005). Ao mesmo tempo, as variâncias das projecções de X no novo sistema de coordenadas são maximizadas de forma a que o primeiro eixo corresponda à variância máxima, o segundo corresponda à variância máxima, ortogonal ao primeiro eixo, e assim sucessivamente (Hyvarinen A, Karhunen J, Oja J, 1994).

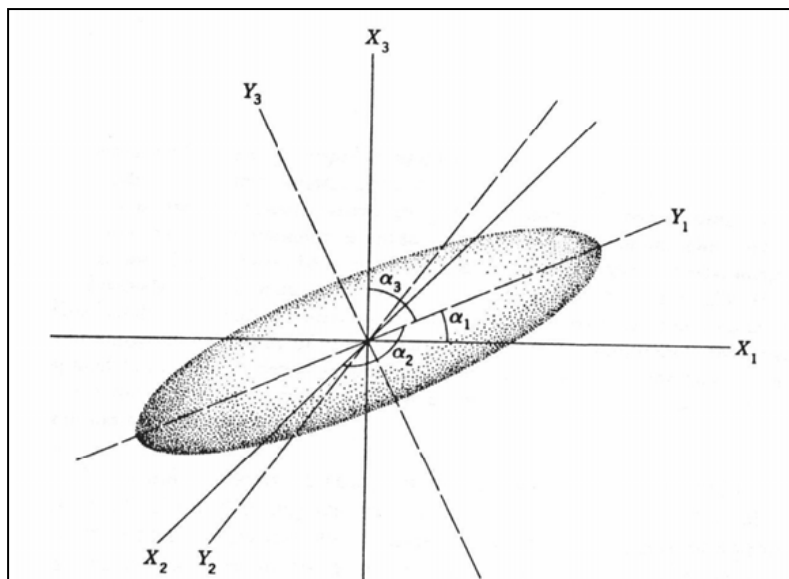


Figura 2.17 Eixos principais com três variáveis.

Fonte: Scremin M (2003).

2.5.1 Matriz das Correlações

Em geral, o primeiro passo a dar na ACP, consiste no exame das relações entre as variáveis, por exemplo, utilizando o coeficiente de correlação linear como medida de associação entre cada par de variáveis. A matriz das correlações poderá permitir identificar subconjuntos de variáveis que estão muito correlacionadas entre si mas pouco associadas a variáveis de outros subconjuntos (Reis E, 2001).

Esta matriz é uma matriz simétrica e mede a associação linear entre as variáveis através do coeficiente de correlação linear r de Pearson. Quanto mais perto de 1 ou de -1 se situar este coeficiente, maior será a associação linear entre as variáveis (Pestana M, Gageiro J, 2005).

2.5.2 Testes para Validação da Análise

O Kaiser-Meyer-Olkin - KMO e o teste de Bartlett são dois procedimentos estatísticos que permitem aferir a qualidade das correlações entre as variáveis de forma a prosseguir com a análise (Pestana M, Gageiro J, 2005).

Um indicador da força da relação entre as variáveis são os coeficientes de correlação parciais, que são estimativas das correlações entre os factores únicos e que se aproximam de zero quando se verificam os pressupostos da análise (Pestana M, Gageiro J, 2005).

O KMO é uma estatística que varia entre zero e um, e compara as correlações de ordem zero com as correlações parciais observadas entre as variáveis. Valores próximos de um indicam coeficientes de correlação parciais pequenos, enquanto valores próximos de zero indicam que a análise factorial pode não ser uma boa ideia, porque existe uma correlação fraca entre as variáveis (Pestana M, Gageiro J, 2005).

Kaiser adjectiva os valores do KMO como se apresentam na seguinte tabela.

Tabela 2.4 Classificação da análise factorial, segundo Kaiser, segundo os valores do KMO.

KMO	Análise factorial
0,9 – 1	Muito Boa
0,8 – 0,9	Boa
0,7 – 0,8	Média
0,6 – 0,7	Razoável
0,5 – 0,6	Má
< 0,5	Inaceitável

Fonte: Baseado em Pestana M, Gageiro J, (2005).

O teste de esfericidade de Bartlett testa a hipótese da matriz das correlações ser a matriz identidade, cujo determinante é 1. Este teste requer que os dados provenham de uma população normal multivariada. No entanto este teste é muito influenciado pelo tamanho da amostra e leva a rejeitar a hipótese nula em grandes amostras, pelo que se torna preferível usar o KMO (Pestana M, Gageiro J, 2005).

Outro tipo de teste para este efeito é o da matriz anti-imagem. Esta matriz forma-se com os simétricos dos coeficientes das correlações parciais. Se apresentar valores pouco elevados em número significativo conclui-se ser válida a aplicação da Análise das Componentes Principais (Reis E, 2001).

2.5.3 Extracção das Componentes Principais

1. Conceitos matemáticos:

Considerando-se uma matriz quadrada $A_{p \times p}$, um vector $x \in C^p$ (com $x \neq 0$) tal que:

$$A x = \lambda x \quad (2.1)$$

para algum escalar $\lambda \in C^p$ diz-se um vector próprio da matriz A . O escalar λ diz-se o valor próprio associado ao vector próprio x .

Se $A_{p \times p}$ for uma matriz simétrica então os vectores próprios associados a valores próprios diferentes são sempre ortogonais. Sendo também sempre possível determinar um conjunto ortonormado de k vectores próprios, mesmo quando há valores próprios repetidos.

Teorema 2.1

Seja $A_{p \times p}$ uma matriz simétrica. Então:

1. A característica de A , $car(A) = k$, é dada pelo número de valores próprios não nulos de A .
2. Os vectores próprios de A associados aos valores próprios não nulos formam uma base ortonormada do subespaço imagem $\Re(A)$.

O resultado seguinte relaciona os valores e vectores próprios de uma matriz simétrica A com os valores da função $\frac{x^t A x}{x^t x}$ ou, equivalentemente, com os vectores da forma quadrática $x^t A x$ para vectores x de norma 1.

Teorema 2.2 (Teorema de Rayleigh-Ritz):

Seja $A_{p \times p}$ uma matriz simétrica, cujos valores próprios são indexados por ordem decrescente:

$$\lambda_{\max} = \lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_{p-1} \geq \lambda_p = \lambda_{\min}. \quad (2.2)$$

1. O maior valor próprio de A é: $\lambda_{\max} = \max_{x \neq 0} \frac{x^t A x}{x^t x}$.
2. O menor valor próprio de A é: $\lambda_{\min} = \min_{x \neq 0} \frac{x^t A x}{x^t x}$.

3. Os restantes λ_j são os sucessivos máximos (ou sucessivos mínimos) do quociente, sujeitos à exigência de que os vectores x considerados sejam ortogonais aos vectores próprios já determinados.

Teorema 2.3 (Decomposição em valores singulares – DVS)

Seja $X \in M_{n \times p}$ uma matriz de característica r . Então pode-se escrever:

$$X = W\Delta V^t \Leftrightarrow X = \sum_{i=1}^r \delta_i w_i v_i^t, \quad (2.3)$$

onde, $\Delta_{r \times r}$ é uma matriz diagonal, $\left. \begin{matrix} V_{p \times r} \\ W_{n \times r} \end{matrix} \right\}$ são matrizes com colunas ortornormadas,

δ_i são os elementos diagonais de Δ (os valores singulares de X), w_i são as colunas de W (os vectores singulares esquerdos de X) e v_i são as colunas de V (os vectores singulares direitos de X).

Este resultado é muito importante, permitindo factorizar qualquer matriz. A DVS desempenha um papel muito importante na ACP e noutras técnicas de estatística multivariada.

Observação:

A matriz V é uma matriz cujas colunas são conjuntos ortornormados de vectores próprios de $X^t X$, associados a valores próprios não nulos. W é a matriz análoga correspondente a $X^t X$. E, Δ é a matriz das raízes quadradas dos valores próprios não nulos $X^t X$.

Teorema 2.4

Seja $X_{n \times p}$ uma matriz de característica r , com a DVS dada por (2.3). Então as colunas $\{w_i\}_{i=1}^r$ da matriz W formam uma base ortornormada do subespaço imagem da aplicação definida por X , $\Re(X) \subset \Re^n$, e as colunas $\{v_i\}_{i=1}^r$ da matriz V formam uma base ortornormada do subespaço imagem da aplicação definida por X^t , $\Re(X^t) \subset \Re^p$.

Teorema 2.5

Seja $X_{n \times p}$ uma matriz de característica k . A matriz $Y_{n \times p}$ de característica $m < k$ que melhor se aproxima de X , no sentido de minimizar a soma de quadrados das diferenças entre os correspondentes elementos de X e Y obtém-se da seguinte forma:

- Seja $X = W\Delta V^t$ a DVS de X .
- Sejam W_m , V_m , as matrizes constituídas pelas m primeiras colunas de W e V , respectivamente. Seja Δ_m a matriz diagonal tipo $m \times m$ resultante de reter apenas as m primeiras linhas e colunas de Δ .
- Então $Y = W_m \Delta_m V_m^t$.

Em $\mathfrak{R}^n$:

1. A média de n observações é o coeficiente da projecção ortogonal do vector de observações x sobre o subespaço $\mathfrak{R}^n(1_n)$, isto é, sobre a bissectriz do primeiro octante de $\mathfrak{R}^n$.
2. A variável centrada em torno da sua média, isto é, o vector com componentes $x_i - \bar{x}$ é a projecção ortogonal de x no complemento ortogonal do subespaço gerado pelo vector dos uns.
3. O desvio padrão das n observações é proporcional à distância do vector x ao subespaço gerado pela coluna de uns, $\mathfrak{R}^n(1_n)$. A distância de um vector a um subespaço é dada pela menor distância entre o referido vector e qualquer vector do subespaço e esta menor distância atinge-se quando se considera a distância do vector à sua projecção ortogonal sobre o subespaço.

De seguida são apresentados dois métodos utilizados para a obtenção das componentes principais baseadas nos apontamentos de Cadima J (2002).

2. Método da máxima verosimilhança:

O ponto de partida para a ACP é uma matriz de dados X , centrada. Nesta abordagem pretende-se determinar o vector $v \in \mathfrak{R}^p$ que maximize a variância da combinação linear Xv das variáveis observadas, isto é, que maximize a forma quadrática $v^t \Sigma v$ - onde $\Sigma = \frac{1}{n} X^t X$ (matriz de variâncias/covariâncias).

Uma vez que a variância depende tanto da norma como da orientação do peso do vector v e cresce ilimitadamente à medida que a norma cresce, impõe-se que a norma de v é constante, e na prática igual a 1. Assim $v^t v = 1$.

Se v não for de norma 1, o vector normalizado $\frac{v}{\|v\|}$ será e a forma quadrática por ele definida tem a forma do quociente de Rayleigh-Ritz (2.2). Sendo assim, maximizar a forma quadrática $v^t \Sigma v$ equivale a maximizar o quociente $\frac{v^t \Sigma v}{v^t v}$. O Teorema Rayleigh-Ritz (**Teorema 2.2**) garante que essa maximização ocorre para o vector próprio associado ao maior valor próprio de Σ , ou seja para v_1 .

Logo, a combinação linear das variáveis originais procurada é a combinação Xv_1 , ou seja, o vector que já foi definido como a primeira componente principal dos dados. O valor próprio λ_1 é assim a variância da primeira componente principal.

Fixada a primeira componente principal, passa-se a procurar uma nova combinação linear Xv de variância máxima, mas agora com a restrição adicional de que a nova combinação linear seja não-correlacionada com a anterior. Portanto, que a covariância entre Xv e Xv_1 seja zero.

Isto equivale à exigência de ortogonalidade entre os vectores v e v_1 , por exigir que o produto interno entre os dois vectores seja nulo, isto é, que $v^t X^t X v_1 = n \cdot v^t \Sigma v_1 = n \cdot \lambda_1 v^t v_1$.

Assim pretende-se maximizar o quociente $\frac{v^t \Sigma v}{v^t v}$, sujeito à restrição $v^t v_1 = 0$. Este problema é resolvido por $v = v_2$, o segundo vector próprio de Σ , associado ao valor próprio λ_2 .

Assim, esta segunda combinação linear é a segunda componente principal, com variância λ_2 .

As restantes componentes principais surgem de forma análoga, como solução do problema de determinar sucessivas combinações lineares de variância máxima, não correlacionadas entre si.

A j -ésima componente principal é dada por Xv_j , onde v_j é o vector próprio de Σ , associado ao j -ésimo maior valor próprio λ_j .

3. Método dos mínimos quadrados:

Com este método escolhe-se o subespaço m -dimensional que minimize a soma dos quadrados das distâncias entre as observações e as respectivas projecções ortogonais sobre o subespaço (ver **Figura 2.18**) (Cadima J, 2002).

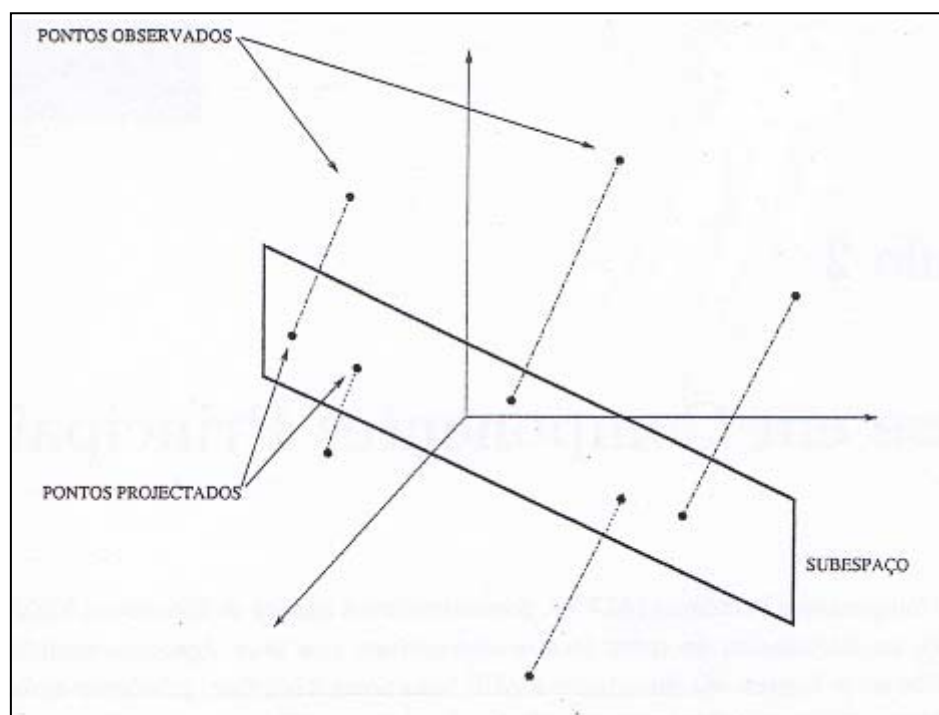


Figura 2.18 A projecção ortogonal de pontos sobre um subespaço.

Fonte: Cadima J, (2002).

Considere-se n pontos num subespaço de dimensão q de $\mathfrak{R}^p$, e coloque-se as respectivas coordenadas nas n linhas de uma matriz $\tilde{X}_{n \times p}$. O facto dos pontos estarem num subespaço de dimensão q vai-se reflectir no facto da matriz $\tilde{X}$ ser de característica q . Por outro lado a soma de quadrados das diferenças entre as coordenadas de cada linha de X e da correspondente linha de $\tilde{X}$ é dada por $\|X - \tilde{X}\|^2$.

Então, considerando o problema em termos de subespaço $\mathfrak{R}^p$, pode-se dizer que através da ACP procura-se a matriz $\tilde{X}_{n \times p}$, de característica q , que minimiza $\|X - \tilde{X}\|$.

Assim, a solução do problema já foi encontrada no **Teorema 2.5**, obtida a partir da DVS, retendo apenas os q primeiros vectores singulares (esquerdos e direitos) e valores singulares.

Ou seja, a matriz $\tilde{X}$ será dada por $\tilde{X} = W_q \Delta_q V_q^t$ em que as colunas W_q e V_q são os q vectores singulares esquerdos e direitos da matriz X associados aos q maiores valores singulares, que constituem os elementos diagonais da matriz diagonal Δ_q .

Esta solução resulta do facto das linhas de $\tilde{X}$ resultarem de se projectar ortogonalmente as colunas de X sobre o subespaço (de dimensão q) de $\mathfrak{R}^p$ gerado pelos q primeiros vectores próprios – representados por $\{v_j\}_{j=1}^q$ - de $\Sigma = \frac{1}{n} X^t X$ (matriz de variâncias/covariâncias); e das colunas de X resultarem de se projectar ortogonalmente as colunas de X sobre o subespaço (de dimensão q) de $\mathfrak{R}^n$ gerado pelos q primeiras colunas de $X V_q$, isto é, pelas q primeiras colunas de $W_q \Delta_q$.

Assim, as componentes principais das variáveis $\{x_j\}_{j=1}^p$ são as variáveis em $\mathfrak{R}^n$, $\{X v_j\}_{j=1}^p$, obtidas efectuando combinações lineares das variáveis originais, utilizando como coeficientes nessas combinações lineares os elementos dos vectores próprios v_j , que costumam-se ordenar pela ordem decrescente dos valores próprios de Σ associados aos vectores próprios v_j .

2.5.4 Selecção do Número de Componentes Principais

Uma das maiores dificuldades no uso da ACP está na selecção do número de componentes, por esta ser uma decisão muito subjectiva. A selecção do número de componentes principais ou factores depende além da subjectividade, também dos objectivos da pesquisa. Por isso, torna-se difícil seleccionar adequadamente estas componentes para que representem o máximo de informação contida nos dados originais, atendendo às necessidades do problema.

A questão de quantas componentes reter está ligada à estabilidade dos resultados obtidos e à finalidade do estudo, como por exemplo, se a finalidade for uma simples representação gráfica para examinar visualmente a estrutura dos dados, o número de componentes principais mais adequado é duas; se o objectivo for reter componentes principais para utilizar como suporte numa classificação, deve-se reter o número suficiente de componentes para que estas representem ao máximo os dados originais,

eliminando as que reflectam flutuações aleatórias dos dados; ou se as componentes principais forem utilizadas como variáveis explicativas em uma análise discriminante ou regressão, procura-se reter um número pequeno de componentes, não necessariamente as primeiras. Mas, de um modo geral, o que se quer é determinar o menor número de componentes principais que melhor representem as relações entre o conjunto de variáveis. Isto envolve ponderar duas necessidades: ter o menor número de componentes possível e explicar o máximo da variância nos dados originais - devendo-se realizar uma análise exploratória até ser encontrada uma solução satisfatória.

Na procura da solução para a questão do número de componentes a seleccionar, muitos métodos têm sido desenvolvidos. Existem, no entanto, quatro métodos mais frequentemente utilizados na decisão do número de componentes principais a reter:

1. Método de Kaiser:

Um dos principais critérios utilizados é o método de Kaiser ou método de Kaiser-Guttman (Cramer D, 2003). Neste método, Kaiser sugere que a melhor resposta para a questão do número de componentes, quando se utiliza a matriz de correlação, será dada pelo número de componentes cujos valores próprios são maiores ou iguais a 1.

Este é um dos mais utilizados na prática, tornando-se importante salientar que este método encontra-se implementado em alguns pacotes computacionais como o STATISTICA™ e o SPSS™ (Statistical Package for Social Sciences).

Geralmente, utilizando a regra do método de Kaiser, retêm-se apenas as componentes mais importantes, no entanto o seu uso sem atenção pode levar a reter componentes que não têm praticamente nenhuma significância.

Apesar de ser um dos métodos mais utilizados por ser de fácil aplicação, é muito criticado. Segundo Stevens J (1996) os estudos realizados por Cattell e Jaspers (1967), Browne (1968) e Linn (1968) avaliaram a acuracidade do critério utilizado neste método. Sabendo que a comunalidade (*communality*) de uma variável é o total da variância dessa variável, explicada pelo conjunto de componentes, que varia entre 0 e 1 e que nestes estudos o número de variáveis encontrava-se entre 10 e 40, tornou-se possível perceber que este critério é mais fiável quando o número de variáveis é pequeno (entre 10 e 15) ou moderado (entre 20 e 30) e quando as comunalidades são altas (aproximadamente 0,70), tornando-se menos fiável com um grande número de variáveis (40) e baixas comunalidades (0,40) (Stevens J, 1996).

2. Método do Diagrama de Valores Próprios:

Este método gráfico, também chamado *scree test*, foi proposto por Cattell R em 1966 (Stevens J, 1996). Cattell sugeriu que o método de Kaiser retém demasiadas componentes quando existem muitas variáveis e poucas componentes quando existem poucas variáveis, acabando por propor um método alternativo (Cramer D, 2003).

Neste método representam-se graficamente a amplitude dos valores próprios (conceito explicado na fórmula 2.1) em função do número dos valores próprios, ordenados crescentemente e seleccionam-se as componentes anteriores ao ponto de ruptura da curva, que ocorre quando há uma queda brusca na amplitude dos valores próprios. Na **Figura 2.19** exemplifica-se uma situação em que 2 componentes principais seriam seleccionadas se usado este método.

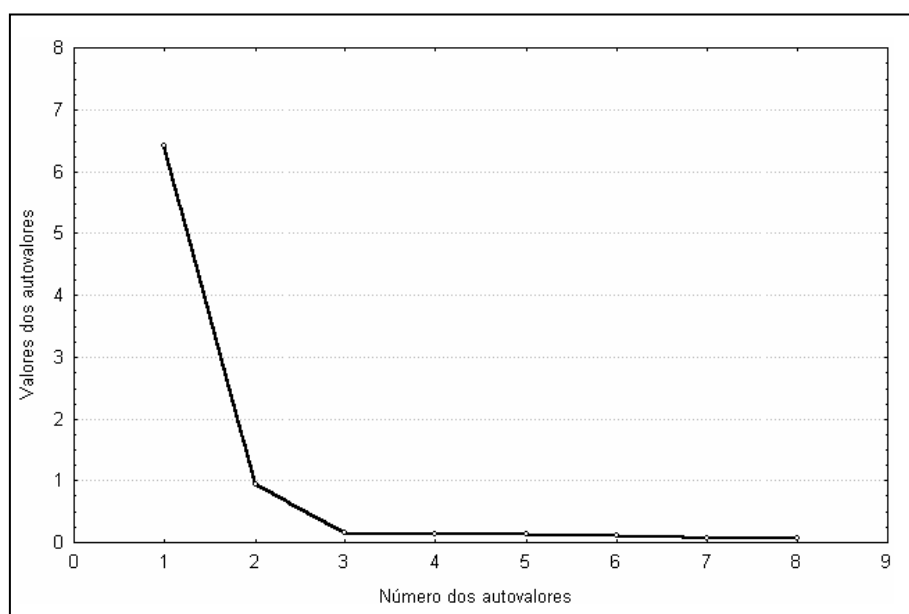


Figura 2.19 Selecção do número de componentes principais pelo método *scree test*.

Fonte: Scremin M, (2003).

Scree é um termo geológico para os sedimentos que se acumulam na base de uma declive acentuado e que escondem a verdadeira base do declive. Daí resulta o nome do teste, visto que o número de componentes a reter é dado pelo número do factor que aparenta representar a linha do fim do declive (Cramer D, 2003).

Este método, geralmente, retém componentes responsáveis por totais de variância grandes ou razoavelmente grandes e distintos. No entanto, o seu uso incorrecto pode levar à não retenção de componentes que, embora retenham um valor menor do total de

variância, podem ser significativos (Stevens J, 1996). A utilização deste teste quando o tamanho da amostra (N) é superior a 200 não será tão proveitosa, pois as maiorias das acumulações são razoavelmente grandes (Stevens J, 1996).

Num estudo sobre este problema, Hakstian A, Rogers W et al. (1982), citados em Stevens J (1996), mostram que quando o N é maior que 250 e quando a média das comunalidades é maior ou igual a 0,60, tanto o método de Kaiser como este método irão dar uma estimativa exacta para o número de verdadeiros componentes. Acrescentando ainda que esta estimativa se torna ainda mais precisa se o quociente entre o número de variáveis e o número de componentes é inferior a 0,30 (tornando-se o método de Kaiser menos preciso e este método muito menos preciso se o quociente tomar um valor superior ao considerado) (Stevens J, 1996).

3. Método do teste de significância estatística:

Sengudo Stevens J (1996), este método foi desenvolvido por Lawley D (1940), para determinar o número de componentes a reter. É um método pouco utilizado porque é excessivamente dependente do tamanho da amostra, tendo como grande desvantagem o facto de uma amostra grande poder levar à retenção de demasiadas componentes (Stevens J, 1996).

4. Método da variância pré-determinada:

Por último, este método baseia-se em reter tantas componentes quantas as necessárias para se atingir um determinado total de variância desejado, definido *à priori*. Geralmente, alguns autores consideram apenas 70% da total da variância, enquanto que outros utilizam valores entre os 80 e os 85%. Esta restrição pode ter como desvantagem a retenção de componentes que se correlacionam fortemente com uma única variável (Stevens J, 1996).

2.5.5 Métodos de Rotação

Apesar das componentes principais serem boas para sumariar a variância contida num grande conjunto de variáveis através de um menor número de componentes, nem sempre estas componentes são de fácil interpretação. Os métodos de rotação são habitualmente adoptados como uma tentativa de superar os problemas de interpretação

física das componentes principais (Stevens J, 1996). Existem duas grandes classes de rotações:

1- Rotações ortogonais (rígidas) – são utilizadas quando as novas componentes não são correlacionadas (Stevens J, 1996). Estas rotações são apenas rotações para um novo conjunto de eixos coordenados (perpendiculares entre si) no mesmo subespaço “produzido” pelas componentes principais (Dunteman G, 1989).

2- Rotações oblíquas – podem ser utilizadas mesmo no caso das as novas componentes estarem correlacionados.

A rotação embora não altere a correlação entre as variáveis observáveis, perde uma das propriedades das componentes principais, nomeadamente a que explica a máxima proporção da variação das variáveis observáveis. Com a rotação, a primeira componente deixa de ser a combinação linear das variáveis que explicam a maior variância dos dados (Pestana M, Gageiro J, 2005).

1. Rotações ortogonais:

A rotação ortogonal produz componentes que não se correlacionam entre si, as quais são interpretadas a partir dos seus pesos ou contribuições (*loadings*), que variam entre zero e 1 ou -1 (Pestana M, Gageiro J, 2005). Este método inclui as rotações *varimax*, *quartimax* e *equimax*.

Serão abordadas de forma resumida estas três rotações:

Quartimax – Aqui a ideia é “limpar” as variáveis. Isto é, a rotação é feita de modo a que cada variável se correlacione maioritariamente com uma componente. De seguida essa variável pode ser considerada como sendo uma medida relativamente pura da componente. O problema desta abordagem acontece quando a maioria das variáveis tende a correlacionar-se com uma componente única, tornando a interpretação desta componente difícil (Stevens J, 1996).

Varimax – Segundo Dunteman G (1989), este é possivelmente o método de rotação ortogonal mais popular. Como refere Stevens J (1996), Kaiser H (1960) tomou um caminho diferente, ao determinar uma rotação para “limpar” as componentes, ou seja, com a sua rotação cada componente irá correlacionar-se mais com um pequeno número de variáveis e correlacionar-se pouco ou muito pouco com as restantes variáveis. Geralmente isto torna a interpretação das componentes resultantes mais fácil. De

salientar que quando este tipo de rotação é feita a propriedade relacionada com a máxima variância das componentes originais é destruída. Essencialmente esta rotação recoloca as contribuições de cada variável para aquela componente. Assim, a primeira componente “rodada” não será necessariamente considerada para o máximo total da variância, ou seja, o total da variância considerada para cada componente “rodada” terá de ser calculada de novo (Stevens J, 1996). O método *varimax* tenta maximizar a variância explicada pelas componentes (daí o nome *varimax*), aumentando a contribuição das variáveis que se correlacionam com eles e diminuindo a contribuição das variáveis que se correlacionam menos com estas variáveis (Cramer D, 2003).

Equimax – este método pretende ser uma solução de compromisso entre os dois métodos anteriores. Em vez de se concentrar nas linhas ou nas colunas da matriz das contribuições, o seu objectivo é simplificar simultaneamente linhas e colunas (Reis E, 2001).

2. Rotações oblíquas:

Nestes métodos o pressuposto de independência entre as componentes é retirado. É permitido às componentes que rodem livremente de maneira a simplificarem o agrupamento das variáveis e a interpretação das componentes. Este método tem a vantagem de poder ser aplicado graficamente com relativa segurança (Reis E, 2001).

Depois de efectuada a rotação, torna-se mais simples identificar e interpretar cada componente principal a partir das contribuições das variáveis que a compõem. Quanto mais próximo de um estiver essa contribuição, mais forte a associação entre a dita variável e a componente, enquanto que uma contribuição próxima de zero permite, desde logo, concluir que essa variável pouco contribui para a formação da componente. Em geral são consideradas como significativas as contribuições superiores a 0,5 (Reis E, 2001).

Têm sido propostas inúmeras rotações oblíquas como por exemplo *oblimaz*, *quartimin*, *maxplane*, *orthoblique* (ou rotação Harris-Kaiser), *promax* e *oblimin*. Vários são os autores que concordam que a possibilidade das componentes estarem correlacionadas é muito mais aceitável de assumir na maioria dos casos e, portanto, que as rotações oblíquas são bastante aceitáveis. No entanto a sua interpretação torna-se mais difícil, sendo preciso examinar duas matrizes, pois existem duas formas de

apresentar os resultados. A primeira é uma matriz de padrão das componentes (*factor pattern matrix*), geralmente a única apresentada que mostra a contribuição que cada variável “tem” com cada componente mas não as contribuições que são partilhadas entre as componentes se estas forem correlacionadas. A outra matriz, a matriz da estrutura das componentes (*factor structure matrix*), indica a contribuição total que cada variável “tem” com uma componente. Se as componentes não estiverem correlacionadas estas duas matrizes deverão ser idênticas e nesse caso será mais simples e mais apropriado utilizar o método *varimax* (Cramer D, 2003). Para além destas duas tabelas existe ainda uma tabela de correlação entre as componentes. É a partir da análise desta tabela que se decide se se usa a rotação ortogonal ou oblíqua. Se a correlação for baixa (abaixo de 0,3) então justifica-se o uso da ortogonal, porque os resultados são semelhantes à oblíqua. Para correlações acima de 0,3 os resultados seriam discrepantes.

2.6 Análise de *Clusters*

A análise de *clusters* (aglomerados) abrange diversas técnicas numéricas cujo objectivo principal é classificar os valores de uma matriz de dados em grupos discretos. A técnica classificatória multivariada da análise de *clusters* pode ser utilizada quando se deseja explorar as similaridades entre indivíduos ou entre variáveis definindo-os em grupos, considerando simultaneamente, no primeiro caso, todas as variáveis medidas em cada indivíduo e, no segundo, todos os indivíduos nos quais foram feitas as mesmas medidas. Segundo esse método, desenvolvido inicialmente em Zoologia, por taxonomistas numéricos, procuram-se *clusters* homogêneos de itens, representados por pontos num espaço, num número conveniente de grupos, relacionando-os através de coeficientes de similaridades ou de correspondência. A contribuição mais expressiva para aplicação destes métodos foi dada por Sokal R, Sneath P, em 1963, com o seu livro *Principles of Numerical Taxonomy* (Aldenderfer M, Blashfield R, 1984). Reis E (2001) mostra que um método eficiente para se proceder à classificação biológica, consiste em juntar toda a informação existente sobre um conjunto de organismos, determinar a semelhança existente entre esses organismos e através de um método de análise de *clusters* colocar organismos semelhantes num mesmo grupo. Desta forma, as características de cada grupo seriam analisadas de modo a determinar se se trata ou não de espécies diferentes (Reis E, 2001).

Assim, se duas ou três componentes principais são responsáveis por uma proporção substancial da variação total, pode-se utilizar as representações (*plots*) para identificar visualmente os *clusters*. Considerando então como *cluster* um grupo de observações que estão mais próximas umas das outras do que das observações de outro *cluster* ou grupo (ver **Figura 2.20**) (Dunterman G, 1989).

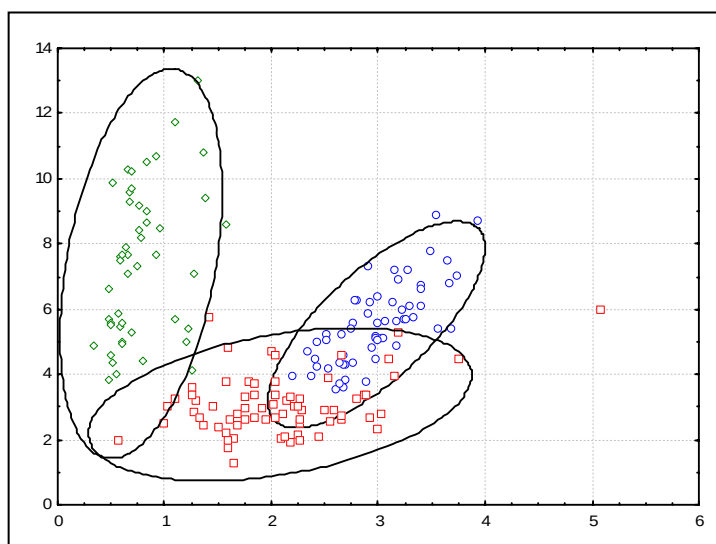


Figura 2.20 Diagrama de Dispersão entre duas componentes principais, onde é possível observar três *clusters*.

A utilização principal da análise de *clusters* é encontrar grupos de entidades similares num conjunto de dados. Não existe uma definição padrão para o termo *cluster*, mas os *clusters* têm certas propriedades, como a densidade, a variância (ou seja, o grau de dispersão dos pontos no espaço desde o centro do *cluster*), a dimensão, a forma e a separação. Com base nestas propriedades Everitt B (1974) definiu *cluster* como sendo uma região contínua de um espaço, contendo uma densidade de pontos relativamente alta, separados por outras regiões contendo uma densidade de pontos relativamente baixa.

Várias técnicas são utilizadas na análise de *clusters*, incluindo a técnica da análise das componentes principais, cuja vantagem é tornar possível representar os *scores* das componentes e procurar visualmente os *clusters* de observações. É possível observar se os *clusters* previamente definidos são homogéneos, distintos ou esteticamente apelativos. Isto porque os algoritmos de aglomeração irão sempre definir *clusters*, mesmo que estes de facto não existam, quando as observações estão igualmente espalhadas pelo espaço. Quando isto ocorre os algoritmos de aglomeração definem apenas uma parte arbitrária das observações em *clusters*. (Dunteman G, 1989).

Os diversos algoritmos de aglomeração não apresentam como resultado, necessariamente, o mesmo número de *clusters*, nem mesmo classificam as observações nos mesmos *clusters*. Uma das razões para esta diferença da solução reside no facto dos métodos de aglomeração terem evoluído com base em fontes díspares, que enfatizam diferentes regras para a formação dos *clusters* (Aldenderfer M, Blashfield R, 1984). Por

estes motivos, a representação dos dados através das duas primeiras componentes principais pode dar muita informação, se estes forem responsáveis por uma grande parte da variância total. Se o objectivo for fazer uma aglomeração arbitrária dos dados, então a aglomeração visual no espaço das duas primeiras componentes principais é tão boa como qualquer outro procedimento. Se as duas primeiras componentes principais não forem responsáveis por uma grande parte da variância total torna-se útil suplementar a informação contida na representação com uma terceira componente principal (Dunteman G, 1989).

A análise de *clusters* é um método que pode ser descrito da seguinte forma: dada um conjunto de n observações para as quais existe informação sobre a forma de p variáveis, o método de análise de *clusters* procede ao agrupamento das observações em função da informação existente, de tal modo que as observações pertencentes a um mesmo grupo sejam tão semelhantes quanto possível e sempre mais semelhantes aos elementos do mesmo grupo do que a elementos dos restantes grupos (Reis E, 2001).

Uma dificuldade inicial é a de não existir uma única via de definição de grupos, isto é, um único critério de partição e/ou agrupamento dos indivíduos ou observações com base numa única medida de (dis)semelhança (Reis E, 2001).

2.6.1 Selecção das Variáveis

A selecção de variáveis comporta dois problemas: um problema substantivo que terá de ser resolvido com o conhecimento prévio sobre o assunto a estudar e que permitirá escolher de entre os dados disponíveis quais os mais importantes na abordagem do problema e um outro, de ordem mais estatística, que tem a ver com o tipo de variáveis utilizadas, sobretudo quando estas estão definidas em diferentes unidades de medida (Reis E, 2001).

Quando as variáveis se encontram definidas na mesma escala de medida (contínuas, rácios, ordinais, nominais ou binárias) é necessário escolher um algoritmo de aglomeração que mais se adequa às variáveis. Quando, pelo contrário, as variáveis se encontram definidas em diferentes escalas de medida e se aplica a análise de *clusters* sem uma padronização prévia, qualquer que seja o método utilizado vai reflectir sobretudo o peso das variáveis que maiores valores e maior dispersão apresentam (Reis E, 2001).

2.6.2 Medidas de Semelhança

Uma das etapas compreendidas na análise de *clusters* é a selecção de uma medida de semelhança ou distância entre cada par de observações. Quando duas observações são semelhantes, o valor da medida das distâncias é pequeno e o valor da medida das semelhanças é grande, porque enquanto que as distâncias medem o afastamento entre duas observações, as semelhanças medem quão perto estão essas observações entre si (Pestana M, Gageiro J, 2005).

Muitos têm sido os termos criados para descrever as características importantes relativas à estimação da similitude. O próprio termo “semelhança” não está imune aos variados significados e os seus sinónimos são “proximidade” e “associação” (Aldenderfer M, Blashfield R, 1984).

Segundo Tversky A (1977), citado em Reis E (2001), a análise teórica das relações de semelhança tem sido dominada pelos modelos geométricos. Estes modelos representam objectos como pontos num qualquer espaço de coordenadas de forma que as dissemelhanças observadas entre objectos correspondam a distâncias métricas entre os respectivos pontos (Aldenderfer M, Blashfield R, 1984 e Reis E, 2001). Os métodos de classificação exigem que os índices de semelhança respeitem as seguintes propriedades métricas:

- Simetria: dados dois objectos, x e y , a distância entre eles verifica a propriedade

$$d(x,y) = d(y,x) \geq 0. \quad (2.4)$$

Ou seja, a distância entre x e y é igual à distância entre y e x e é sempre maior ou igual a 0.

- Desigualdade triangular: dados três objectos x , y e z , as distâncias entre eles satisfaçam a propriedade

$$d(x,y) \leq d(x,z) + d(z,y). \quad (2.5)$$

Ou seja, geometricamente, esta desigualdade diz que o comprimento de qualquer lado de um triângulo é menor ou igual à soma dos comprimentos dos outros dois.

- Diferenciabilidade de não idênticos: dados dois objectos x e y ,

$$d(x,y) \neq 0 \Rightarrow x \neq y. \quad (2.6)$$

Ou seja, se a distância entre x e y for diferente de zero, então x é diferente de y .

- Indiferenciabilidade de idênticos: dados dois objectos idênticos x e x' ,

$$d(x,x') = 0. \quad (2.7)$$

Ou seja, a distância entre dois objectos idênticos é zero.

Sneath P, Sokal R (1973) subdividiram as medidas de semelhança em quatro categorias:

- coeficientes de correlação;
- medidas de distância;
- coeficientes de associação; e
- medidas de semelhança probabilística.

Todas estas medidas têm vantagens e desvantagens, mas as mais utilizadas nas ciências sociais são as duas primeiras, pelo que serão de seguida resumidamente descritas, com base em Aldenrfer M, Blashfield R (1984) e Reis E (2001).

1. Coeficientes de correlação:

Estes coeficientes são caracterizados por serem de fácil interpretação geométrica e são das medidas mais utilizadas nas ciências sociais, nomeadamente o coeficiente de Pearson.

Este coeficiente, em geral, mede o grau de associação linear entre duas variáveis e o seu valor varia entre -1 e 1, com o valor zero a significar não existência de correlação entre as observações. A correlação é elevada quando existe uma relação forte de tipo linear entre os perfis das observações.

Este coeficiente é particularmente insensível às diferenças de escalas das variáveis uma vez que o cálculo da média de todas as variáveis para cada observação impõe a padronização prévia dessas variáveis. No entanto, é sensível às diferenças de forma de cada observação e à dispersão dos valores das variáveis em torno das respectivas médias.

2. Medidas de distância:

Devido ao seu apelo intuitivo estas medidas são as mais populares. Tecnicamente, são melhores descritas como medidas de dissemelhança, uma vez que os coeficientes mais conhecidos demonstram semelhança com grandes valores, mas as medidas de distância operam em escala inversa. Existem várias medidas que podem ser utilizadas como medidas de distância ou dissemelhança entre os elementos de uma matriz de dados.

Sendo i e j duas observações e X_i as variáveis com $(i = 1, 2, \dots, p)$, então podem-se destacar como mais utilizadas as seguintes:

- Distância euclideana, onde a distância entre as duas observações é dada por

$$d_{ij} = \sqrt{\sum_{i=1}^p (X_{li} - X_{ji})^2}; \quad (2.8)$$

- Quadrado da distância euclideana, onde a distância entre as duas observações é dada por

$$d_{ij}^2 = \sum_{i=1}^p (X_{li} - X_{ji})^2; \quad (2.9)$$

- Distância absoluta, onde a distância entre as duas observações é dada por

$$d_{ij} = \sum_{i=1}^p |X_{li} - X_{ji}|; \quad (2.10)$$

- Distância de Minkowski, onde a distância entre as duas observações é dada por

$$d_{ij} = \left[\sum_{i=1}^p |X_{li} - X_{ji}|^r \right]^{1/r}; \quad (2.11)$$

- Distância de Mahalanobis ou distância generalizada, onde é considerada a matriz de covariância, Σ , para o cálculo das distâncias; e

- Distância de Chebishev, onde a distância entre as duas observações é dada por

$$d_{ij} = \max_i |X_{li} - X_{ji}|. \quad (2.12)$$

Apesar da importância destas medidas de distância, existem vários problemas na sua utilização, sendo o mais importante o efeito que as diferenças das escalas das variáveis provocam sobre os valores das distâncias. As variáveis que apresentam variações e unidades de medida elevadas, facilmente anularão o efeito das outras variáveis. Mais uma vez a padronização das variáveis possibilita resolver este problema.

2.6.3 Métodos de Aglomeração

Os métodos de aglomeração (*clustering*) têm como objectivo dividir em subconjuntos, homogêneos, conjuntos de observações, de maneira que as tais observações pertencentes a um mesmo aglomerado (ou *cluster*) sejam similares com respeito às características (atributos) que forem medidas em cada elemento. Por outras palavras, os métodos de aglomeração podem ser caracterizados como qualquer procedimento estatístico que, utilizando um conjunto finito e multi-dimensional de informações, classifica seus elementos em grupos restritos homogêneos internamente.

Segundo Reis E (2001) na aplicação dos métodos de aglomeração, é necessário identificar a técnica de análise mais apropriada. É possível dividir as técnicas em vários grupos:

- Técnicas hierárquicas – aglomeração hierárquica;
- Técnicas de optimização – partição iterativa;
- Técnicas de densidade;
- Outras técnicas

Cada uma destas técnicas representa uma perspectiva diferente na criação dos *clusters*, e os resultados obtidos através destes métodos, com os mesmos dados, podem ser diferentes. De seguida são apresentadas de forma resumida as técnicas e mais detalhadamente os dois métodos mais utilizados em ciências sociais, com base na abordagem de Aldenrfer M, Blashfield R (1984) e Reis E (2001).

1. Aglomeração hierárquica:

Os métodos de análise de *clusters* hierárquicos são os mais frequentemente utilizados. Estes métodos processam-se através de uma sucessão de aglomeração de observações. Parte-se de uma situação de máxima desagregação, em que o número de *clusters* é idêntico ao número de espaços elementares e, através da aglomeração dos espaços mais semelhantes, vai-se progressivamente diminuindo o número de *clusters* até uma máxima agregação em que todas as observações sejam agrupadas num único *cluster*. Esta sequência de aglomerações pode ser representada visualmente por um diagrama chamado dendograma, que é uma representação gráfica da estrutura hierárquica. De entre os vários métodos desta família todos requerem exactamente $N - 1$ passos (onde N se refere ao número de observações).

Apesar da simplicidade destes métodos existem alguns problemas como o facto de requererem o cálculo e armazenamento da chamada matriz de similitude que é potencialmente grande.

Estes métodos são distintos primariamente pelas diferentes regras de formação dos *clusters*. Existem pelo menos doze formas de ligação propostas, sendo as quatro mais populares as seguintes:

- *single linkage* ou *critério do vizinho mais próximo* – segundo Aldenrfer M, Blashfield R (1984), este método, descrito por Sneath P (1957), forma os *clusters*

seguindo a seguinte regra: uma observação é associada a um *cluster* existente se pelo menos um dos membros desse *cluster* existente está ao mesmo nível de similaridade que a observação considerada. As conexões são baseadas apenas em ligações simples entre as observações e os *clusters*. Ou dito de outro modo, dados dois grupos: um formado pelas observações l e j e outro grupo formado pela observação k , a distância entre os dois grupos é a menor distância entre os elementos dos dois grupos:

$$d = \min \{d_{lk}; d_{kj}\}. \quad (2.13)$$

Desta forma, qualquer *cluster* é definido como o conjunto de observações em que qualquer observação é mais semelhante a pelo menos uma outra observação do mesmo *cluster* do que a qualquer observação de outro *cluster*. Apesar da sua simplicidade e de ser um dos poucos métodos que não será afectado pela transformação dos dados, tem-se mostrado que tem a tendência de formar longos *clusters* (no sentido, em que agregam muito e diferenciaram pouco) porque, segundo Lance G e Williams W (1967), citados em Reis E (2001), cada observação terá mais tendência para se agrupar a um grupo já definido do que para formar o núcleo de um novo *cluster*.

- *complete linkage* ou *critério do vizinho mais afastado* – este método é logicamente oposto ao anterior, aqui a regra é mais rigorosa e dita que a observação candidata à inclusão num *cluster* existente deve ter um determinado nível de similitude com todos os membros desse *cluster*. Ou seja, dados dois grupos: um formado pelas observações l e j e outro grupo formado pela observação k , a distância entre os dois grupos é a maior distância entre os seus elementos:

$$d = \max \{d_{lk}; d_{kj}\}. \quad (2.14)$$

Desta forma, cada *cluster* passa a ser definido como um conjunto de observações em que cada uma é mais semelhante a todas as restantes observações do *cluster* do que a qualquer das observações dos restantes *clusters*. Sendo assim com este método tende-se a encontrar *clusters* relativamente compactos compostos por casos altamente similares.

- *average linkage* ou *critério da média* – segundo Aldenrfer M, Blashfield R (1984), este método proposto por Sokal R e Michener C (1958) foi desenvolvido como antídoto aos dois métodos extremos anteriores. Apesar de existir um grande número de variantes deste método, cada um deles essencialmente computa uma média de similitude da observação, em consideração, com todas as observações existentes no *cluster* e,

subsequentemente, aglomera essa observação no *cluster* se um dado nível de similitude for atingido utilizando esse valor médio. A variante mais comumente utilizada é a média aritmética das similaridades entre as observações.

- *Ward's method* – este método é designado para otimizar a mínima variância nos *clusters* e baseia-se na perda de informação resultante do agrupamento das observações e medida através da soma dos quadrados dos desvios das observações individuais relativamente à médias dos grupos em que são classificadas. Esta função objectiva é dada por:

$$x_i^2 - \frac{1}{(\sum x_1)^2}, \quad (2.15)$$

onde x_i é o score da i -ésima observação.

No primeiro passo deste processo de aglomeração, quando cada observação pertence ao seu próprio *cluster*, o valor da equação (2.15) é 0. Este método opera juntando os grupos ou observações que resultam no aumento mínimo de (2.15). A principal desvantagem apontada por alguns autores, como Williams W (1971), nesta metodologia é a sua tendência em criar *clusters* de tamanhos semelhantes e pequenos e a sua tendência para encontrar grupos esféricos, mesmo quando a representação gráfica dos dados revela grupos com diferentes formas. Este método tem sido virtualmente ignorado pelas ciências biológicas e utilizado em larga escala pelas ciências sociais.

2. Partição Iterativa:

Nestes métodos de optimização é definido um critério de agrupamento que indica qual deverá ser o *cluster* onde cada observação será incluída, pressupondo que todas as observações pertencem a um número predefinido de *clusters*.

Por outras palavras este método baseia-se directamente na escolha antecipada de um número de *clusters* que conterão todas as observações. Procede-se, em seguida, a uma divisão de todas as observações pelos k *clusters* preestabelecidos e a melhor partição das observações será aquela que optimiza o critério escolhido. Uma descrição sumária destes métodos é feita em seguida:

- Inicialmente efectua-se uma partição dos dados num número específico de *clusters* e calculam-se os núcleos de cada *cluster*.

- Calculam-se as distâncias euclidianas entre as observações e os núcleos dos *clusters*, e associam-se as observações aos *clusters* que têm os núcleos mais próximos.

- Ao se adicionarem as observações aos *clusters* estes vão se transformar, tendo, portanto, um novo núcleo. Desta forma calculam-se os novos núcleos dos *clusters*, sendo apenas obtidos novos *clusters* após uma passagem completa por todas as observações. E alterna-se entre os dois passos anteriores até que nenhum ponto modifique os *clusters*. Ao contrário da aglomeração hierárquica estes métodos não necessitam do cálculo de matrizes de similitude e, mais particularmente, o facto de estes métodos requererem mais do que uma passagem pelos dados, compensa a pobre partição inicial destes dados. Contudo, estes métodos têm uma grande limitação que é o facto de que para se encontrar a partição óptima do conjunto de dados é necessário formar todas as possíveis partições, o que torna necessária uma capacidade computacional considerável.

As propriedades destes métodos podem ser resumidas nos seguintes pontos:

- Partição inicial - existem duas maneiras para começar um método iterativo: a definição das estimativas dos núcleos dos *clusters* (*seed points*) ou a selecção de uma partição de começo (*starting partition*) apropriada. Na primeira, as observações são associadas aos *clusters* com os núcleos mais próximos na primeira passagem pelos dados. A segunda, envolve a especificação da primeira cessão dos *clusters*, em cujo procedimento, os núcleos de cada *cluster* são definidos como a média multivariada das observações desse *cluster*.

- Tipo de passagem – este ponto está relacionado com a forma como as observações são associadas aos *clusters*, existindo, dois tipos principais: *K-means* e *hill climbing*.

O método *K-means* envolve a nova cessão das observações para o *cluster* com o núcleo mais próximo. Estas passagens podem ser combinatórias ou não-combinatórias, sendo que as combinatórias requerem o cálculo dos núcleos dos *clusters* sempre que um novo “membro” lhe é associado e as não-combinatórias apenas calculam os novos núcleos quando é feita uma passagem completa pelos dados. Outra importante distinção é que podem ser inclusivos ou exclusivos. Os métodos exclusivos removem a observação em consideração do *cluster* “pai” quando o núcleo é computado e os métodos inclusivos, incluem-no. Estes procedimentos tentam minimizar a variância em cada *cluster*.

Nas passagens *hill climbing*, antes de se atribuírem novas observações a um *cluster* com base nas suas distâncias dessas observações ao núcleo do *cluster*, movimentam-se as observações de forma a ver se essa movimentação proposta optimiza ou não o valor do critério estatístico escolhido.

- Critério estatístico - as passagens requerem critérios estatísticos e todos os critérios relacionam-se com a detecção da homogeneidade dos *clusters* representados num espaço multivariado.

Para além da limitação já referida, este método de aglomeração tem outra desvantagem: o facto de não existir nenhuma garantia de que o algoritmo forneça uma optimização global e não apenas um optimização local. A sua aplicabilidade só se torna possível com a definição, *à priori*, do número de grupos pretendido.

3. Técnicas de densidade:

Estas técnicas são resultado natural do conceito de que um *cluster* é uma região de grande densidade de pontos. Desta forma, nesta metodologia os grupos são formados através da procura de regiões que contenham uma concentração relativamente densa de observações. Duas estratégias básicas destas técnicas são técnicas baseadas numa variante do *critério do vizinho mais próximo* (em que em vez de juntarem observações a grupos já formados, têm regras que permitem iniciar novos *clusters*) e técnicas baseadas em distribuições de probabilidade multivariada.

4. Outras técnicas:

Este grupo inclui um conjunto de outras técnicas, menos utilizadas, sendo algumas das quais, técnicas que permitem que haja sobreposição dos grupos (*fuzzy clusters*).

2.6.4 Selecção do Número de Clusters

Uma vez que a análise de *clusters* tem por objectivo criar grupos homogéneos, é necessário encontrar um conjunto de procedimentos que possam ser utilizados para determinar quantos *clusters* existirão. Por exemplo, a estrutura em árvore de um dendograma sugere que podem estar presentes nos dados vários grupos diferentes, a questão é em que ponto “cortar” essa árvore de forma a encontrar um número óptimo de grupos.

Infelizmente, este passo fundamental é um dos problemas ainda não resolvidos da análise de *clusters* e duas das mais importantes razões para este facto é a falta de aptidão das chamadas hipóteses nulas e a natureza complexa da distribuição aleatória das observações. Em algumas áreas, como as ciências biológicas este problema de definir o número de *clusters* não tem uma grande importância uma vez que o objectivo da análise é explorar os padrões gerais das relações entre as entidades. No caso das ciências sociais existem duas aproximações para resolver este problema: os procedimentos heurísticos e os testes formais (Aldenderfer M, Blashfield R, 1984).

Os procedimentos heurísticos são os procedimentos mais comuns. No nível mais básico, tem-se o corte, numa árvore hierárquica, através de uma inspecção subjectiva dos diferentes níveis da árvore. Este procedimento básico é pouco satisfatório pois depende das necessidades e das opiniões do investigador. Um procedimento heurístico mais formal consiste em representar graficamente o número de *clusters* implícitos num dendograma em função do coeficiente de fusão, que é o valor numérico em que as observações se agrupam num *cluster*, e é representado ao longo do eixo dos *yy* do dendograma (ver **Figura 2.21** onde o abrandamento na curva mostra a existência de três – ou dois – *clusters*).

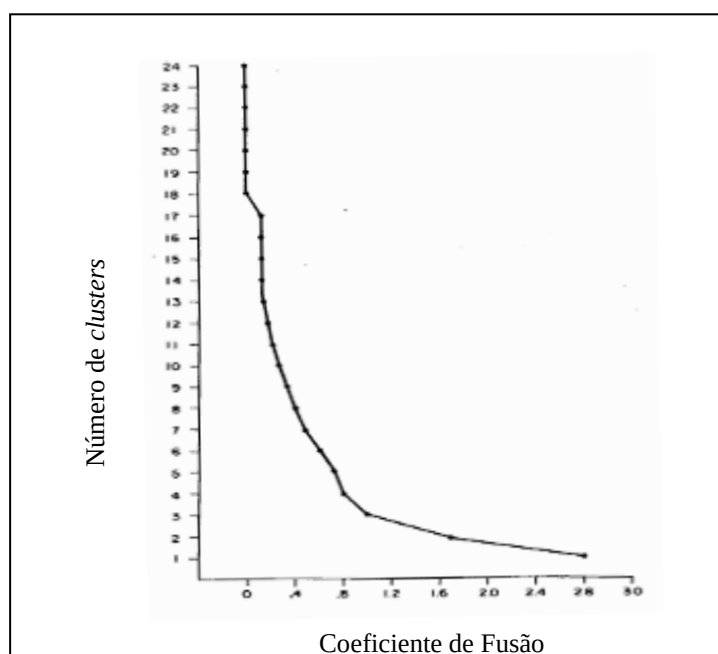


Figura 2.21 Representação gráfica do número de *clusters* versus coeficiente de fusão.

Fonte: Aldenderfer M, Blashfield R (1984).

Quanto aos testes estatísticos formais, estes não têm sido muito apreciados devido aos problemas da complexidade da distribuição das observações multivariadas, mas alguns têm sido aceites. A hipótese nula que é mais comum utilizar-se nestes testes é a que pressupõe que os dados em estudo são uma amostra ao acaso de uma distribuição normal multivariada. Segundo Aldenrfer M, Blashfield R (1984) uma hipótese alternativa, desenvolvida por Lee K (1979), parte do princípio que os dados são uma amostra ao acaso de uma distribuição uniforme. Os testes baseados nesta hipótese utilizam a função referida no método de *Ward's* que parece ser um ponto de partida útil na determinação de possíveis diferenças entre os *clusters*. Como aspecto negativo, este teste só é capaz de lidar com uma variável única (Aldenrfer M, Blashfield R, 1984).

Qualquer que seja o procedimento a escolher deve-se ter em atenção que estes testes têm sido pouco estudados intensivamente, poucos são compreendidos ou então são heurísticos e que os seus resultados devem ser utilizados com os cuidados necessários.

3. MATERIAIS E MÉTODOS

3.1 Materiais

Os materiais utilizados neste estudo podem ser divididos em três grupos: bases de dados, mapas e software. As bases de dados por sua vez podem ser divididos em três subgrupos: dados relacionados com as fracturas do colo do fémur, dados demográficos e dados relacionados com a mortalidade infantil.

3.1.1 Bases de Dados

1. Dados relacionados com as fracturas de colo do fémur:

Os dados sobre as fracturas do colo do fémur utilizados referem-se ao período de 2000 a 2002 e foram disponibilizados pelo Instituto de Gestão Informática e Financeira da Saúde – IGIF, do Ministério da Saúde de Portugal. Contêm informações sobre todos os episódios de internamento hospitalar, em hospitais da rede pública, de doentes com fracturas do colo do fémur, em Portugal Continental. Esta base de dados contém as seguintes variáveis: código do doente (aleatório de maneira a manter o anonimato), género, idade, local de residência (freguesia, concelho e distrito), data de internamento e data de saída, causa principal do internamento e mais 19 causas secundárias, classificadas de acordo com o Código Internacional de Doenças, 9ª revisão, Modificação Clínica – CID9-MC, diagnóstico principal e mais 19 secundários, também classificados pela CID9-MC, procedimentos clínicos durante o internamento, tipo de internamento (se urgente ou programado), destino após a alta (óbito, para casa, para outra instituição de internamento), Grupo de Diagnóstico Homogéneo (GDH) e custos do internamento.

O GDH, segundo o Tribunal de Contas português, é um sistema de classificação de doentes internados em hospitais de agudos, em classes clinicamente coerentes e homogéneas, do ponto de vista do consumo de recursos. Estas classes são definidas a partir de um conjunto de variáveis que caracterizam clinicamente os doentes e que explicam os custos associados à sua estadia no hospital – diagnósticos, intervenções cirúrgicas e outros actos médicos relevantes.

Os dados utilizados referem-se aos internamentos em todas as instituições do Serviço Nacional de Saúde (SNS). Como as fracturas do colo do fémur requerem hospitalização e quase sempre intervenção cirúrgica, com elevados custos associados, a maioria dos

doentes utilizam hospitais do SNS e desta maneira a base de dados do IGIF abrange quase a totalidade de fracturas de colo do fémur, em Portugal Continental. Com relação aos arquipélagos dos Açores e da Madeira optou-se pela exclusão destas regiões autónomas do estudo, porque a base de dados do IGIF não é completa no que diz respeito ao número de ocorrências de fracturas de colo do fémur nestas regiões.

Os dados utilizados neste estudo, já tinham sido anteriormente trabalhados por Alves S (2004) e por este motivo já estavam georeferenciados até ao nível de freguesia, classificados em grupos etários de 5 em 5 anos a partir dos 50 anos de idade, bem como seleccionados apenas os registos referentes às fracturas do colo do fémur (CID9-MC: 820-x), causadas por baixo impacto (também classificadas de acordo com a CID9-CM).

2. Dados demográficos:

Foram utilizados os dados provenientes do Censo Demográfico de 2001 adquiridos ao Instituto Nacional de Estatística – INE, a nível de concelho, com variáveis sobre a habitação, as famílias e a população.

O Censo demográfico é um inquérito sistemático, realizado normalmente a cada 10 anos, a um universo de indivíduos que vivem num determinado país e que fornece informação demográfica e socioeconómica sobre as condições de vida dos habitantes. Em Portugal, O INE é o órgão responsável pela realização dos Censos, tendo o último sido realizado em 2001.

As variáveis seleccionadas e posteriormente trabalhadas para elaborar um perfil socioeconómico dos concelhos, estão relacionadas abaixo:

Habitação - Tipo de Ocupação:

- Alojamentos clássicos arrendados e subarrendados, ocupados como residência habitual, segundo o escalão de renda;
- Alojamentos clássicos propriedade dos ocupantes, ocupados como residência habitual, segundo a existência de encargos por compra e respectivos escalões;
- Alojamentos clássicos, ocupados como residência habitual, segundo a época de construção dos edifícios;
- Alojamentos clássicos, ocupados como residência habitual, segundo o número de divisões;

- Alojamentos clássicos ocupados como residência habitual, segundo o número de famílias clássicas e o número de pessoas residentes;
- Alojamentos clássicos, segundo a forma de ocupação.

Habitação - Características da Construção:

- Edifícios, segundo os principais materiais utilizados na construção;
- Edifícios, segundo o número de pavimentos;
- Alojamentos familiares, ocupados como residência habitual, segundo instalações existentes (electricidade e sanitárias) nos alojamentos;
- Alojamentos clássicos, ocupados como residência habitual, divisões, famílias clássicas, pessoas residentes e indicadores de ocupação.

Famílias:

- Famílias clássicas, segundo a condição perante a actividade económica e o género do representante da família;
- Famílias clássicas, segundo a dimensão;
- Famílias clássicas, segundo o número de pessoas com actividade económica e pessoas a cargo;
- Famílias clássicas, segundo o tipo de família na base da estrutura etária dos seus membros e número de crianças;
- Famílias clássicas, segundo o tipo de família.

População:

- População residente, por concelhos, por sexo e grupos etários de 5 anos.
- População residente desempregada em sentido lato, segundo o principal meio de vida;
- População residente empregada segundo grupos de profissões;
- População residente por condição perante a actividade económica, género e grupos etários;
- População residente, desempregada em sentido lato, segundo o grupo etário;
- População residente, desempregada em sentido lato, segundo o nível de escolaridade;
- População residente em 1991 e 2001, segundo os grupos etários e sua evolução entre 1991 e 2001;

- População residente segundo o nível de ensino atingido, género e taxa de analfabetismo (1991 e 2001);
- População residente, segundo a nacionalidade e género;
- População residente, segundo as migrações (relativamente a 1999/12/31), por concelho de residência habitual em 2001/03/12;
- População residente, segundo o estado civil e género, por grupo socioeconómico;
- População residente;
- Índice *Per Capita*;
- Índice de dependência de jovens;
- Índice de dependência de idosos.

3. Dados relacionados com a mortalidade infantil:

Os dados relacionados com a mortalidade infantil referem-se às décadas de 30, 40 e 50 do século XX e foram adquiridos ao INE. A mortalidade infantil relaciona o número de óbitos em bebés com menos de um ano de idade, dividido pelo número de nados vivos no mesmo período de tempo. Utilizou-se os concelhos como unidade geográfica de análise.

A justificação para a escolha destes dados deve-se ao facto de se pretender com este estudo caracterizar os concelhos socioeconomicamente não apenas nos anos actuais mas também há setenta, sessenta e cinquenta anos atrás. Esta análise socioeconómica do início do século XX é importante pois as pessoas que estão a sofrer fracturas do colo do fémur agora, eram crianças e estavam na fase de adquirir massa óssea, precisamente à setenta, sessenta e cinquenta anos atrás. Os dados censitários de 2001 foram utilizados para realizar a caracterização socioeconómica actual.

Sendo, sabidamente, a mortalidade infantil um bom indicador das condições socioeconómicas da população e por não estarem disponíveis os dados censitários do início e meados do século passado, optou-se por utilizar a mortalidade infantil como indicador socioeconómico.

A mortalidade infantil foi analisada ano a ano, desde 1930 até 1959. Os dados relativos aos anos 30 encontravam-se agrupados por distrito, enquanto que os outros anos estavam agrupados por concelho. Estas bases de dados continham informação relativa ao número de nados vivos e ao número de óbitos até ao primeiro ano após o nascimento.

3.1.2 Mapas

Foi utilizada a Carta Administrativa Oficial de Portugal, por concelhos, em formato digital, obtida no site do Instituto Geográfico, IGEO

(<http://www.igeo.pt/produtos/cadastro/caop/inicial.htm>).

3.1.3 Software

Para a entrada, gestão, análise e apresentação dos dados foram utilizados os seguintes programas:

- Microsoft EXCELTM e Microsoft ACCESSTM como ferramentas de suporte para o trabalho das bases de dados;
- SPSSTM como ferramenta para a análise das variáveis seleccionadas para estudo; e
- ArcMapTM (9.1) para o mapeamento dos resultados.

3.2 Métodos

As etapas de trabalho consistiram na estruturação, geocodificação, georeferenciação, análise dos dados e apresentação dos resultados em mapas, gráficos e tabelas.

3.2.1 Cálculo dos Indicadores demográficos

Os dados provenientes dos Censos foram estruturados em tabelas bidimensionais, onde cada linha corresponde a um concelho e cada coluna corresponde a um atributo do concelho, ou seja, todas as informações referentes aos concelhos foram armazenadas nas colunas da tabela. A geocodificação consistiu na introdução dos códigos (geocódigos) oficiais dos concelhos (também chamados chaves primárias), que foram utilizados para efectuar o relacionamento entre as tabelas e para relacionar os dados das tabelas aos mapas.

O georeferenciamento dos dados consistiu em, como o próprio nome indica, referenciar os dados a uma localização na superfície terrestre, que no caso deste estudo, foi o limite do concelho. Assim, através dos geocódigos, os dados censitários foram ligados ao mapa de concelhos, ou seja, foram georeferenciados por concelho.

Após a estruturação e geocodificação de todas as tabelas provenientes do Censo 2001, foram seleccionadas as variáveis que teriam mais interesse para este trabalho, de acordo com as informações obtidas durante a fase de revisão bibliográfica de maneira a calcular os indicadores a serem usados na Análise de Componentes Principais.

Inicialmente foram seleccionadas as tabelas relacionadas com:

- Famílias clássicas: Segundo o tipo de família na base da estrutura etária dos seus membros e número de crianças.

De todas as variáveis contidas nesta tabela foi apenas utilizada aquela relacionada com o número de indivíduos (do género masculino, feminino e de ambos os sexos) com 65 anos ou mais que viviam sozinhos e a partir desta variável foi calculada a percentagem de indivíduos com 65 anos ou mais que viviam sozinhos (em relação à população total do concelho), para todos os concelhos, de Portugal Continental.

- Alojamentos familiares: Ocupados como residência habitual, segundo instalações existentes. Estas variáveis focam o tipo de instalação da água (canalizada ou não canalizada) e o tipo de instalação de aquecimento.

Desta tabela foram utilizadas todas as variáveis e calculadas as respectivas percentagens em relação ao total geral de alojamentos ocupados, por concelho.

- Alojamentos clássicos: ocupados como residência habitual, divisões, famílias clássicas, pessoas residentes e indicadores de ocupação.

Desta tabela foram utilizadas as variáveis referentes às médias de: divisões por alojamentos; famílias por alojamento; pessoas por alojamento; e pessoas por divisão. Estas variáveis já se encontravam calculadas pelo que não foi necessário nenhum tipo de preparação.

- População residente: segundo o nível de escolaridade. Estas variáveis indicam o número de indivíduos com os diferentes níveis de escolaridade, desde os indivíduos sem nenhum nível de ensino até ao nível do ensino superior e pós-graduações. Em cada nível de escolaridade existe ainda a distinção entre completo, incompleto ou a frequentar.

Aqui foram utilizadas todas as variáveis para calcular as respectivas percentagens em relação ao número total de indivíduos por concelho.

- População residente segundo o estado civil (solteiro, casado com ou sem registo, viúvo, separado e divorciado) e género.

Desta tabela foram utilizadas todas as variáveis para calcular as percentagens em relação ao número total de pessoas, por género, por concelho.

- População residente com 15 ou mais anos, segundo o principal meio de vida e género.

Desta tabela foram utilizadas todas as variáveis para calcular as respectivas percentagens em relação ao número de indivíduos com 15 anos ou mais, por género e por concelho.

- População residente, segundo a nacionalidade e género. Esta tabela contém informação referente à população portuguesa e estrangeira residente em cada concelho. A população estrangeira apresenta-se dividida em: população Europeia (proveniente da

Alemanha, Espanha, França, Reino Unido ou outros países da União Europeia), população Africana (proveniente do países PALOP ou outros países africanos) e outros.

Desta tabela foram utilizadas todas as variáveis para calcular as respectivas percentagens em relação ao número total de indivíduos por concelho.

Além das informações sobre a estrutura da população, dos alojamentos e das famílias, provenientes do Censo, foram ainda seleccionadas outras variáveis, também disponíveis no sítio do INE, relativas ao concelho. Este conjunto de variáveis utilizadas é:

- Índice de dependência de jovens (ou seja, a relação entre a população jovem e a população em idade activa, definida habitualmente como o quociente entre o número de pessoas com idades compreendidas entre os 0 e os 14 anos e o número de pessoas com idades compreendidas entre os 15 e os 64 anos) e índice de dependência de idosos (ou seja, a relação entre a população idosa e a população em idade activa, definida habitualmente como o quociente entre o número de pessoas com 65 ou mais anos e o número de pessoas com idades compreendidas entre os 15 e os 64 anos).
- Índice Per Capita: esta variável representa o poder de consumo e é calculada em função de uma média de consumo nacional, com a qual todos os concelhos são comparados. Assim, concelhos com valor abaixo de 1, têm um poder de consumo menor do que a média nacional, enquanto concelhos com índice *per capita* maior do que 1, têm um poder de consumo maior do que a média nacional.
- Alojamentos clássicos arrendados e subarrendados, ocupados como residência habitual, segundo o escalão de renda: estas variáveis foram reagrupadas em três variáveis, cada uma correspondendo ao nível das rendas – baixo (de 14,96€ até 199,51€), médio (de 199,52€ a 399,03 €) e alto (399,04€ ou mais) – a partir das quais foram calculadas as percentagens de alojamentos em cada uma das classes com relação ao número total de alojamentos clássicos ocupados, arrendados e subarrendados, como residência habitual.
- População residente empregada segundo grupos de profissões. A classificação Nacional das Profissões (CNP) é constituída por sete grandes grupos incluindo, cada um deles, as profissões de algum modo relacionadas pela formação exigida ou pelo tipo de actividade:

- 1- Pessoal de Profissões Científicas, Técnicas, Artísticas e de Profissões Similares;
- 2- Directores e Quadros Superiores Administrativos;
- 3- Pessoal Administrativo e Trabalhadores Similares;
- 4- Pessoal do Comércio e Vendedores;
- 5- Pessoal dos Serviços de Protecção e Segurança, dos Serviços Pessoais e Domésticos e Trabalhadores Similares;
- 6- Agricultores, Criadores de Animais, Trabalhadores Agrícolas e Florestais, Pescadores e Caçadores;
- 7/8/9- Trabalhadores das Indústrias Extractiva e Transformadora e Condutores de Máquinas Fixas e de Transporte.

Optou-se por utilizar estas variáveis reagrupadas em três grandes grupos: um grupo que agrega as profissões com CNP 1,2 e 3; outro grupo que agrega as profissões com CNP 4 e 5 e o terceiro grupo agregando as restantes profissões. Em seguida foram calculadas as respectivas percentagens de pessoas empregadas em cada um destes grandes grupos em relação ao número total de pessoas residentes empregadas por concelhos.

- Estrutura etária da população. Estes dados encontravam-se divididos em grupos de 5 anos e foram reagrupados, por género, em grupos etários mais alargados. Primeiramente dividiu-se a população em dois grupos etários: população com menos de 50 anos, e população com 50 anos ou mais. Posteriormente dividiu-se em quatro grupos etários: população com menos de 25 anos, população com idades compreendidas entre os 25 e os 49 anos, população com idades compreendidas entre os 50 e os 74 anos, e população com 75 anos ou mais.

A **Tabela 3.1** apresenta o dicionário das variáveis utilizadas, ou seja, o nome abreviado e o respectivo significado de todas as variáveis utilizadas neste trabalho.

Tabela 3.1 Dicionário de variáveis.

Nome	Designação
P_H/M65	Percentagem de homens/mulheres com 65 anos ou mais que vivem sozinhos
P_CRPU	Percentagem de alojamentos familiares, ocupados como residência habitual, com água canalizada proveniente da rede pública
P_CRPA	Percentagem de alojamentos familiares, ocupados como residência habitual, com água canalizada proveniente de rede particular

P_CFA	Percentagem de alojamentos familiares, ocupados como residência habitual, com água canalizada fora dos mesmos mas no edifício
P_SCFB	Percentagem de alojamentos familiares, ocupados como residência habitual, sem água canalizada nos mesmos ou nos edifícios, proveniente de fontanário ou bica
P_SCP	Percentagem de alojamentos familiares, ocupados como residência habitual, sem água canalizada nos mesmos ou nos edifícios, proveniente do poço ou furo particular
P_SCOT	Percentagem de alojamentos familiares, ocupados como residência habitual, sem água canalizada nos mesmos ou nos edifícios, proveniente de outra forma
P_IBD	Percentagem de alojamentos familiares, ocupados como residência habitual, com instalações de banho ou duche
P_SIBD	Percentagem de alojamentos familiares, ocupados como residência habitual, sem instalações de banho ou duche
P_AC	Percentagem de alojamentos familiares, ocupados como residência habitual, com aquecimento central
P_ANCL	Percentagem de alojamentos familiares, ocupados como residência habitual, com aquecimento não central – lareira
P_ANCF	Percentagem de alojamentos familiares, ocupados como residência habitual, com aquecimento não central – fixo
P_ANCM	Percentagem de alojamentos familiares, ocupados como residência habitual, com aquecimento não central – aparelhos móveis
P_SA	Percentagem de alojamentos familiares, ocupados como residência habitual, sem aquecimento
P_CRPU_P	Percentagem de população residente com água canalizada proveniente da rede pública
P_CRPA_P	Percentagem de população residente com água canalizada proveniente de rede particular
P_CFA_P	Percentagem de população residente com água canalizada fora dos mesmos mas no edifício
P_SCFB_P	Percentagem de população residente sem água canalizada nos mesmos ou nos edifícios, proveniente de fontanário ou bica
P_SCP_P	Percentagem de população residente sem água canalizada nos mesmos ou nos edifícios proveniente, do poço ou furo particular
P_SCOT_P	Percentagem de população residente sem água canalizada nos mesmos ou nos edifícios proveniente, de outra forma
P_IBD_P	Percentagem de população residente com instalações de banho ou duche
P_SIBD_P	Percentagem de população residente sem instalações de banho ou duche
P_AC_P	Percentagem de população residente com aquecimento central
P_ANCL_P	Percentagem de população residente com aquecimento não central – lareira
P_ANCF_P	Percentagem de população residente com aquecimento não central – fixo
P_ANCM_P	Percentagem de população residente com aquecimento não central – aparelhos móveis
P_SA_P	Percentagem de população residente sem aquecimento
P_2_3F_A	Percentagem de alojamentos onde vivem 2 ou 3 famílias
P_DIV	Número médio de divisões por alojamento
P_RES	Número médio de pessoas por alojamento

P_RES_D	Número médio de pessoas por divisão
Tx_A01	Taxa de analfabetismo em 2001
P_1C	Percentagem de população residente com o 1º ciclo completo
P_2C	Percentagem de população residente com o 2º ciclo completo
P_3C	Percentagem de população residente com o 3º ciclo completo
P_SC_C	Percentagem de população residente com o ensino secundário completo
P_SUP_C	Percentagem de população residente com o ensino superior completo
P_SP_IF	Percentagem de população residente com o ensino superior incompleto ou a frequentar
P_VH/M	Percentagem população residente, homem/mulher, viúvo
P_SD_H/M	Percentagem de população residente, com mais de 15 anos, que recebe o subsídio de desemprego
P_RM_H/M	Percentagem de população residente, com mais de 15 anos, que recebe o rendimento mínimo garantido
P_REF_H/M	Percentagem de população residente, com mais de 15 anos, que recebe pensão/reforma
P_AS_H/M	Percentagem de população residente, com mais de 15 anos, que recebe apoio social
CNP_123	Percentagem de população residente empregada nos grupos de profissões CNP 1, 2 ou 3
CNP_45	Percentagem de população residente empregada nos grupos de profissões CNP 4 ou 5
CNP_6_9	Percentagem de população residente empregada nos grupos de profissões CNP 6, 7, 8 ou 9
IDEP_I	Índice de dependência de idosos
IDEP_J	Índice de dependência de jovens
IPERC	Índice <i>Per Capita</i>
P_PT	Percentagem de população residente, em Portugal, portuguesa
P_UE	Percentagem de população residente, em Portugal, europeia
P_AFR	Percentagem de população residente, em Portugal, africana
P_OUT	Percentagem de população residente, em Portugal, de outros países
NIV_B/M/A	Percentagem de alojamentos clássicos arrendados e subarrendados, ocupados como residência habitual, segundo o escalão de renda (baixo, médio ou alto)
P_HM0_49	Percentagem de homens e mulheres com idade inferior a 50 anos
P_HM50_	Percentagem de homens e mulheres com idade igual ou superior a 50 anos
P_H0_49	Percentagem de homens com idade inferior a 50 anos
P_H50_	Percentagem de homens com idade igual ou superior a 50 anos
P_M0_49	Percentagem de mulheres com idade inferior a 50 anos
P_M50_	Percentagem de mulheres com idade igual ou superior a 50 anos
P_HM0_49	Percentagem de homens e mulheres com idade inferior a 50 anos

P_HM50_	Percentagem de homens e mulheres com idade igual ou superior a 50 anos
P_H0_24	Percentagem de homens com idade inferior a 25 anos
P_H25_49	Percentagem de homens com idade compreendida entre os 25 e os 49 anos
P_H50_74	Percentagem de homens com idade compreendida entre os 50 e os 74 anos
P_H75_	Percentagem de homens com idade igual ou superior a 75 anos
P_M0_24	Percentagem de mulheres com idade inferior a 25 anos
P_M25_49	Percentagem de mulheres com idade compreendida entre os 25 e os 49 anos
P_M50_74	Percentagem de mulheres com idade compreendida entre os 50 e os 74 anos
P_M75_	Percentagem de mulheres com idade igual ou superior a 75 anos
P_HM0_24	Percentagem de homens e mulheres com idade inferior a 25 anos
P_HM25_49	Percentagem de homens e mulheres com idade compreendida entre os 25 e os 49 anos
P_HM50_74	Percentagem de homens e mulheres com idade compreendida entre os 50 e os 74 anos
P_HM75_	Percentagem de homens e mulheres com idade igual ou superior a 75 anos

Todas as tabelas foram relacionadas usando o código do concelho de maneira a obter uma única tabela, com o conjunto de variáveis seleccionadas, que foi então utilizado na análise estatística.

Diversos têm sido os estudos que explicam e definem o estatuto socioeconómico através de variáveis relacionadas com o nível de escolaridade, o estado civil, a habitação, a ocupação e/ou rendimento. Desta forma este conjunto de variáveis foi seleccionado com base na revisão bibliográfica feita e na optimização da análise realizada. Por exemplo, não sendo as variáveis relacionadas com as canalizações e as instalações de banho e duche as ideais decidiu-se juntar esta informação devido à possível contribuição que podem ter para a caracterização das condições habitacionais em que a população vive.

3.2.2 Cálculo da Mortalidade Infantil

A mortalidade infantil foi calculada dividindo o número de óbitos abaixo de 1 ano de idade pelo número de nados vivos, para o mesmo ano, por distrito ou concelho conforme a década.

Como de 1930 a 1939 os dados de óbitos abaixo de 1 ano e nados-vivos encontravam-se agrupados por distrito, agruparam-se também nesta unidade geográfica os dados de

mortalidade infantil das décadas de 40 e 50 e calculou-se, para todos os anos, a mortalidade infantil. Calculou-se ainda a mortalidade infantil média, por década, paracada distrito

3.2.3 Análise

A análise dos dados pode ser dividida em dois grandes grupos: a análise dos dados populacionais com os quais se procurou caracterizar socioeconomicamente os concelhos e relacionar o resultado encontrado com a taxa acumulada de incidência das fracturas do colo do fémur; e a análise dos dados da mortalidade infantil, relacionando-se com a incidência das fracturas e também com cada uma das componentes principais

1. Análise dos dados demográficos:

Após uma primeira selecção de variáveis, anteriormente descrita, a primeira etapa desta análise consistiu na Análise de Componentes Principais, de forma a obter-se um pequeno número de combinações lineares (componentes principais) do conjunto de variáveis original, que retenham o máximo possível da informação contida nessas variáveis.

Desta forma começou por ser analisada a matriz de correlações o que permitiu identificar as variáveis que estavam mais correlacionadas entre si e que, conseqüentemente, deveriam ser excluídas. A análise desta matriz, permitiu identificar desde logo aquelas variáveis que possuíam o coeficiente de correlação mais baixo entre si, permitindo refinar a escolha final do conjunto de variáveis pré-seleccionadas e avaliar, dessa forma, a pertinência na especificação do uso do modelo de componentes principais. Foram eliminadas da análise as variáveis que apresentavam coeficientes de correlação muito altos, de tal maneira que do conjunto inicial de 88 variáveis retiveram-se apenas 30, não correlacionadas entre si. Uma série de simulações foi realizada até se chegar a este conjunto ideal de variáveis.

Desta forma obteve-se o seguinte conjunto final de variáveis que foi analisado:

Tabela 3.2 Tabela das variáveis escolhidas para realizar a análise.

Variáveis	Variáveis
P_h65	Cnp_45
P_m65	Cnp_6_9
P_vh	Iperc
P_vm	P_res
P_ref_hm	P_res_d
Tx_a01	P_crpa
Idep_i	P_cfa
Idep_j	P_scfb
P_div	P_scp
P_sd_hm	P_scot
P_1c	P_ibd
P_3c	P_ac
P_sc_c	P_ancl
P_sup_c	P_ancm
Cnp_123	P_hm50_

A etapa seguinte foi a extracção das componentes principais. Para validar esta análise foram realizados os testes KMO e de Bartlett para identificar se o conjunto de variáveis seleccionadas é adequado para a utilização da Análise de Componentes Principais. O resultado deste teste permitiu considerar a qualidade das correlações entre as variáveis de boa para se prosseguir com a análise, já que o teste KMO teve um resultado superior a 0,80 e o teste de esfericidade de Bartlett foi significativo. O resultado deste teste pode ser visto na **Tabela 3.3**.

Tabela 3.3 Resultado do teste de Bartlett e KMO.

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		0,856
Bartlett's Test of Sphericity	Approx. Chi-Square	13577,429
	df	435
	Sig.	,000

A selecção do número de componentes principais a reter foi baseada em dois métodos: no método de Kaiser e no método em que se retêm tantas componentes quantas as necessárias para se atingir um determinado total de variância desejado. Neste caso o total de variância pretendido estava acima dos 70%. No primeiro método, Kaiser sugere que a melhor resposta para a questão do número de componentes, quando se utiliza a matriz de correlação, será dada pelo número de componentes cujos valores próprios são maiores ou iguais a 1.

Também foi feito e analisado o *scree plot*. O resultado deste método sugeria a retenção de cinco componentes principais, mas a decisão de se usar apenas as quatro componentes foi tomada depois de se verificar que a quinta componente não diferenciava quase nenhuma variável e o nível de explicação da variância total só aumentava 3%.

Através da análise da **Tabela 3.4** e com base nos métodos referidos, foram extraídas 4 componentes principais (valor próprio igual a 1,611) e, que explicam, aproximadamente, 75,778% da variância total.

Tabela 3.4 Total da variância explicada.

Componente	Valores Próprios		
	Total	% de Variância	% acumulada
1	11,659	38,862	38,862
2	6,558	21,862	60,724
3	2,906	9,685	70,409
4	1,611	5,369	75,778
5	,941	3,135	78,913
6	,893	2,978	81,891
7	,817	2,724	84,615
8	,635	2,116	86,732
9	,593	1,978	88,709
10	,528	1,758	90,468
11	,484	1,614	92,082
12	,474	1,580	93,662
13	,365	1,218	94,880
14	,322	1,074	95,953
15	,222	,741	96,695
16	,201	,671	97,366
17	,184	,613	97,979
18	,114	,380	98,359
19	,104	,345	98,704
20	,089	,298	99,002
21	,067	,222	99,224
22	,056	,185	99,409
23	,040	,134	99,543
24	,036	,119	99,663
25	,032	,106	99,769
26	,027	,091	99,860
27	,025	,082	99,942
28	,013	,044	99,986
29	,003	,010	99,995
30	,001	,005	100,000

Através da análise da tabela de correlações entre as componentes (**Tabela 3.5**), obtida quando se fazia a rotação oblíqua, decidiu-se pelo uso das rotações ortogonais. Com base nesta tabela realizou-se então a rotação ortogonal *varimax*, uma vez que as correlações são baixas (apesar da correlação das componentes 1 e 4 ser ligeiramente superior a 0.3). Na **Tabela 3.6** é apresentada a matriz das componentes rotacionadas.

Tabela 3.5 Total de correlações das componentes principais.

Componente	1	2	3	4
1	1,000	0,161	0,237	0,331
2	0,161	1,000	0,105	0,051
3	0,237	0,105	1,000	-0,040
4	0,331	0,051	-0,040	1,000

Tabela 3.6 Matriz das componentes rotacionada.

	Componentes			
	1	2	3	4
P_h65	0,765			0,330
P_m65	0,792			
P_vh	0,821	-0,421		
P_vm	0,895	-0,338		
P_ref_hm	0,871	-0,389		
Tx_a01	0,702	-0,513		0,327
Idep_i	0,876	-0,340		
Idep_j	-0,811			
P_div		-0,386	-0,706	-0,359
P_sd_hm			0,669	
P_1c		-0,799		
P_3c		0,644	0,438	
P_sc_c		0,846	0,380	
P_sup_c		0,906		
Cnp_123		0,907		
Cnp_45		0,792	0,332	
Cnp_6_9		-0,955		
Iperc		0,910		
P_res	-0,953			
P_res_d	-0,853		0,372	
P_crpa	-0,684	-0,331		
P_cfa				0,544
P_scfb			-0,352	0,669
P_scp				0,622
P_scot				0,687
P_ibd		0,524		-0,674
P_ac			-0,721	
P_ancl	0,307	-0,515	-0,563	
P_ancm		0,520	0,686	
P_hm50_	0,941			

Após a sua extracção, estas componentes principais guardaram-se como novas variáveis.

Concluída esta etapa, passou-se à análise de *clusters*, de maneira a poder agrupar os diversos municípios em áreas homogéneas de acordo com os factores socioeconómicos. Para isso foram identificados os *clusters* de concelhos homogéneos no que se refere às componentes principais encontradas.

O método de aglomeração escolhido foi a aglomeração hierárquica, tendo sido o *Ward's method* a regra de formação de *clusters* escolhida. A determinação do número de *clusters* fez-se com base no conhecimento empírico. Foram assim determinados 6 *clusters*, número considerado bastante e suficiente para diferenciar as regiões, sem excesso de generalização por um lado, nem excessivas desagregações por outro. Procedeu-se em seguida à interpretação e mapeamento destes *clusters*.

Após esta análise, que permitiu caracterizar socioeconomicamente os concelhos, foram relacionados os resultados encontrados com a incidência das fracturas do colo do fémur. Assim, através da agregação dos concelhos pertencentes a cada *cluster*, sumariou-se o total de casos de fracturas osteoporóticas do colo do fémur por sexo e grupos etários de 5 anos e população residente também por sexo e grupo etário de 5 anos. Em seguida foram calculadas as taxas de incidência acumuladas, de 2000 a 2002, (e respectivos intervalos de confiança) das fracturas de colo do fémur, por sexo e padronizadas por idade pelo método directo, para cada um dos 6 *clusters* encontrados. Para verificar se os valores das taxas encontrados para cada *cluster* eram ou não estatisticamente diferentes foram analisados os intervalos de confiança (com 95% de confiança).

Para além desta análise realizou-se também a correlação linear entre cada componente principal encontrada e a incidência de fracturas, por concelho.

Uma variável, pelo seu especial interesse, foi também analisada isoladamente e correlacionada com as fracturas de colo do fémur, por concelho. Foi este o caso da percentagem de pessoas, (separadamente para o género masculino e feminino), com 65 anos ou mais que viviam sozinhas. A decisão de tratar esta variável de maneira distinta deveu-se ao facto de existirem estudos que demonstram a existência de uma associação entre o facto de se viver sozinho e o aumento do risco de sofrer uma fractura (Farahmand BY, Persson PG, 2000, Vestergaard P, Rejnmark L et al., 2006).

2. Análise dos dados de mortalidade infantil:

A análise destes dados consistiu no estudo da regressão entre a mortalidade infantil e a incidência de fracturas e entre a mortalidade infantil e as componentes principais. Como os dados da mortalidade infantil foram calculados por distrito e as taxas de incidência de fracturas de colo do fémur estavam calculadas por concelho, foi necessário recalcular as taxas de incidência acumuladas (2000 a 2002) das fracturas por distrito. Seguiu-se a mesma metodologia que tinha sido usada no cálculo destas taxas por concelho, ou seja, padronizou-se por idade, usando o método directo, a população portuguesa de 2001, como população padrão e grupos etários de 5 em 5 anos a partir dos 50 anos de idade.

Após este primeiro passo foi feita a regressão entre a mortalidade infantil por ano (de 1930 a 1950) e a taxa de incidência acumulada das fracturas do colo do fémur, para ambos os sexos. Foi ainda realizada a regressão entre a mortalidade infantil média de cada década (décadas de 30, 40 e 50) e a taxa de incidência acumulada das fracturas de colo do fémur.

Por fim e da mesma forma calculou-se a regressão da mortalidade infantil da década de 40 e 50 (por concelhos) com as componentes principais. Só foi possível fazer a regressão utilizando os dados destas duas décadas, porque os dados referentes à década de 30 estavam agrupados por distritos e não por concelhos.

4. RESULTADOS

4.1 Resultados da Análise dos Dados Demográficos

4.1.1 Análise de Componentes Principais

A base de dados de onde a análise partiu tinha um total de 88 variáveis do Censo Demográfico de 2001 (fornecido pelo INE), distribuídas da seguinte forma: 3 variáveis relacionadas com a estrutura familiar (3%); 33 variáveis relacionadas com a habitação (38%); e 52 variáveis relacionadas com a população (59%).

O primeiro passo desta análise, a construção da matriz de correlações permitiu reduzir este conjunto inicial de variáveis. A matriz de correlações pode ser vista no **Anexo B**.

Da análise das variáveis iniciais foram seleccionadas 30 variáveis: 2 dizem respeito à estrutura familiar (7%); 12 à habitação (40%) e 16 à população (53%). As variáveis seleccionadas para a análise estão apresentadas na **Tabela 4.1**, onde estão também apresentadas as respectivas médias das percentagens por concelho e desvios padrões.

Tabela 4.1 Tabela das variáveis escolhidas para realizar a análise, com as respectivas médias e desvios padrão.

Variável	Média (%)	Desvio Padrão (%)	Variável	Média (%)	Desvio Padrão (%)
P_h65	10,862	2,54	Cnp_45	22,441	4,845
P_m65	25,794	4,157	Cnp_6_9	57,922	9,735
P_vh	3,287	1,132	Iperc	70,464	27,302
P_vm	12,576	3,183	P_res	2,789	0,26
T_ref_hm	29,381	7,685	P_res_d	0,586	0,062
Tx_a01	13,521	5,543	P_crpa	15,957	18,983
Idep_i	33,831	12,911	P_cfa	0,813	0,698
Idep_j	22,583	3,193	P_scfb	1,617	2,450
P_div	4,775	0,328	P_scp	0,7	1,042
P_sd_hm	1,946	1,022	P_scot	0,258	0,217
P_1c	25,430	2,962	P_ibd	90,787	5,666
P_3c	4,244	0,85	P_ac	5,258	4,157
P_sc_c	5,173	2,134	P_ancl	48,852	24,169
P_sup_c	3,7	2,191	P_ancm	27,04	17,443
Cnp_123	18,879	5,85	P_hm50_	38,457	7,679

A análise da **Tabela 3.6** (matriz das componentes rotacionadas) permitiu compreender o conjunto de variáveis que cada componente melhor explica, ou seja, o conjunto de variáveis que têm mais peso na formação da componente, facilitando a interpretação destas. Assim conclui-se que:

- A componente 1 está mais associada a uma maior percentagem de indivíduos com mais de 50 anos, de indivíduos com 65 anos ou mais que vivem sozinhos, de viúvos, de indivíduos reformados e com uma maior taxa de analfabetismo. Está ainda associada a um maior índice de dependência de idosos e, consequentemente, com um menor índice de dependência de jovens, a um menor número médio de pessoas por alojamento e divisões e uma menor percentagem de alojamentos familiares com água canalizada proveniente de rede particular, ocupados como residência habitual. Ou seja, esta componente está relacionada com a população mais envelhecida, a viver mais sozinha e com maior percentagem de analfabetismo.

- A componente 2 está mais associada a um maior nível de escolaridade, uma maior percentagem de indivíduos com profissões científicas, técnicas, artísticas, directores e quadros superiores administrativos e pessoal administrativo (ou seja, indivíduos empregados nos grupos de profissões CNP 1, 2 ou 3) e um pouco menos relacionada com a percentagem de indivíduos que trabalham no comércio ou como vendedores, pessoal dos serviços de protecção e segurança, dos serviços pessoais e domésticos (ou seja, empregados nos grupos de profissões CNP 4 e 5), e relacionada negativamente com a percentagem de indivíduos que trabalham como agricultores, criadores de animais, trabalhadores agrícolas e florestais, pescadores e caçadores ou trabalhadores das indústrias extractiva e transformadora e condutores de máquinas fixas e de transporte; está também relacionada com um índice *per capita* mais alto.

Em suma, é uma componente que reflecte um estatuto socioeconómico alto.

- A componente 3 está mais relacionada com algumas variáveis que dizem respeito às condições dos alojamentos e ao subsídio de desemprego. Esta componente encontra-se associada a um menor número de divisões por alojamento e um maior número de pessoas por divisão; a uma menor percentagem de utilização de aquecimento central e maior percentagem de utilização de aquecimento móvel; e a uma maior percentagem de indivíduos que recebem o subsídio de desemprego. Pode-se considerar que esta é uma componente associada com estatuto socioeconómico baixo e com más condições de moradia.

- A componente 4 está mais relacionada com as variáveis referentes às condições dos alojamentos, como as canalizações. Esta componente encontra-se relacionada com uma maior percentagem de alojamentos familiares sem água canalizada, nem nos próprios alojamentos nem nos edifícios e com uma maior percentagem de alojamentos sem instalações de banho e duche, ocupados como residência habitual. Em suma, esta componente está mais associada com más condições dos alojamentos.

As próximas tabelas (**Tabelas 4.2 a 4.4**) apresentam o resultado das regressões realizadas entre cada uma das componentes principais, por concelho, e a taxa de incidência das fracturas osteoporóticas do colo do fémur, também por concelho.

Tabela 4.2 Correlações e significâncias resultantes das regressões entre as componentes principais e a taxa de incidência de fracturas osteoporóticas do colo do fémur para homens.

Componente	Correlação (r)	Significância
1	0,014	0,813
2	0,059	0,324
3	0,058	0,335
4	0,055	0,358

Tabela 4.3 Correlações e significâncias resultantes das regressões entre as componentes principais e a taxa de incidência de fracturas osteoporóticas do colo do fémur para mulheres.

Componente	Correlação (r)	Significância
1	-0,007	0,906
2	0,264	0,000
3	-0,1	0,866
4	0,02	0,741

Tabela 4.4 Correlações e significâncias resultantes das regressões entre as componentes principais e a taxa de incidência de fracturas osteoporóticas do colo do fémur para o total de homens e mulheres.

Componente	Correlação (r)	Significância
1	0,002	0,702
2	0,245	0,000
3	0,012	0,837
4	0,002	0,971

Também foi estudada a regressão entre as variáveis referentes às percentagens de homens e mulheres com 65 anos ou mais que vivem sozinhos, por concelho e a taxa de incidência destas fracturas, também por concelho, para cada um dos sexos em separado. Na **Tabela 4.5** estão apresentados os resultados encontrados.

Tabela 4.5 Correlações e significâncias resultantes das regressões entre a variável referente à percentagem de homens e mulheres, com 65 anos ou mais que vivem sozinhos e a taxa de incidência de fracturas osteoporóticas do colo do fémur, para homens e mulheres.

Variável	Correlação	Significância
P_H65	0,012	0,846
P_M65	0,018	0,77

4.1.2 Análise de *Clusters*

Uma vez obtidas as componentes principais, estas passaram a ser tratadas como variáveis na análise de *clusters*, ou seja, as variáveis utilizadas na análise de *clusters* foram as componentes principais. Na **Tabela 4.6** é apresentada a distribuição da população, dos concelhos e do número de fracturas osteoporóticas do colo do fémur, em cada um dos seis *clusters* e na **Tabela 4.7** é possível perceber-se como cada componente se comporta em cada *cluster*, através das estatísticas descritivas.

Tabela 4.6 Distribuição da população, dos concelhos e das fracturas osteoporóticas do colo do fémur por *cluster*.

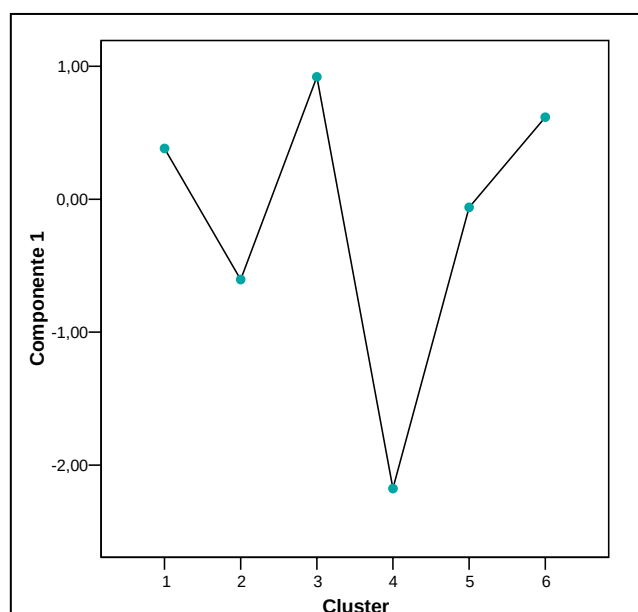
Cluster	População total (%)	Total de concelhos	Total de fracturas (%)
1	430020 (4%)	25	1489 (6%)
2	2207783 (22%)	71	5484 (21%)
3	616478 (6%)	64	2624 (10%)
4	1251059 (13%)	17	1805 (7%)
5	5154564 (52%)	75	13497 (53%)
6	209439 (2%)	26	735 (3%)

Tabela 4.7 Descrição de cada *cluster* com base nas componentes principais.

		Média	Desvio Padrão
Comp. 1	1	0,382	0,926
	2	-0,605	0,63
	3	0,921	0,691
	4	-2,177	0,384
	5	-0,061	0,504
	6	0,617	0,333
Comp. 2	1	0,108	0,723
	2	-0,054	0,769
	3	-0,666	0,341
	4	-,96	0,547
	5	1,04	0,97
	6	-0,69	0,365
		Média	Desvio Padrão
Comp. 3	1	0,208	0,85
	2	-0,935	0,5
	3	-0,552	0,576
	4	0,691	0,427
	5	0,607	0,675
	6	1,509	0,531
Comp. 4	1	2,451	1,08
	2	-0,147	0,629
	3	-0,38	0,542
	4	-0,189	0,567
	5	-0,217	0,486
	6	-0,272	0,725

No quadro de aglomeração (ver **Anexo C**) pode-se identificar quais os casos ou *clusters* que são combinados em cada etapa. O dendograma (ver **Anexo D**) representa a aglomeração feita numa escala 0 a 25, que é diferente da escala das distâncias e serve como um dos processos de escolha do número adequado de *clusters*.

Para a interpretação dos *clusters*, ou seja para identificar as variáveis (neste caso componentes) que se distinguem entre os *clusters*, recorreu-se ao *One-Way Anova* e à análise da **Tabela 4.7** e das figuras seguintes (**Figuras 4.1 a 4.4**). Estas figuras representam para cada *cluster* o valor médio de cada componente.

**Figura 4.1** Representação das médias da componente 1 para cada *cluster*.

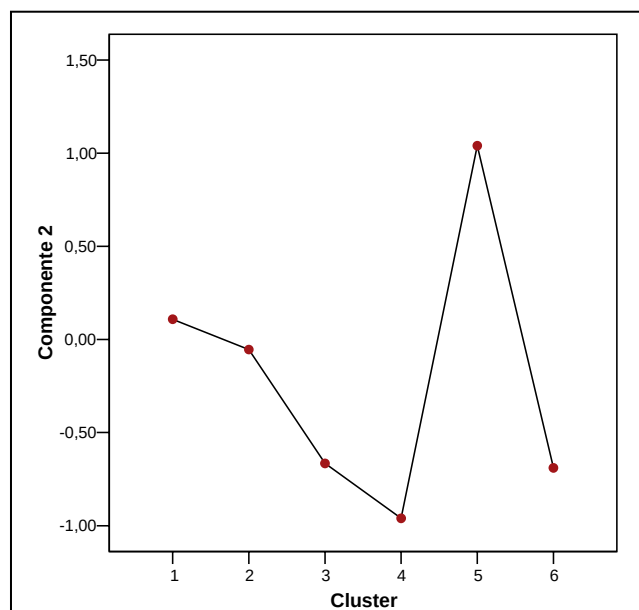


Figura 4.2 Representação das médias da componente 2 para cada *cluster*.

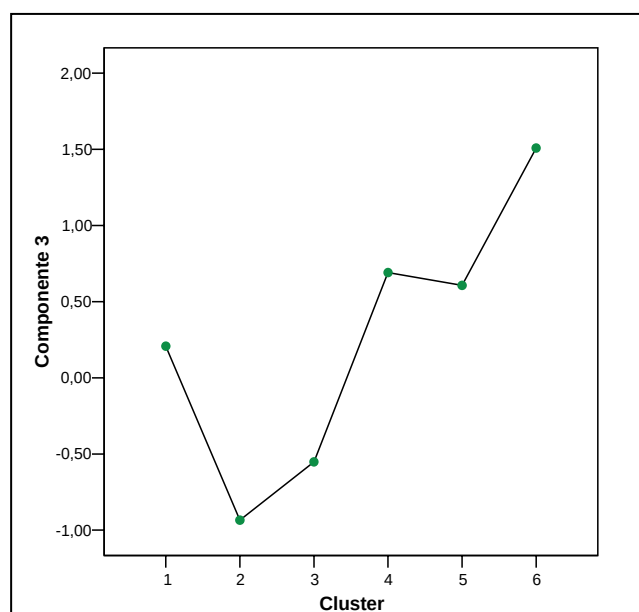


Figura 4.3 Representação das médias da componente 3 para cada *cluster*.

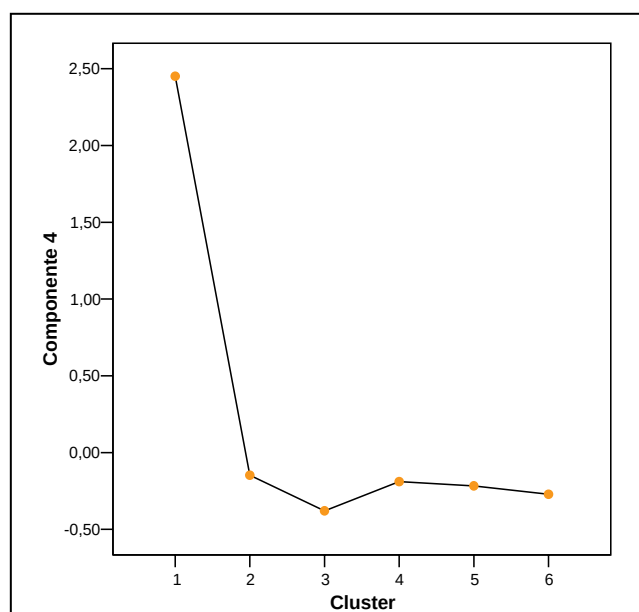


Figura 4.4 Representação das médias da componente 4 para cada *cluster*.

Através da análise da **Tabela 4.7** e das **Figuras 4.1 a 4.4** pode-se verificar que:

- O cluster 5 está associado ao estatuto socioeconómico mais alto: agrupa os concelhos com população menos envelhecida, baixa taxa de analfabetismo e menor número de pessoas a viver sozinhas; está ainda relacionado com alto nível de escolaridade, alto índice *per capita* e maior percentagem de indivíduos empregados nos grupos de profissões CNP 1-3 (Quadro superior, científico e pessoal administrativo) ou 4-5 (comércio e serviços); com um menor número de divisões por alojamento, um maior número de pessoas por divisão e com uma maior percentagem de indivíduos que recebem o subsídio de desemprego. Este *cluster* encontra-se relacionado com os maiores centros urbanos, como o Porto e os arredores mais próximos (como Matosinhos, Rio Tinto e Vila Nova de Gaia), Lisboa e arredores (que incluem alguns concelhos de Setúbal e Santarém), os concelhos da costa do Algarve, Coimbra, Évora e Beja. Mas deste *cluster* também fazem parte concelhos como Figueira da Foz, Castelo Branco, Covilhã, Portalegre, Castelo de Vide, Campo Maior, Elvas, Reguengos de Monsaraz e Castro Verde.

- O cluster 4 está associado ao estatuto socioeconómico médio - alto: agrupa os concelhos com a população mais jovem e menor taxa de analfabetismo, no entanto, está relacionado também com os níveis mais baixos de escolaridade (como o 1º ciclo), com

baixo índice *per capita*, grande percentagem de indivíduos empregados nos grupos de profissões CNP 6 a 9 (que incluem empregados na indústria) e com boas condições a nível dos alojamentos (canalização, aquecimento e instalações de banho e duche). A este *cluster* pertencem os concelhos dos arredores do Porto, como Braga, Guimarães, Paredes, Penafiel, Marco de Canaveses, Santa Maria de Feira e Ovar, portanto, áreas industriais, medianamente urbanas. Os resultados indicam que, em princípio, estas populações são jovens, com um baixo nível de escolaridade, com empregos ligados à própria indústria, e com boas condições a nível de alojamentos.

- O cluster 2 está associado ao estatuto socioeconómico médio: agrupa concelhos com boas condições de alojamentos, mas uma população um pouco menos jovem do que o *cluster* anterior e com uma maior taxa de analfabetismo. Está relacionado com um baixo número de pessoas a viverem sozinhas e com um nível baixo de escolaridade e de índice *per capita*. Este *cluster* está mais relacionado com altas percentagens de indivíduos empregados nos grupos de profissões CNP 6 a 9 (além da indústria, inclui pescadores e agricultores). Este *cluster* agrega concelhos que estão em regiões pesqueiras e agrícolas, maioritariamente no litoral centro-norte, mas também inclui Bragança, Macedo de Cavaleiros, Mirandela, Chaves e Guarda. Estes últimos talvez não tenham populações tão envelhecidas como os que os concelhos que os rodeiam, e por isso se agrupem a este *cluster*.

- O cluster 1 está associado ao estatuto socioeconómico médio - baixo: agrega concelhos com a população mais envelhecida, com uma maior taxa de analfabetismo e maior número de pessoas que vivem sozinhas; com um nível baixo de escolaridade, baixo índice *per capita* e pouca actividade ligadas às profissões mais científicas, técnicas ou administrativas (CNP 1,2 e 3), mais actividade na área agrícola, florestal e trabalhadores de indústrias (CNP 6 a 9); e com as piores condições a nível de alojamentos (alojamentos sem água canalizada nem nos alojamentos, nem nos edifícios e com uma maior percentagem de alojamentos sem instalações de banho e duche). Deste *cluster* fazem parte os concelhos do baixo Alentejo e as regiões interiores do Algarve, mas também Coruche, Montemor-o-Novo, Estremoz, Arronches, Marvão, Arcos de Valdevez, Baião e Celorico de Bastos.

- O cluster 6 está associado a um estatuto socioeconómico baixo: agrega concelhos do alto Alentejo mas também Chamusca, Alpiarça, Salvaterra de Magos e Cadaval. Os resultados mostram que estas são populações envelhecidas, com altas taxas de analfabetismo, com um nível médio-baixo de escolaridade. É grande a percentagem de pessoas que vivem sozinhas neste *cluster*, o índice *per capita* é baixo e a pessoas têm uma actividade mais rural (CNP 6 a 9). Estes concelhos têm um menor número de divisões por alojamento e o maior número de pessoas por divisão; têm ainda uma menor percentagem de utilização de aquecimento central e maior percentagem de utilização de aquecimento móvel. São ainda populações com a maior percentagem de indivíduos a receber o subsídio de desemprego.

- O cluster 3 está associado a um estatuto socioeconómico muito baixo: agrega concelhos com a população mais envelhecida, maior taxa de analfabetismo e maior número de pessoas a viverem sozinhas; com um baixo nível de escolaridade, baixo índice *per capita* e actividades mais rurais (CNP 6 a 9); com a maior percentagem de alojamentos sem água canalizada e sem instalações de banho e duche. Deste *cluster* fazem parte os concelhos do interior centro - norte. Este foi o *cluster* considerado mais desfavorecido nesta análise.

Na figura seguinte (**Figura 4.5**) é apresentado o resultado do mapeamento dos 6 *clusters*.

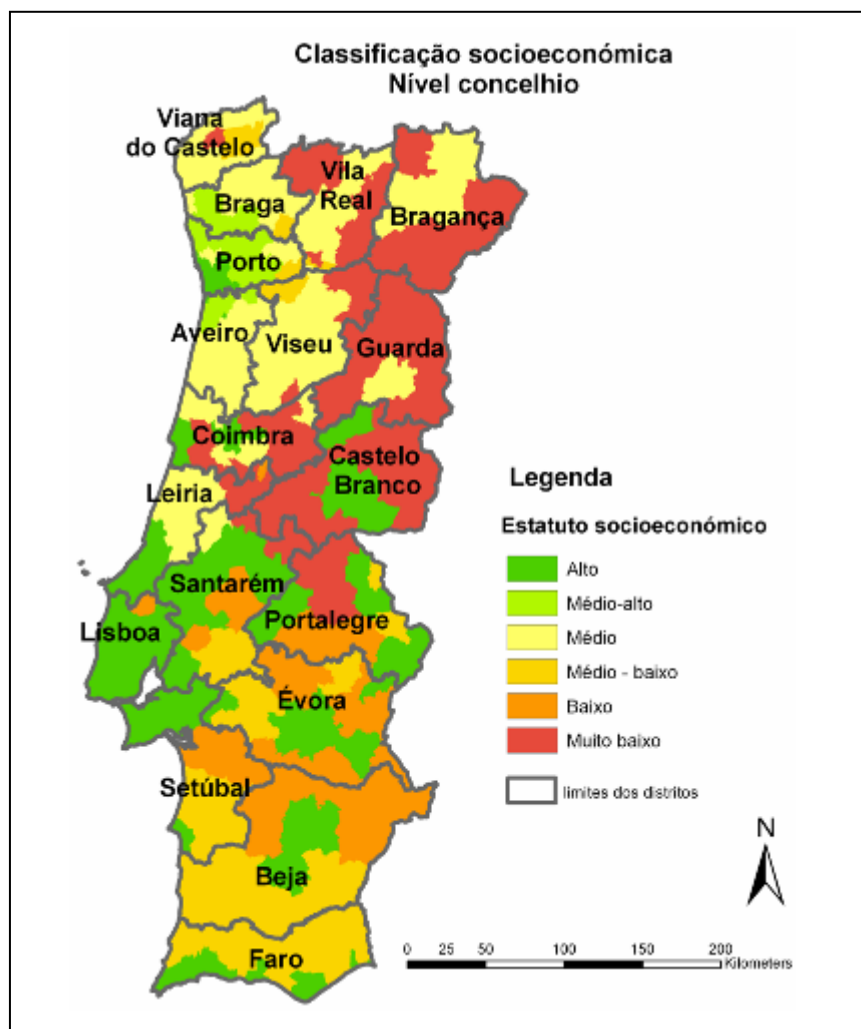


Figura 4.5 Mapa dos *clusters* da análise dos dados demográficos.

As tabelas e figuras seguintes (Tabela 4.8 a 4.10 e Figuras 4.6 a 4.8) apresentam as taxas de incidência das fracturas osteoporóticas do colo do fémur, por género e total, padronizadas por idade, calculadas para cada cada *cluster* com os respectivos intervalos de confiança e o risco relativo (RR) em relação ao *cluster* com a menor taxa de incidência de fracturas.

Tabela 4.8 Taxa de incidência das fracturas osteoporóticas do colo do fémur, para homens.

<i>Cluster</i>	Taxa de incidência (IC95%)	RR
Estatuto Alto	400,104 (385,656 - 414,552)	1,10
Estatuto Médio - Alto	364,613 (330,582 - 398,644)	1 (ref)
Estatuto Médio	380,174 (359,419 - 400,928)	1,04
Estatuto Médio - Baixo	389,059 (350,095 - 428,024)	1,07
Estatuto Baixo	382,378 (328,149 - 436,607)	1,05
Estatuto Muito Baixo	376,132 (346,643 - 405,621)	1,03

Tabela 4.9 Taxa de incidência das fracturas osteoporóticas do colo do fémur, para mulheres.

<i>Cluster</i>	Taxa de incidência (IC95%)	RR
Estatuto Alto	1108,468 (1087,3 - 1129,6)	1,20
Estatuto Médio - Alto	927,581 (878,35 - 976,81)	1 (ref)
Estatuto Médio	1012,182 (981,55 - 1042,8)	1,09
Estatuto Médio - Baixo	1057,53 (995,18 - 1119,8)	1,14
Estatuto Baixo	953,672 (873,53 - 1033,8)	1,03
Estatuto Muito Baixo	1017,193 (972,6 - 1061,8)	1,10

Tabela 4.10 Taxa de incidência das fracturas osteoporóticas do colo do fémur, para ambos os sexos.

<i>Cluster</i>	Taxa de incidência (IC95%)	RR
Estatuto Alto	798,795 (785,32 - 812,27)	1,18
Estatuto Médio - Alto	678,179 (646,89 - 709,47)	1 (ref)
Estatuto Médio	727,809 (708,55 - 747,07)	1,07
Estatuto Médio - Baixo	736,2 (698,81 - 773,59)	1,09
Estatuto Baixo	688,668 (638,88 - 738,46)	1,02
Estatuto Muito Baixo	719,743 (692,2 - 747,28)	1,06

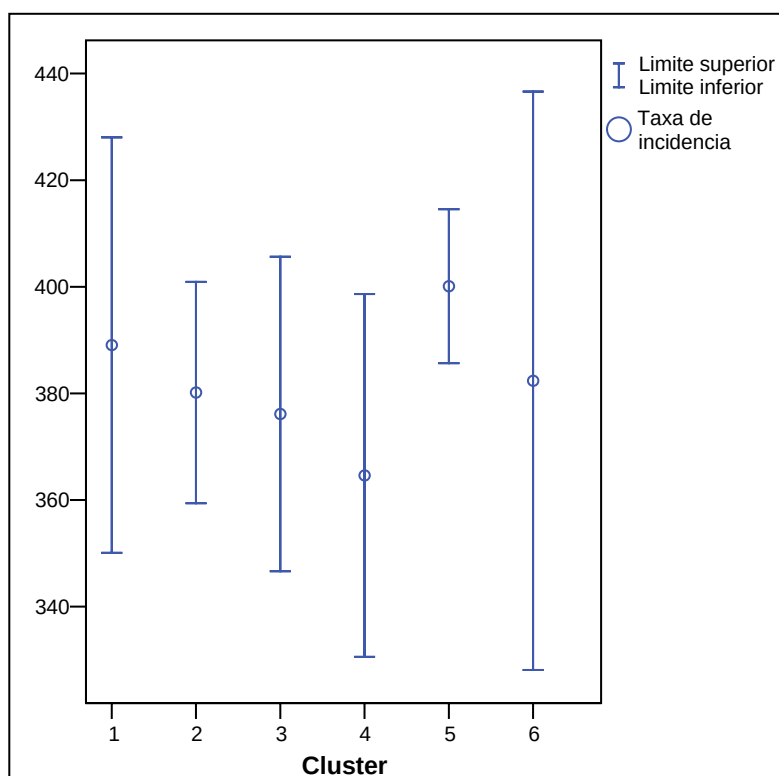


Figura 4.6 Representação das taxas de incidência das fracturas osteoporóticas do colo do fémur para homens e intervalos de confiança para cada *cluster*.

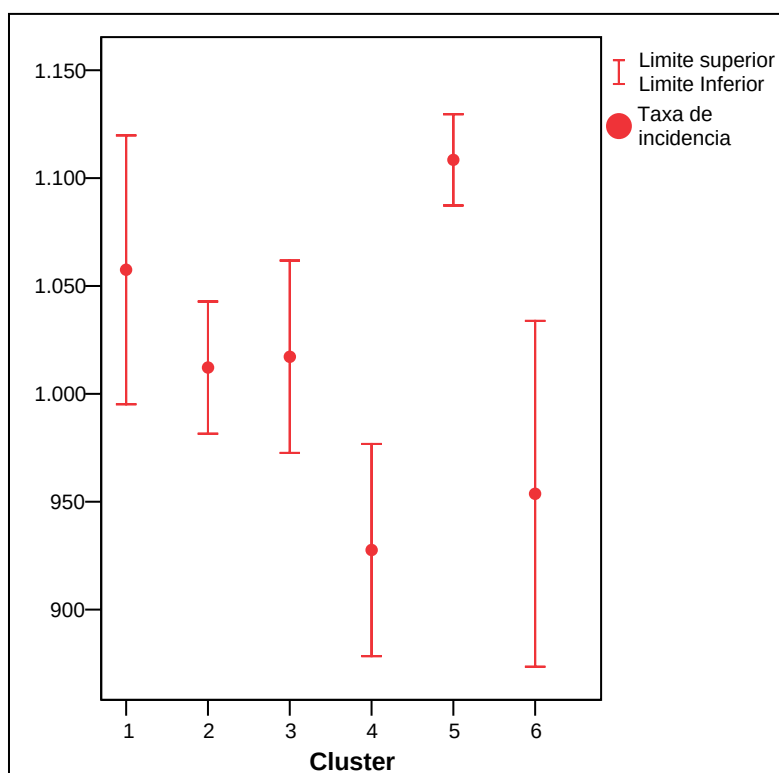


Figura 4.7 Representação das taxas de incidência das fracturas osteoporóticas do colo do fémur para mulheres e intervalos de confiança para cada *cluster*.

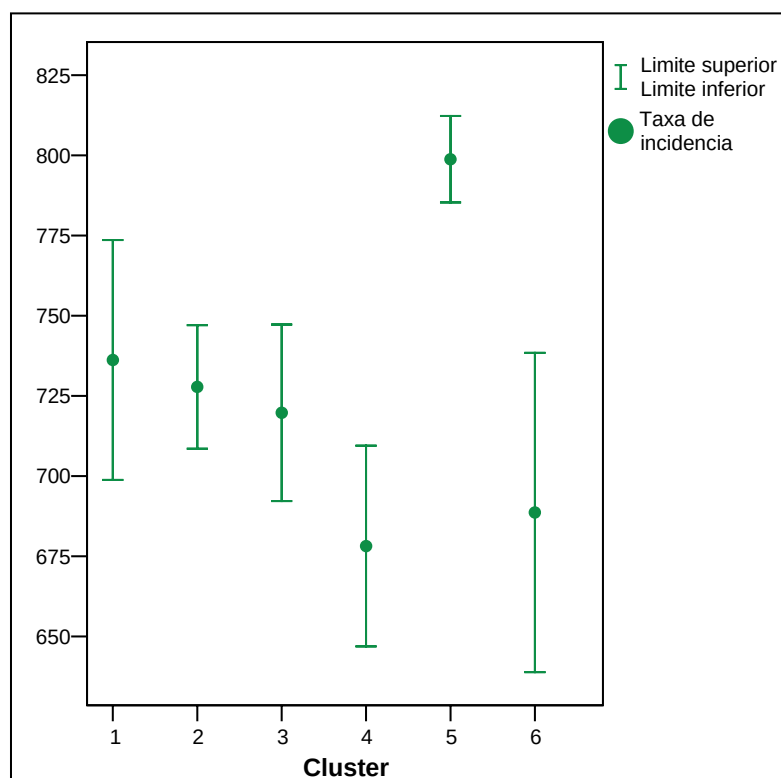


Figura 4.8 Representação das taxas de incidência das fracturas osteoporóticas do colo do fémur para ambos os sexos e intervalos de confiança para cada *cluster*.

4.2 Resultados da Análise dos Dados de Mortalidade Infantil

Durante as três décadas a que os dados da mortalidade infantil se referem (1930 a 1959), segundo os dados do INE, ocorreram 617.572 óbitos infantis (menos de 1 ano de idade) e um total de 5.576.504 nados vivos. As percentagens de óbitos por década, em relação ao total de óbitos no período completo, foram as seguintes: 38%, 35% e 27%, respectivamente para os anos 30 (ou seja, entre 1930 e 1939), anos 40 (ou seja, entre 1940 e 1949) e anos 50 (entre 1950 e 1959) (ver **Tabela 4.11**).

Tabela 4.11 Distribuição dos óbitos infantis e de nados vivos, nas três décadas.

Década		Total de óbitos	Mortalidade infantil (por mil nados vivos)
1930 - 1939	Óbitos infantis	235204	125,699
	Nados vivos	1871168	
1940 - 1949	Óbitos infantis	217350	117,7224
	Nados vivos	1846292	
1950 - 1959	Óbitos infantis	165018	88,76498
	Nados vivos	1859044	

A mortalidade infantil foi calculada ano a ano para as três décadas, no entanto a unidade geográfica de análise para os anos 30 foi apenas distrito, uma vez que os dados do INE estavam agregados desta maneira, enquanto que para os anos 40 e 50 os dados estavam mais desagregados, por concelhos. Sendo assim, para as décadas de 40 e 50 a mortalidade infantil foi calculada tanto por distrito como por concelho. Na **Tabela 4.12** estão apresentados os valores médios da mortalidade infantil, por cada década, por distrito.

Tabela 4.12 Mortalidade infantil em cada década, por distrito (por mil nados vivos).

Distritos	Mortalidade infantil		
	1930 - 39	1940-49	1950-59
Aveiro	101,957	104,52	78,904
Beja	122,339	114,5296	80,537
Braga	125,71	131,958	107,67
Braganca	140,561	131,37	96,954
Castelo Branco	111,838	106,421	84,962
Coimbra	91,5675	82,43	78,753

Evora	123,077	116,091	81,877
Faro	114,742	100,165	83,232
Guarda	133,177	136,819	85,568
Leiria	102,753	86,961	81,661
Lisboa	160,164	122,627	65,544
Portalegre	114,209	111,613	72,27
Porto	159,69	152,359	125,145
Santarem	99,488	87,342	83,381
Setubal	153,738	130,136	90,473
Viana Do Castelo	103,941	103,907	103,232
Vila Real	117,475	118,982	96,73
Viseu	99,8546	102,471	83,556

Na **Tabela 4.13** estão apresentados os resultados da regressão entre a mortalidade infantil em cada ano, por distrito e a taxa de incidência das fracturas do colo do fémur, no período de 2000 a 2002, também por distrito.

Tabela 4.13 Correlações e significâncias resultantes das regressões entre a mortalidade infantil por ano e a taxa de incidência das fracturas do colo do fémur.

Ano	Correlação	Significância	Ano	Correlação	Significância	Ano	Correlação	Significância
1930	0,270	0,279	1940	0,216	0,388	1950	0,579	0,012
1931	0,394	0,106	1941	-0,013	0,96	1951	-0,308	0,213
1932	0,269	0,28	1942	0,219	0,383	1952	-0,141	0,577
1933	0,286	0,25	1943	-0,026	0,939	1953	-0,208	0,408
1934	0,218	0,385	1944	-0,087	0,732	1954	-0,283	0,256
1935	0,349	0,155	1945	0,291	0,241	1955	-0,386	0,113
1936	0,360	0,142	1946	-0,045	0,86	1956	0,130	0,608
1937	0,275	0,269	1947	-0,193	0,444	1957	-0,193	0,444
1938	0,237	0,343	1948	0,055	0,828	1958	-0,167	0,509
1939	0,139	0,582	1949	0,003	0,99	1959	-0,179	0,478

A **Figura 4.9** representa graficamente as mesmas informações contidas na **Tabela 4.13**. Os pontos verdes representam a década de 50, os pontos azuis representam a década de 40 e os pontos vermelhos a década de 30. Ainda que os resultados não sejam significativos como se pode verificar na **Tabela 4.13**, é possível observar que os anos da década de 30 apresentam as maiores correlações e os anos da década de 50 as correlações mais baixas, à excepção do ano de 1950 que foge a este padrão.

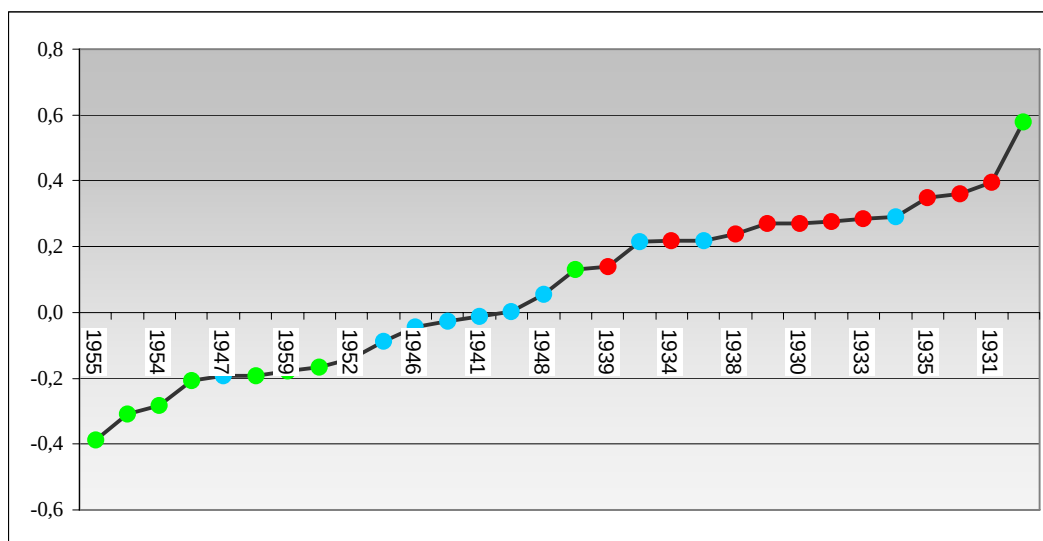


Figura 4.9 Representação das correlações entre a mortalidade infantil em cada ano, por distrito, e a taxa de incidência das fracturas do colo do fémur, também por distrito, ao longo dos anos.

As tabelas seguintes (**Tabelas 4.14 a 4.17**) mostram os resultados, em termos de correlações e significâncias, das regressões entre a mortalidade infantil, por ano nas décadas de 40 e 50 e cada componente principal (a amarelo estão assinalados os resultados mais significativos). Nas **Figuras 4.10 a 4.13** são apresentados os gráficos das correlações entre a mortalidade infantil em cada ano, por distrito e a cada componente principal (mais uma vez, os pontos verdes representam a década de 50 e os pontos azuis representam a década de 40). Não foram realizadas estas análises para a década de 30, porque as componentes foram calculadas por concelho, enquanto que a mortalidade infantil na década de 30 está por distrito.

Tabela 4.14 Correlações e significâncias resultantes das regressões entre a mortalidade infantil por ano nas décadas de 40 e 50, e a componente principal 1.

Ano	Correlação	Significância	Ano	Correlação	Significância
1940	0,020	0,745	1950	-0,043	0,484
1941	-0,096	0,115	1951	-0,840	0,166
1942	-0,020	0,739	1952	-0,054	0,372
1943	-0,092	0,131	1953	-0,152	0,012
1944	-0,027	0,656	1954	-0,056	0,355
1945	0,037	0,543	1955	-0,073	0,233
1946	0,015	0,812	1956	-0,096	0,113
1947	-0,06	0,325	1957	-0,089	0,144
1948	-0,118	0,053	1958	-0,102	0,093
1949	-0,069	0,256	1959	-0,058	0,337

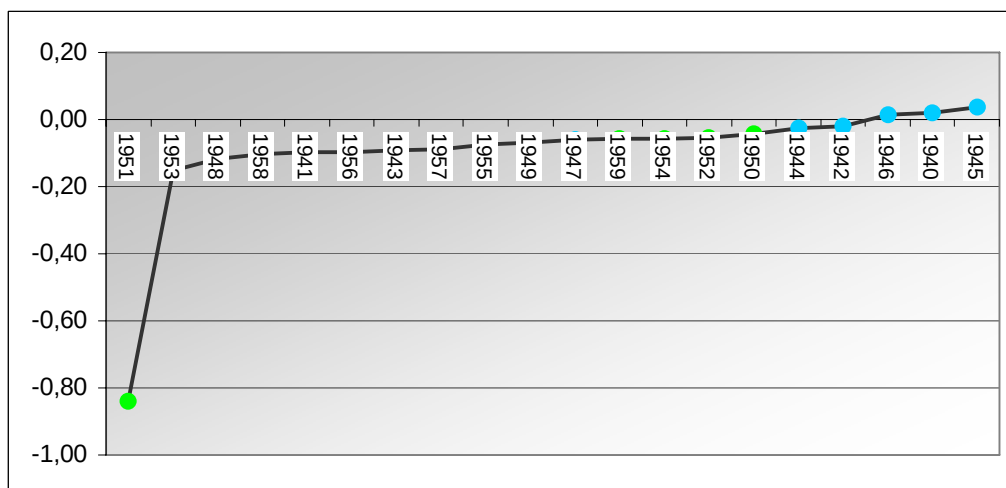


Figura 4.10 Representação das correlações entre a mortalidade infantil por ano nas décadas de 40 e 50, e a componente principal 1.

Tabela 4.15 Correlações e significâncias resultantes das regressões entre a mortalidade infantil por ano nas décadas de 40 e 50, e a componente principal 2.

Ano	Correlação	Significância	Ano	Correlação	Significância
1940	0,281	0,000	1950	0,211	0,000
1941	0,172	0,004	1951	-0,045	0,464
1942	0,114	0,061	1952	0,099	0,104
1943	-0,056	0,356	1953	-0,080	0,187
1944	0,140	0,021	1954	0,064	0,294
1945	0,159	0,009	1955	0,042	0,492
1946	-0,052	0,390	1956	0,093	0,128
1947	0,076	0,211	1957	0,041	0,496
1948	0,060	0,321	1958	0,026	0,672
1949	-0,008	0,893	1959	-0,032	0,598

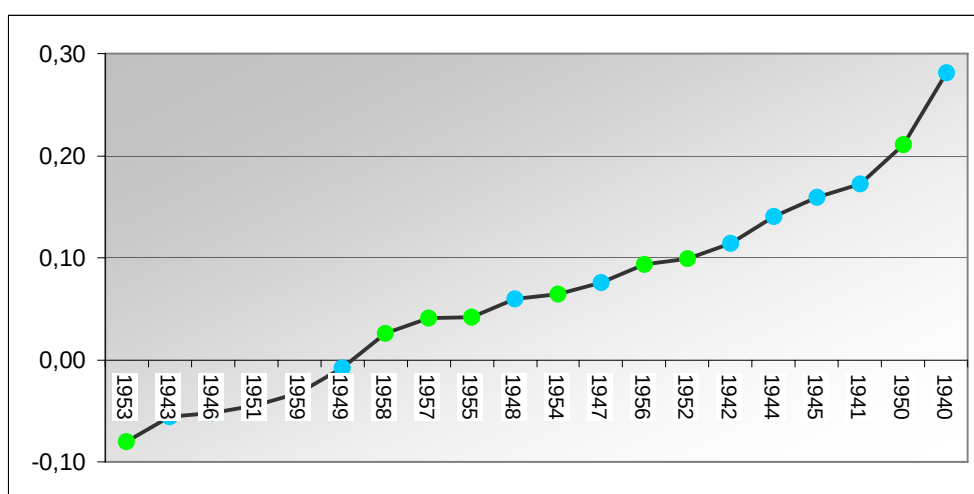


Figura 4.11 Representação das correlações entre a mortalidade infantil por ano nas décadas de 40 e 50, e a componente principal 2.

Tabela 4.16 Correlações e significâncias resultantes das regressões entre a mortalidade infantil por ano nas décadas de 40 e 50, e a componente principal 3.

Ano	Correlação	Significância	Ano	Correlação	Significância
1940	0,248	0,000	1950	0,121	0,046
1941	0,225	0,000	1951	0,143	0,018
1942	0,067	0,268	1952	-0,139	0,021
1943	-0,114	0,061	1953	0,049	0,423
1944	0,160	0,008	1954	0,235	0,000
1945	0,266	0,000	1955	-0,049	0,418
1946	0,141	0,020	1956	0,174	0,004
1947	0,088	0,150	1957	-0,074	0,226
1948	0,103	0,089	1958	-0,035	0,565
1949	0,078	0,197	1959	0,021	0,735

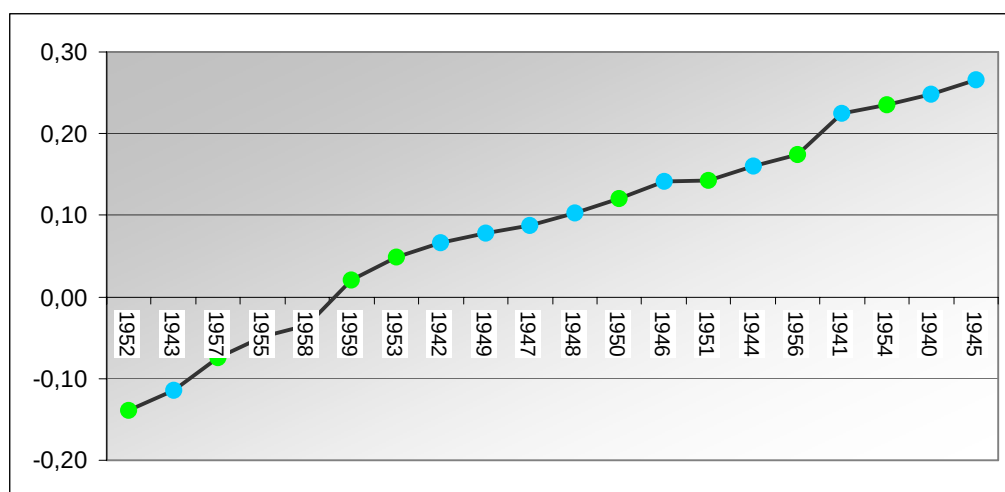


Figura 4.12 Representação das correlações entre a mortalidade infantil por ano nas décadas de 40 e 50, e a componente principal 3.

Tabela 4.17 Correlações e significâncias resultantes das regressões entre a mortalidade infantil por ano nas décadas de 40 e 50, e a componente principal 4.

Ano	Correlação	Significância	Ano	Correlação	Significância
1940	-0,180	0,003	1950	-0,043	0,477
1941	-0,080	0,191	1951	0,005	0,932
1942	0,003	0,962	1952	0,006	0,922
1943	-0,098	0,106	1953	-0,046	0,445
1944	-0,162	0,008	1954	-0,109	0,073
1945	0,023	0,704	1955	-0,041	0,496
1946	-0,109	0,072	1956	-0,045	0,456
1947	-0,131	0,030	1957	-0,135	0,026
1948	-0,082	0,177	1958	-0,068	0,261
1949	-0,086	0,156	1959	-0,083	0,171

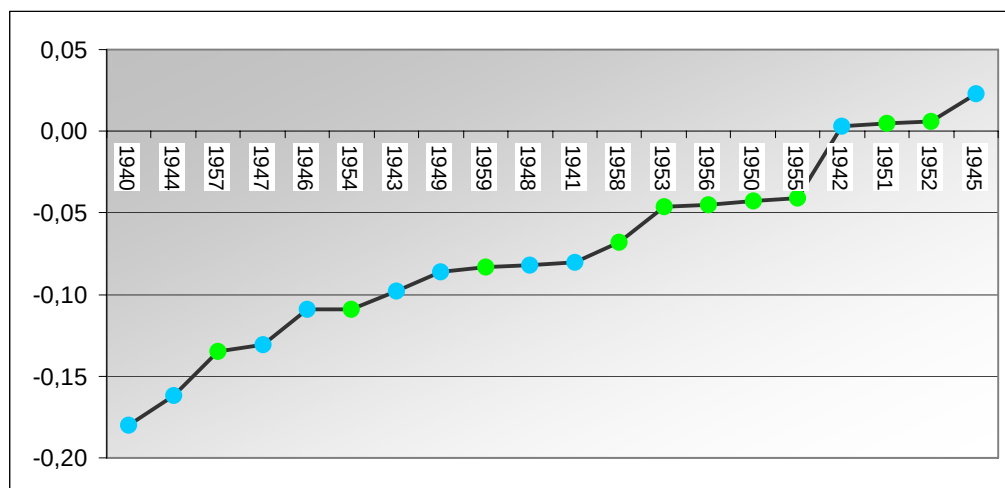


Figura 4.13 Representação das correlações entre a mortalidade infantil por ano nas décadas de 40 e 50, e a componente principal 4.

As correlações entre as taxas de incidência acumuladas das fracturas osteoporóticas do colo do fémur e a mortalidade infantil, foram também analisadas por distrito, excluindo os distritos do Porto e de Lisboa, por apresentarem mortalidades muito elevadas, que podiam estar a comprometer a análise de regressão dos demais distritos. Na **Tabela 4.18** estão apresentados os resultados desta regressão.

Tabela 4.18 Correlações e significâncias resultantes das regressões entre a mortalidade infantil por distrito e a taxa de incidência das fracturas osteoporóticas do colo do fémur, sem os distritos de Porto e Lisboa.

Ano	Correlação	Significância	Ano	Correlação	Significância	Ano	Correlação	Significância
1930	0,221	0,412	1940	0,194	0,472	1950	0,604	0,013
1931	0,451	0,08	1941	0,000	0,999	1951	-0,269	0,314
1932	0,260	0,33	1942	0,234	0,382	1952	-0,062	0,819
1933	0,305	0,251	1943	-0,006	0,981	1953	-0,180	0,505
1934	0,220	0,413	1944	-0,115	0,671	1954	-0,318	0,23
1935	0,402	0,123	1945	0,297	0,264	1955	-0,399	0,125
1936	0,378	0,149	1946	-0,007	0,981	1956	0,267	0,318
1937	0,297	0,263	1947	-0,206	0,444	1957	-0,118	0,663
1938	0,232	0,388	1948	0,106	0,696	1958	-0,081	0,765
1939	0,188	0,487	1949	0,046	0,866	1959	-0,084	0,757

A **Figura 4.14** representa o gráfico das correlações entre a mortalidade infantil em cada ano, por distrito, e a taxa de incidência das fracturas do colo do fémur, também por distrito, ao longo dos anos (os pontos verdes representam a década de 50, os pontos azuis representam a década de 40 e os pontos vermelhos a década de 30).

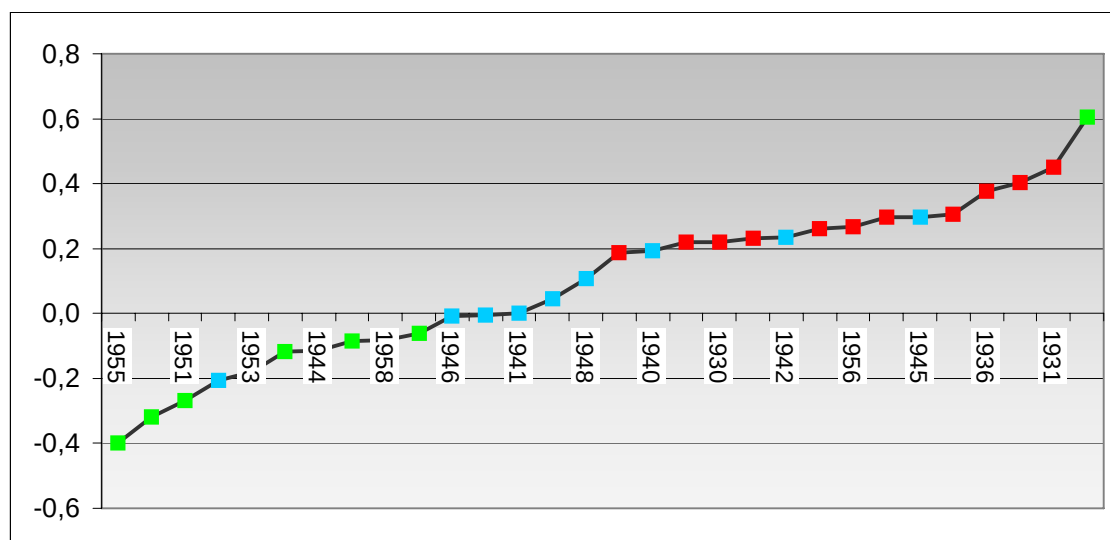


Figura 4.14 Representação das correlações entre a mortalidade infantil em cada ano, por distrito, e a taxa de incidência das fracturas do colo do fémur, também por distrito, ao longo dos anos, sem os distritos de Porto e Lisboa.

No **Anexo E** encontram-se as tabelas e os gráficos correspondentes à mortalidade infantil, por distrito, ao longo dos 30 anos a que se referem os dados.

5. DISCUSSÃO

5.1 Discussão

Os resultados encontrados, neste trabalho, indicam que existe uma associação entre factores socioeconómicos e fracturas de colo do fémur, em Portugal, nos anos de 2000 a 2002.

Em Portugal as diferenças geográficas encontradas nas taxas de incidência das fracturas do colo de fémur em indivíduos com 50 anos ou mais, por Alves S (2004) são claras. Considerando os resultados obtidos neste trabalho pode-se afirmar que existem diferenças entre as classes socioeconómicas quanto à incidência das fracturas osteoporóticas do colo do fémur. O risco relativo, em relação ao *cluster* com a menor taxa de incidência de fracturas, chega a ser 20% superior nas mulheres, nas áreas associadas ao estatuto socioeconómico mais alto. Encontrou-se que os grandes centros urbanos e regiões com um estatuto socioeconómico mais alto estão associados a taxas de incidência mais elevadas, em ambos os géneros. Mas os resultados revelam também a existência de diferenças entre os géneros no que diz respeito a esta associação entre características socioeconómicas das regiões e as taxas de incidência das fracturas osteoporóticas do colo do fémur.

No que diz respeito à caracterização socioeconómica foram encontrados *clusters* geográficos distintos, nomeadamente, entre o norte e o sul de Portugal e taxas diferentes entre os *clusters* associadas a populações mais envelhecidas e mais jovens. Por último encontrou-se que a mortalidade infantil do início do século e, portanto, as características socioeconómicas principalmente dos anos 30, podem estar relacionadas com a incidência de fracturas osteoporóticas do colo de fémur de agora.

Estes resultados sugerem que o padrão da incidência das fracturas osteoporóticas do colo do fémur em Portugal, e o risco individual de sofrer fracturas pode estar associado ao nível socioeconómico em que essa população está inserida, indo de encontro ao que afirmam Jones S e Johansen A (2004) e à quase totalidade dos estudos que relacionam características socioeconómicas e o risco de sofrer fracturas osteoporóticas. Suen L 1998 e Vestergaard P, Rejnmark L et al. 2006 justificam esta relação com base no facto de um melhor estatuto socioeconómico poder estar associado a uma melhor dieta, a um menor consumo de tabaco ou álcool, melhores habitações, melhores cuidados ou melhores tratamentos e acessos a cuidados de saúde (preventivos ou não). Por outro

lado também tem sido reportado que os grandes centros urbanos, quando comparados com áreas rurais, apresentam maiores taxas de incidência de fracturas osteoporóticas do colo do fémur, o que vai de encontro aos resultados encontrados. Também o facto de que a maior parte da população viver nas áreas associadas ao estatuto socioeconómico mais alto, portanto, áreas mais urbanas, pode estar associado à diferença entre os meios urbanos e rurais já demonstrada em outros trabalhos.

5.2 Análise demográfica *versus* fracturas osteoporóticas do colo do fémur

Em relação às diferenças geográficas encontradas nas taxas de incidência das fracturas osteoporóticas do colo de fémur, a análise das **Tabela 4.8 a 4.10** e as **Figuras 4.6 a 4.8**, permite identificar diferenças entre o género masculino e o feminino.

Quanto ao género masculino, a análise dos intervalos de confiança mostra que os valores das taxas de incidência calculadas não são estatisticamente diferentes, no entanto é possível salientar que o *cluster* associado ao estatuto socioeconómico alto é o que apresenta uma taxa maior, seguido dos *clusters* associados aos estatutos médio-baixo, baixo, médio, muito baixo e médio-alto. Estes resultados sugerem que as regiões do Porto, Lisboa e arredores de Lisboa, alto e baixo Alentejo e Algarve são as que apresentam taxas mais elevadas quando comparadas com as taxas das restantes zonas do país. O que de certa forma vai de encontro a alguns resultados obtidos por Alves S (2004). Alves S (2004) encontrou que as zonas onde a taxa de incidência das fracturas osteoporóticas do colo do fémur, nos homens, é maior são o Alentejo litoral e alguns dos arredores de Lisboa. Desta forma, com estes dados fica por justificar as também elevadas taxas encontradas em Trás-os-Montes.

Quanto ao género feminino, os intervalos de confiança já mostram que algumas das taxas de incidência encontradas são estatisticamente diferentes umas das outras. Assim, analisando estas taxas é possível verificar que a taxa mais alta é encontrada no *cluster* associado ao estatuto socioeconómico alto, seguida das taxas de incidência encontradas nos *clusters* associados aos estatutos médio-baixo, muito baixo, médio, baixo e médio-alto. Estes resultados, um pouco diferentes dos encontrados para o género masculino, sugerem que as taxas de incidência das fracturas do colo do fémur, para mulheres com 50 anos ou mais, são mais elevadas nas regiões do Porto, Lisboa e arredores e Algarve, no baixo Alentejo e em grande parte do interior Norte e Centro. Estes resultados vão novamente, de alguma forma, ao encontro do que já tinha sido encontrado por Alves S (2004). Os resultados encontrados por Alves S (2004) revelam que as regiões onde as taxas de incidência destas fracturas, em mulheres, são mais altas são o Algarve, os arredores de Lisboa e a região de Trás-os-Montes. Esta autora encontra uma clara diferença entre as taxas de incidência do Algarve e do Baixo Alentejo, diferença esta que talvez justifique a alta incidência encontrada no *cluster* associado ao estatuto

socioeconómico médio-baixo, uma vez que deste *cluster* ainda fazem parte alguns concelhos do Algarve. O que se verifica é que alguns dos concelhos que fazem parte da costa algarvia também se estendem para o interior do Alentejo, o que possivelmente os torna mais parecidos, socioeconomicamente, com as outras regiões do baixo Alentejo do que com os concelhos mais costeiros do Algarve.

É então possível afirmar que tanto nos homens como nas mulheres os grandes centros urbanos e as condições socioeconómicas neles implícita (tanto boas condições de vida, e um alto nível de escolaridade e população pouco envelhecidas e empregada em profissões que implicam um menor esforço físico, como também menos boas condições reveladas por uma grande taxa de subsídio de desemprego) apresentam as maiores taxas de incidência de fracturas. Para o género feminino, os resultados indicam a possibilidade das condições socioeconómicas subentendidas dos grandes centros urbanos e das regiões com as populações mais desfavorecidas poderem estar relacionadas com uma maior incidência de fracturas do colo do fémur, quando comparadas com áreas de condições socioeconómicas mais medianas. Nos homens também se poderá afirmar o mesmo, no entanto, o facto do *cluster* associado ao estatuto socioeconómico mais baixo não apresentar uma taxa tão elevada como as restantes leva a que esta conclusão não se seja tão óbvia, ou seja feita de forma mais reticente.

Os resultados sugerem também que as populações mais envelhecidas e mais desfavorecidas também estão associadas a taxas de incidência de fracturas elevadas, embora não as mais elevadas. Embora, este facto é mais nítido nas mulheres que nos homens.

Também é possível verificar que o *cluster* relacionado com a população mais jovem, associado ao estatuto socioeconómico médio-alto (*cluster* 4), é o *cluster* que, em ambos os sexos, apresenta a taxa de incidência mais baixa. O que sugere que a população mais jovem está relacionada com uma menor incidência destas fracturas, indo de encontro à grande parte da bibliografia que indica um aumento do risco de sofrer fracturas do colo do fémur à medida que se envelhece. No entanto, mais uma vez esta relação não é tão clara nos homens, pois a população mais envelhecida, associada ao estatuto médio-baixo (*cluster* 3) apresenta a segunda taxa mais baixa de incidência destas fracturas.

O facto de quanto ao género masculino se encontrar uma maior taxa de incidência de fracturas nos grandes centros urbanos e a menor taxa numa área mais rurais vai de

encontro aos resultados encontrados por Billsten M, Sernbo I et al. 2000; Sawka A, Thabane L et al. 2005 e Soveid M, Serat A et al. 2005. Estes autores, nos seus trabalhos, reportam a existência de uma menor taxa de incidência em meios rurais, quando comparadas com as taxas dos meios urbanos e justificaram estes resultados no facto do exercício físico promover força muscular e uma postura mais estável, nas grandes diferenças entre as dietas destes dois meios, havendo em meios urbanos há uma maior deficiência de vitamina D e, no facto de em meios urbanos existir uma maior coexistência de problemas de saúde e maior uso frequente de drogas. Estes autores sugerem, no entanto, que outro factor relacionado com estas diferenças entre os meios urbanos e rurais, se relaciona com o facto de nos meios urbanos existirem mais mulheres a viverem sozinhas.

Com relação às diferenças socioeconómicas encontradas em Portugal, torna fácil verificar que existe uma clara diferença entre o norte e o sul do país, existindo dois padrões distintos.

No norte do país é possível verificar a existência três *clusters* claramente distintos que não se encontram no sul, um relativo aos arredores do Porto, outro relativo ao litoral e outro relativo ao interior. Em relação ao sul do país também é possível distinguir três *clusters*, mas não exclusivos ao sul, ou seja, existem regiões mais a norte que também fazem parte *clusters*. Ainda em relação aos *clusters* encontrados no sul do país a clara diferença entre eles torna-se mais esbatida, quando comparando com as diferenças encontradas entre os *clusters* no norte do país. É possível distinguir um *cluster* relativo a Lisboa, arredores de Lisboa e alguns concelhos mais costeiros do Algarve (de salientar que também pertence a este *cluster* a cidade do Porto e alguns dos arredores mais próximos), um que abrange mais o alto Alentejo e um terceiro nitidamente relativo ao baixo Alentejo.

No que se refere à relação entre as características socioeconómicas e a taxa de incidência das fracturas osteoporóticas do colo do fémur também se obtiveram resultados relevantes.

A partir das variáveis seleccionadas provenientes do Censos de 2001 (dados estes fornecidos pelo INE e posteriormente trabalhados) foi realizada a análise de componentes principais. Pela análise das **Tabelas 4.2 a 4.4** é possível verificar que a componente 2 se encontra correlacionada aproximadamente 7% com a taxa de

incidência de fracturas osteoporóticas do colo do fémur em mulheres. Sendo significativo, este resultado sugere que um maior nível de escolaridade, uma maior percentagem de indivíduos com profissões científicas, técnicas, artísticas, directores e quadros superiores administrativos e pessoal administrativo (ou seja, indivíduos empregados nos grupos de profissões CNP 1, 2 ou 3), uma maior percentagem de indivíduos que trabalham no comércio ou como vendedores, pessoal dos serviços de protecção e segurança, dos serviços pessoais e domésticos (ou seja, empregados nos grupos de profissões CNP 4 e 5), ou um índice *per capita* mais alto, podem estar relacionados com a uma maior incidência destas fracturas, nas mulheres.

De certa forma os resultados encontrados vão de encontro aos resultados encontrados por Vestergaard P, Rejnmark L et al. (2006). Estes autores incluíram no seu estudo todos os indivíduos que sofreram uma fractura durante o ano de 2000 e investigaram os efeitos do rendimento, da educação, da ocupação e do estado civil, no risco de sofrer fracturas. Os resultados que encontraram revelam que o rendimento não está associado com o risco de sofrer fracturas, mas que o estatuto de se ter ou não trabalho está, isto é, como os autores referem, talvez não seja o rendimento que determina esse risco mas a estrutura social de ter um trabalho. Vestergaard P, Rejnmark L et al. (2006) reportam também que ter um companheiro é protector no que respeita à incidência de fracturas do colo do fémur. Mas é no nível de escolaridade que os autores encontram os resultados mais interessantes. Como referem, um maior nível de escolaridade pode ser facilmente associado a um aumento do risco de sofrer uma fractura. Contudo estes autores encontraram nos seus resultados uma interessante interacção entre esta relação e a idade: entre os indivíduos mais novos, observaram uma redução do risco de fracturas entre aqueles que tinham um maior nível de escolaridade e, entre indivíduos com 60 anos ou mais encontraram o oposto, ou seja um aumento do risco para aqueles que tinham um maior nível de escolaridade. Vestergaard P, Rejnmark L et al. (2006) não encontraram explicação para este facto mas sugerem que a justificação se prenda nos diferentes tipos de educação que existem entre indivíduos mais novos e entre indivíduos mais velhos.

Apesar destes resultados relacionados com a educação, autores como Benetos I, Babis G et al. (2007) e Suen L (1998) mencionam resultados anteriores encontrados em mulheres Turcas e Espanholas onde são os níveis mais baixos de escolaridade que estão

associados a uma aumento do risco do sofrer uma fractura do colo do fémur. Também Jones S, Johansen A et al. (2004) referem um trabalho italiano em que foram estudadas 6160 mulheres italianas e cujos resultados mostram que a prevalência da osteoporose era cerca de 28,7% entre mulheres com o nível mais baixo de escolaridade e cerca de 18,3% entre mulheres com o nível mais alto de escolaridade. Existem ainda estudos, como o de Farahmand BY, Persson PG et al. (2000) que não encontram nenhuma relação entre a incidência das fracturas e o nível de escolaridade, o que de certa forma poderá sugerir que não foram as variáveis relativas à educação que tiveram influência na existência da correlação entre a componente 2 e a taxa de incidência de fracturas.

Por outro lado, Suen L (1998), encontrou uma relação entre as profissões mais sedentárias e o aumento da incidência das fracturas do colo do fémur. O que confirma os resultados obtidos neste trabalho, ou seja, ao facto de uma maior percentagem de indivíduos empregados nos grupos de profissões CNP 1, 2, 3, 4 ou 5 poder estar relacionado positivamente com a incidência das fracturas. No seu estudo Suen L (1998), analisou em 416 sujeitos, dos subúrbios de Sydney (Austrália), a relação entre o risco de sofrer uma fractura do colo do fémur e a actividade física associada à ocupação e o estatuto socioeconómico definido em termos de ocupação. O autor encontrou um aumento do risco de sofrer fracturas do colo do fémur entre mulheres com um trabalho sedentário aos 50 anos ou um trabalho maioritariamente de transporte de pesos, quando comparado com as ocupações de intermédia actividade física; também observou alguma evidência que os níveis mais altos socioeconómicos estão associados a uma diminuição do risco de sofrer fracturas. Os resultados obtidos relativamente às profissões mais sedentárias confirmam outros encontrados em estudos realizados no Reino Unido. Uma possível explicação para estes resultados reside no facto dos osteócitos responderem melhor a uma diversificação de exercícios físicos, do que, a uma repetição de um mesmo exercício. Este autor refere ainda outros trabalhos em que não são encontradas diferenças na densidade mineral óssea entre empregados de escritório com trabalhos sedentários e os funcionários que distribuem o correio ou notícias, talvez porque o caminhar diário e o subir escadas habitual, são exercícios insuficientes para produzir um efeito anabólico na mineralização do osso. Estes resultados sugerem que talvez a associação encontrada entre a segunda componente principal e a taxa de incidência de fracturas nas mulheres, se deva ao facto desta componente estar relacionada com

pessoal administrativo ou uma maior percentagem de indivíduos que trabalham no comércio ou em serviços pessoais (ou seja, trabalhos mais sedentários).

A análise das regressões realizadas entre as componentes e a taxa de incidência dos homens não sugerem, mas também não excluem, a hipótese desta componente estar ou não relacionada com a taxa de incidência, uma vez que os resultados não foram significativos.

Com base no significado da componente principal 1 teria sido de esperar encontrar uma associação entre a componente principal 1 e a taxa de incidência, no entanto em nenhum dos géneros o resultado foi significativo.

Nos seus estudos, Farahmand BY, Persson P-G (2000), Vestergaard P, Rejnmark L et al. (2006) revelam através dos seus resultados a existência de uma associação entre o facto de se viver sozinho e o aumento do risco de sofrer uma fractura. Também Benetos I, Babis G et al. (2007) fazem referência a esta relação. Analisando a **Tabela 4.5** não é possível confirmar se em Portugal existe ou não alguma associação entre os indivíduos com 65 anos ou mais que vivem sozinhos e o aumento da taxa de incidência das fracturas do colo do fémur, uma vez que os resultados encontrados não são significativos.

Com os fundamentos encontrados nos resultados encontrados por Vestergaard P, Rejnmark L et al. (2006), acima referidos, também poderia ser de esperar uma relação entre a componente 3 e a taxa de incidência de fracturas. Uma vez que a componente 3 está relacionada com o subsídio de desemprego. Esta variável, que diz respeito à percentagem de indivíduos, com mais de 15 anos, que recebem o subsídio de desemprego, foi de facto escolhida para tentar preencher a lacuna que a falta de informação sobre os rendimentos da população suscitou. No entanto, não tendo sido significativo o resultado, não é possível tirar nenhuma conclusão.

Apesar não existido acesso a estudos cuja metodologia se assemelhasse à utilizada, é possível afirmar, que dentro das limitações existentes, estes resultados vão de encontro à grande parte da bibliografia que estuda a relação de factores sociais e económicos com o risco de sofrer fracturas do colo do fémur ou com a taxa de incidência destas fracturas. Estes estudos concluem que existe de facto esta relação, por vezes mais esbatida e noutra vezes mais nítida, tal como existe em outras doenças que não a osteoporose. Assim os resultados encontrados confirmam o que Farahmand BY, Persson P-G et al.

(2000) mencionam, ou seja, que o risco de sofrer uma fractura do colo do fémur, tal como para outras doenças e como acontece com a mortalidade, é modestamente influenciado pelo estatuto socioeconómico.

No entanto, Jones S, Johansen A et al. (2004) advertem que não é fácil generalizar os resultados encontrados sobre a associação do estatuto socioeconómico e a osteoporose. Pois para além do tipo de estudo, a definição de estatuto socioeconómico difere de país para país e pode ser baseado no rendimento, na educação ou no acesso a bens materiais, entre outros. Jones S, Johansen A et al. (2004) referem ainda que as relações entre a incidência de fracturas osteoporóticas e o estatuto socioeconómico, entre vários países, podem tomar direcções diferentes. Por exemplo, Vestergaard P, Rejnmark L et al. (2006) definem o estatuto socioeconómico através dos rendimentos, do estado civil, do nível de escolaridade e do ter ou não ocupação. Já Farahmand BY, Persson P-G et al. (2000) definem o estatuto socioeconómico através da ocupação, do nível de escolaridade, do rendimento e condições de alojamento. Por outro lado, Jones S, Johansen A et al. (2004) definem a privação socioeconómica com base nas taxas de desemprego, no grande número de pessoas por alojamento, no facto do alojamentos terem ou não acesso a carros e no facto de não existir nenhum dono a ocupar o alojamento.

Apesar das diferenças entre as características socioeconómicas seleccionadas, cada estudo analisa a relação entre cada uma destas características e a taxa de incidência de fracturas do colo do fémur, em separado. Chegando a resultados mais ou menos consistentes dependendo da característica em questão. Farahmand BY, Persson P-G et al. (2000) afirmam que os níveis mais baixos de escolaridade e o não se ser casado são factores que têm sido associados de forma mais consistente ao risco de sofrer fracturas, mas outros estudos, que relacionam a ocupação com a densidade mineral óssea, têm encontrado resultados conflituosos. Farahmand BY, Persson P-G et al. (2000) referem ainda que a associação entre as fracturas do colo do fémur e outros aspectos do estatuto socioeconómico, como a habitação e o rendimento familiar, não têm sido tão investigados.

Pela diferença da metodologia utilizada neste trabalho e trabalhos de outros autores e por estes últimos analisarem cada característica socioeconómica separadamente, não existem grandes parâmetros de comparação entre os resultados aqui encontrados e pelos

obtidos por outros autores. O facto da metodologia utilizada nesta dissertação não ser utilizada por outros autores, no estudo das fracturas do colo do fémur, pode ser um indicativo de que o estudo da relação entre aspectos socioeconómicos e as taxas de incidências das fracturas tem sido negligenciado em estudos ecológicos.

5.3 Mortalidade infantil *versus* fracturas osteoporóticas do colo do fémur

A mortalidade infantil foi utilizada como possível e único indicador socioeconómico que poderia permitir a caracterização das condições socioeconómicas em que as pessoas com mais de 50 anos viviam quando nasceram e durante a sua infância ou adolescência.

O estudo da possível associação entre a incidência das fracturas do colo do fémur e a mortalidade infantil é inovador não havendo, portanto, resultados que comprovem ou não o que neste trabalho foi encontrado. No entanto, um breve conhecimento sobre a osteoporose poderia sugerir a existência desta associação. A existência de uma maior mortalidade infantil pode ser indicador de piores condições de vida e piores condições socioeconómicas, o que por sua vez pode estar associado a um crescimento mais pobre em termos de alimentação, higiene e saúde, que está directamente relacionado com o ganho de massa óssea, sendo assim, aumentado o risco de vir a sofrer fracturas.

Analisando a **Tabela 4.13** e a **Figura 4.9** é possível verificar que a década em que há uma maior correlação entre a mortalidade infantil e a incidência de fracturas osteoporóticas do colo do fémur é a década correspondente aos anos 30, tendo sido considerado uma excepção o ano de 1950. Também se pode verificar que na década de correspondente aos anos 40 esta correlação abranda, tornando-se mesmo negativa na grande maioria dos anos da década de 50.

Estes resultados apesar de não serem significativos indicam que as condições de vida à 70 anos atrás podem estar relacionadas com a incidência de fracturas do colo de fémur agora. E, assim a incidência destas fracturas pode ser explicada em parte com base nas condições socioeconómicas do início e meados do século passado.

Analisando as **Figuras 1. a 18.** (no **Anexo E**), que representam a mortalidade infantil em cada ano por distrito, pode verificar-se que o Porto e Lisboa apresentam na maioria dos anos uma elevada mortalidade infantil. Estes resultados sugerem que sendo estas as duas cidades principais muita gente aqui nascia, aumentando assim a mortalidade infantil. Mas as evidências indicam também que, por serem duas cidades grandes e principais, é possível que existissem grandes diferenças socioeconómicas entre determinadas zonas das próprias cidades, existindo zonas de extrema pobreza. Não sendo então clara a justificação para as elevadas taxas encontradas no Porto e em Lisboa, estudaram-se os resultados da mesma regressão, mas desta vez sem os dados referentes a estas duas cidades. Como é possível observar na **Tabela 4.18** e na **Figura**

4.14, a correlação entre a mortalidade infantil e a taxa de incidência das fracturas parece aumentar ligeiramente em alguns anos e diminuir ligeiramente noutros, quando comparados estes resultados com os da análise anterior referida. É de salientar o aumento quase consistentemente todos os anos da década de 50 desta correlação. No entanto, tal como acontecia com os resultados anteriores estes resultados continuaram a não ser significativos (excepto, mais uma vez, o resultado referente ao ano de 1950), não se podendo tirar uma conclusão mais precisa dos resultados.

Não foi possível encontrar uma justificação clara para o facto do ano de 1950 apresentar a correlação mais significativa, sendo provável que se deva meramente ao acaso, razão pela qual foi considerado um ano atípico.

No que diz respeito aos resultados obtidos na associação entre a mortalidade infantil e as componentes principais, pela análise das **Tabela 4.14 a 4.14** é possível perceber que as componentes principais onde são obtidas as maiores correlações são as componentes 2 e 3. De entre estas duas componente, observando as significâncias, a componente 3 é a que apresenta melhores resultados, por serem mais significativos.

Estes resultados sugerem que algumas condições de alojamento (como um menor número de divisões por alojamento, um maior número de pessoas por divisão e a uma menor percentagem de utilização de aquecimento central) ou uma maior percentagem de indivíduos que recebem subsídio de desemprego são mais relacionadas com as características socioeconómicas de às sessenta e cinquenta anos atrás que as outras variáveis. E que, embora um pouco menos, um nível de escolaridade mais alto, ou indivíduos com profissões científicas, técnicas, artísticas e administrativas (ou seja, indivíduos empregados nos grupos de profissões CNP 1, 2 ou 3), ou indivíduos que trabalham no comércio ou como vendedores, pessoal dos serviços de protecção e segurança, dos serviços pessoais e domésticos (ou seja, empregados nos grupos de profissões CNP 4 e 5), ou um índice *per capita* mais alto, também possam estar relacionados com essas mesmas condições.

Estes resultados permitem afirmar que, de certa forma, as características socioeconómicas do início e meados do século passado podem ser, com os devidos cuidados, analisadas com base nas componentes principais seleccionadas. Ou seja, os resultados encontrados, em termos das diferenças socioeconómicas nas diferentes regiões de Portugal, podem ser generalizados para o passado.

5.4 Conclusão

A intensidade real das fracturas do colo do fémur em Portugal já é conhecida e conhecido também é o padrão geográfico da incidência destas fracturas. Então tornava-se urgente entender o porquê deste padrão e quais os factores que contribuem para o seu aparecimento. Para isso foram estudadas as características socioeconómicas do país.

Com base nas informações disponíveis conclui-se que os factores socioeconómicos explicam uma parte das diferenças geográficas encontradas na incidência das fracturas do colo do fémur e que existem diferenças entre os géneros feminino e masculino. Nas mulheres as maiores taxas são encontradas nas maiores áreas urbanas mas também nas regiões mais desfavorecidas, o que indica que as características associadas aos dois extremos (áreas mais urbanas e, portanto, menor exercício físico e estilos de vida sedentários, e áreas mais rurais e, portanto, população mais envelhecida, maior taxa de analfabetismo e trabalhos mais pesados) podem justificar as taxas de incidência, quando comparadas com características mais medianas. Nos homens foi encontrada uma maior relação entre as taxas de incidência e as características socioeconómicas do sul do país, o que sugere que as diferenças geográficas nas taxas de incidência das fracturas de colo do fémur, encontradas nos homens, se devam no norte a motivos diferentes que não as características socioeconómicas. Em ambos os sexos a população mais jovens pode ser associada a uma menor taxa de incidência destas fracturas. É possível concluir que existem diferenças espaciais relacionadas com as características socioeconómicas da população, em Portugal. Nomeadamente, diferenças entre o norte e o sul do país, onde é possível observar dois padrões muito distintos e, apesar de no norte do país seja possível encontrarem-se algumas características socioeconómicas do sul, o contrário não acontece, o que sugere que o norte e o sul são socioeconomicamente diferentes.

A análise dos dados relacionados com a mortalidade infantil permite concluir que as diferenças socioeconómicas entre os concelhos talvez se tenham mantido ao longo do tempo e que as características socioeconómicas do início e meados do século passado podem, com os devidos cuidados, ser analisadas com base nas variáveis disponíveis. É possível ainda concluir que as condições de vida à 70 anos atrás podem estar relacionadas com a incidência de fracturas do colo de fémur agora.

O sucesso deste estudo depende da qualidade das variáveis disponíveis para o realizar e da qualidade das bases de dados disponíveis. O trabalho debateu-se com algumas limitações como a falta de variáveis censitárias mais discriminantes, como por exemplo falta de informação relativa a rendimentos. O facto da análise ter sido realizada a nível concelhio pode ser a razão de parte das análises não terem apresentado significância estatística, uma vez que a amostra é relativamente pequena.

Conclui-se assim, que em Portugal, tal como em outros países, o risco de sofrer fracturas é moderadamente influenciado pelo estatuto socioeconómico da população. O padrão existente em Portugal a nível das fracturas do colo do fémur é explicado em parte através dos factores socioeconómicos. No entanto, este padrão pode também estar relacionado com factores genéticos, ambientais, culturais, entre outros. Ao longo deste processo surgiram indicações que estas diferenças podem ser atribuídas a diferentes causas, não só a factores socioeconómicos, mas também a factores relacionados com a qualidade da água ou relacionados com a população imigrante (particularmente, a proveniente dos países nórdicos e que imigra, principalmente, para o sul do país). Pelo que é necessário continuar a investigar as razões que estão por detrás do padrão geográfico encontrado por Alves S (2004).

REFERÊNCIAS

- Aldenderfer M, Blashfield R (1984). Clusters Analysis. Sage publications.
- Almeida N (1991). The scientific statute of epidemiology. *Revista de Saúde Pública*, **35** (5): 339-340.
- Alves S (2004). Epidemiological Descriptive Study of Osteoporosis in Portugal, Using Geographical Information System. [dissertação de mestrado] *Mestrado de Engenharia Biomédica – Universidade do Porto*.
- Araújo D, Barros H, Silva C, Pereira J (1995). Osteoporosis in Portuguese Healthy Women. 3º Congresso Brasileiro de Epidemiologia e 2º Congresso Ibero-Americano de Epidemiologia.
- Baron YM, BrinCAT M., Galea R, Baron AM (1994). The epidemiology of osteoporotic fractures in a Mediterranean country. *Calcified Tissue International*, **54** (5): 365-369.
- Barros H (1997). A prática e a boa prática da epidemiologia. A necessidade de uma reflexão em Portugal. *Revista de Epidemiologia, Arquivos de Medicina É a revista de Epidemiologia ou os Arquivos de Medicina? São duas revistas diferentes!!!*, **11** (5): 2-2.
- Benetos I, Babis G, Zoubos A, Benetou V, Soucacos P (2007). Factors affecting the risk of hip fractures. *Injury*, **38**: 735-744.
- Billsten M, Sernbo I, Ornstein E, Rogmark C, Johnell O (2000). The incidence and mortality of hip fractures in two regions of southern Sweden 1992-1997. *Current Orthopaedics*, **14** (3): 182-188.
- Burstyn I (2004) Principal Component Analysis is a Powerful Instrument in Occupational Hygiene Inquiries. *Annals of Occupational Hygiene*, **48** (8): 655-661.
- Cadima J (2002). Apontamentos de estatística multivariada descritiva. Departamento de Matemática – Mestrado em Matemática Aplicada às Ciências Bilógicas, Universidade do Porto.
- Canhão H, Fonseca J, Queiroz M (2005). Epidemiologia da osteoporose. Mecanismos de remodelação óssea e factores protectores do osso. Serviço de Reumatologia do Hospital de Santa Maria Faculdade de Medicina de Lisboa, *Acta Reumatológica Portuguesa*, **30**: 225-240.
- Carvalho MS, Nobre FF (2001). Análise de dados espaciais em saúde. *Cadernos de Saúde Pública* **17** (5).
- Carvalho MS, Pina MF, Santos S (2000). Conceitos básicos de Sistemas de Informação Geográfica e Cartografia aplicados à saúde. *Organização Panamericana da Saúde - Ministério da Saúde*.
- Court-Brown C, Caesar B (2006). Epidemiology of adult fractures: A review. *International Journal Care Injured*, **31** (8): 691-697.

- Crabtree N, Lunt M, Holt G, Kroger H, Burger H, Grazio S, Khan K-T, Lorenc RS, Nijs J, Stepan J, Falch JA, Miazgowski T, Raptou P, Pols HA, Dequeker J, Havelka S, Hoszowski K, Jajic I, Czekalki S, Lyritis G, Silman AJ, Reeve J (2000). Hip geometry, bone mineral distribution, and bone strength in European men and women: The EPOS Study. *Bone*, **27** (1): 151-159.
- Cramer D (2003). Advanced quantitative data analysis. Open University Press.
- Cummings SR, Melton L (2002). Epidemiology and outcomes of osteoporotic fractures. *The Lancet*, **359** (9319): 1761-1767.
- Cummings SR, Nevitt M, Browner W, Stone K, Fox D, Ensrud K, Cauley J, Black D, Vogt T (1995). Risk factors for hip fracture in white women. *New England Journal Medicine*, **332** (12): 767-773.
- Dennison E, Cooper C (2000). Epidemiology of osteoporotic fractures. *Hormone Research*, **54** (1): 58-63.
- Diamond T, Eisman J, Mason R, Nowson C, Pasco J, Sambrook P, Wark J - Working Group of the Australian and New Zealand Bone and Mineral Society, Endocrine Society of Australia and Osteoporosis Australia (2005). Vitamin D and adult bone health in Australia and New Zealand: a position statement. *The Medical journal of Australia*, **183** (1): 52-53.
- Dias A. (2000). Epidemiologia da osteoporose em Portugal: Análise comparativa com outros países. *Acta Reumatológica Portuguesa*, **97**: 21-31.
- Dias A, Ventura S, Fernandes M, Brito I, Vaz C, Vaz A (1995). Osteoporose vertebral – Protocolo do estudo epidemiológico multicêntrico na Europa. *Revista Portuguesa de Reumatologia*, **6** (55): 1377-1397.
- Dicker R (2005). Principles of epidemiology. An introduction to applied epidemiology and biostatistics. *American Public Health Association*, **202**: 789-5600.
- Dunteman G (1989). Principal Components Analysis. Sage Publications.
- Elliot P, Wakefield J, Best N, Briggs D (2000). Spatial epidemiology: methods and applications. Oxford University Press.
- EPOS - The European Prospective Osteoporosis Study Group (2002). Incidence of Vertebral Fractures in Europe: Results From the European Prospective Osteoporosis Study (EPOS). *Journal of Bone and Mineral Research*, **17** (4): 716-724.
- Everitt B (1974). Cluster Analysis. Ed: Heinemann.
- Farahmand BY, Persson P-G, Michaelsson K, Baron JA, Parker MG, Ljunghall S (2000). Socioeconomic status, marital status and hip fracture risk: A population-based case-control study. *Osteoporosis International*, **11**: 803-808.
- Finnern H, Sykes D (2003). The hospital cost of vertebral fractures in the EU: estimates using national datasets. *Osteoporosis International*, **14** (5): 429-436.

- Feldstein AC, Nichols G, Rowel E, Elmer PJ, Smith DH, Herson M, Aickin M (2005). The near absence of ostoporosis treatment in older men with fractures. *Osteoporosis International*, **16**: 953-962.
- Felsenberg D, Silman AJ, Lunt M, Armbrecht G, Ismail AA, Finn J, Cockerill WC, Banzer D, Benevolenskaya LI, Bhalla A, Bruges AJ, Cannata JB, Cooper C, Dequeker J, Eastell R, Felsch B, Gowin W, Havelka S, Hoszowski K, Jajic I, Janott J, Johnell O, Kanis JA, Kragl G, Lopes VA, Lorenc R, Lyritis G, Masaryk P, Matthis C, Miazgowski T, Parisi G, Poor G, Raspe HH, Reid DM, Reisinger W, Scheidt-Nave C, Stepan JJ, Todd C, Weber K, Woolf AD, Yershova OB, Reeve J, O'Neill TW (2002). Incidence of vertebral fracture in Europe: Results from the European Prospective Osteoporosis Study (EPOS). *Journal of Bone and Mineral Research*, **17** (4): 716-724.
- Fleck M, Bourdel M (1998). Method of simulation and choice of factors in the analysis of principal components. *Revista Saúde Pública*, **32** (3): 267-272.
- Giversen I (2006). Time trends of age-adjustes incidence rates of first hip fractures: A register-based study among older people in Virbog County, Denmark, 1987-1997. *Osteoporosis International*, **17** (4): 552-564.
- Hashefi M, Burke M, Dempster D, Field E, Sebba A (2005). Practical guide to preventing, recognizing, and managing osteoporosis. *George Washington University Medical Center*.
- Hernández JL, Olmos JM, Alonso MA, González-Fernández CR, Martínez J, Pajarón M, Llorca J, González-Macías J (2005). Trend in hip fracture epidemiology over a 14-year period in a Spanish population. *Osteoporosis International*, **17** (3): 464-470.
- Hunter D, Sambrook P (2000). Bone loss, Epidemiology of bone loss. *Arthritis Research and Therapy*, **2** (6): 441 – 445.
- Hyvarinen A, Karhunen J, Oja J (1994). Independent Component Analysis. *John Wiley & Sons, Inc*, capítulo 6.
- Jéquier V, Burnand B, Vader J-P, Paccaud F (1995). Hip fracture incidence in the Conton of Vaud, Switzerland, 1986-1991. *Osteoporosis International*, **5** (3): 191-195.
- Johnell O, Kanis J (2004). An estimate of the worldwide prevalence, mortality and disability associated with hip fracture. *Osteoporosis International*, **15**: 897-902.
- Johnston C, Jr., Slemenda C (1995). Pathogenesis of osteoporosis. *Bone*, **17** (2): 195-225.
- Jolliffe I (1986). Principal Component Analysis. *Springer*.
- Jones S, Johansen A, Brennan J, Butler J, Lyons R (2004). The effect of socioeconomic deprivation on fracture incidence in the United Kingdom. *Osteoporosis International*, **15**: 520-524.
- Kanis J, Delmas P, Burckhardt P, Cooper C, Torgerson D (1997). Guidelines for diagnosis and management of osteoporosis. *Osteoporosis International*, **7**: 390-406.

- Kanis J, Johnell O, Oden A, Jonsson B, Laet C, Dawson A (2000). Risk of hip fracture according to the World Health Organization criteria for osteopenia and osteoporosis. *Bone*, **27** (5): 585-590.
- Kannus P, Niemi S, Parkkari J, Palvanen M, Vuor I, Jarvinen M (1999). Hip fractures in Finland Between 1970 and 1997 and predictions for the future. *The Lancet*, **353**: 802-805.
- Karagas MR, Baron JA, Barret JA, Jacobsen SJ (1996). Patterns of fractures among the United States Elderly: Geographic and Fluoride Effects. *Annals of Epidemiology*, **6** (3): 209-216.
- Kawachi I (2002). Social epidemiology. *Social Science & Medicine*, **54** (12): 1739-1741.
- Krieger N (2001). A glossary for social epidemiology. *Journal of Epidemiology and Community Health*, **55**: 693-700.
- Krieger N (2001). Theories for social epidemiology in the 21st century: an ecosocial perspective. *International Epidemiological Association*, **30**: 668-677.
- Krzanowski W, Marriot F (1994). Multivariate analysis Part.1 – Distributions, ordination and inference. *Edward Arnold*.
- Laet C, Pols H (2000). Fracture in the Elderly: epidemiology and demography. *Baillière's Clinical Endocrinology and Metabolism*, **14** (2): 171-179.
- Lane N (2006). Epidemiology, etiology, and diagnosis of osteoporosis. *American Journal of Obstetrics and Gynaecology*, **194** (2): S3-11.
- Lashas L, Masanauskaite D, Lashene D, Masanauskiene E (1996). Rate of characteristics for osteoporotic fractures in the elderly population of Kaunas Region, Lithuania. *Scandinavian Journal of Rheumatology*, **103**: 16-20.
- Last J (2001). A dictionary of epidemiology. *Oxford University Press*.
- Lilienfeld D, Stolley P (1994). Foundations of epidemiology. *Oxford University Press*.
- Lips P (1997). Epidemiology and predictors of fractures associated with osteoporosis. *The American Journal of Medicine*, **103**: 3S-11S.
- Lesic A, Jarebinski M, Pekmezovic T, Bumbasirevic M, Spasovski D, Atkinson HD (2006). Epidemiology of hip fractures in Belgrade, Serbia Montenegro, 1990-2000. *Archives of Orthopaedic and Trauma Surgery*, **127** (3): 179-183.
- Lodthus C, Osnes E, Falch J, Kaastad T, Kristiansen I, Nordsletten L, Stensvold I, Meyer H (2001). Epidemiology of hip fractures in Oslo, Norway. *Bone*, **29** (5): 413-418.
- Lonnroos E, Kautiainen H, Karppi P, Huusko T, Hartikainen S, Kiviranta I, Sulkava R (2006). Increased incidence of hip fractures. A population based-study in Finland. *Bone*, **39** (3): 623-627.
- Marcus R, Feldman D, Kelsey J (1996). Osteoporosis. *Academic Press*.

- Marques J, Marques M (2005). The principal components in the reduction of variables in a multiple regression model. *Revista FAE*, **8** (1): 93-101.
- Melton LJ, Achenbach SJ, O'fallon WM, Khosla S (2002). Secondary osteoporosis and the risk of distal forearm fractures in men and women. *Bone*, **31** (1): 119-125.
- Orimo H, Hoshimoto T, Sakata K, Yoshimura N, Suzuki T, Hosoi T (2000). Trends in the incidence of hip fracture in Japan, 1987-1997: the third nation survey. *Journal of Bone and Mineral Metabolism*, **18**: 126-31.
- Paspati I, Galanos A, Lyritis GP (1998). Hip fracture epidemiology in Greece during 1977-1992. *Calcified Tissue International*, **62** (6): 542-547.
- Parkkari J, Kannus P, Niemi S, Pasanen M, Järvinen M, Lühje P, Vuori I (1994). Increasing age-adjusted incidence of hip fractures in Finland: the number and incidence of fractures in 1970-1991 and prediction for the future. *Calcified Tissue International*, **55**: 342-345.
- Pearce N (2005). A short introduction to epidemiology. Wellington, Centre for Public Health Research.
- Pestana M, Gageiro J (2005). Análise de dados para ciências sociais: A complementaridade do SPSS. Edições Sílabo, 4ª Edição.
- Pina MF (2003³). Uso dos Sistemas de Informação Geográficos na Análise de Dados da Saúde. *Arquivos de Medicina*, **17** (4).
- Pina MF (1998). Potencialidades dos sistemas de informação geográfica na área da saúde. Em: Najar AL. & Marques EC. Saúde e espaço: Estudos metodológicos e técnicas de análise. Fiocruz.
- Pina MF, Alves S, Barros H, Barbosa M. Hip fractures are clustering in space: an epidemiological analysis in Portugal. *Osteoporosis International*, in Press.
- Prado A, Baranauskas M, Medeiros C (2000). Cartografia e sistemas de informação geográfica como sistemas semióticos: uma análise comparativa. *Anais II workshop brasileiro de geoinformatica*, 88-95.
- Press J (1972). Applied Multivariate Analysis. Holt, Rinehart and Winston, capítulo 9: 283-301.
- Ramalho W, Skaba D, Barcellos C, Pina MF, Magalhães M (2007). Conceitos e Arquitetura de Sistemas de Informações Geográficas. Em: Santos S, Souza-Santos R (orgs) Sistemas de Informações Geográficas e Análise Espacial na Saúde Pública. Série Capacitação e Atualização em Geoprocessamento em Saúde, vol. 2, pp. 15-35.
- Reginster JY, Gillet P, Gosset C (2001). Secular increase in the incidence of hip fractures in Belgium between 1984 and 1996: need for a concerted public health strategy. *Bulletin of the World Health Organization*, **79** (10): 942-946.
- Reis E (2001). Estatística multivariada aplicada. Edições Sílabo, 2ª Edição.

- Resendes A, Barcellos C, Skaba D, Oliveira E, Gondim G, Rojas L, Pina MF, Magalhães M, Peiter P, Santos R, Gracie R, Santos S (2006). Abordagens espaciais na saúde pública. Série B. Textos básicos de Saúde, 1.
- Ricketts T (2003). Geographic Information Systems and Public Health. Annual Review of Public Health, 24: 18.1-18.6.
- Rocha L (2000). Sistemas de Informação Geográfica em Saúde: conceitos básicos. Organização Pan-Americana da Saúde.
- Sawka A, Thabane L, Papaioannou A, Gafni A, Ioannidis G, Papadimitropoulos E, Hopman W, Cranney A, Hanley D, Pickard L, Adachi J (2005). Health-related quality of life measurements in elderly Canadians with osteoporosis compared to other chronic medical conditions: a population-based study from the Canadian Multicentre Osteoporosis Study (CaMos). Osteoporosis International, 16 (12):1836-1840.
- Scremin M (2003). Método para a selecção do número de componentes principais com base na lógica difusa. [dissertação] Universidade de Santa Catarina.
- Schutte H (1995). Socioeconomic impact of osteoporosis. A review of the literature. European Journal of Radiology, 20: 165-169.
- Sneath P, Sokal R (1973). Numerical Taxonomy. The principles of numerical classification. W. H. Freeman.
- Souza E, Grundy E (2004). Health promotion, social epidemiology, and social capital: associations and perspectives for public health. Cadernos de Saúde Pública, 20 (5): 1354-1360.
- Soveid M, Serat AR, Massompoor M (2005). Incidence of hip fracture in Shiraz, Iran. Osteoporosis International, 16 (11): 1412-1416.
- Stevens J (1996). Applied multivariate statistics for the social sciences. 3ª Edição. Lawrence Erlbaum Associates.
- Suen L (1998). Occupation and risk of hip fracture. Journal Public Health Medicine, 20 (4): 428-433.
- Van Stan TP, Dennison EM, Leufkens H, Cooper C (2001). Epidemiology of fractures in England and Wales. Bone, 29 (6): 517-522.
- Vestergaard P, Rejnmark L, Mosekilde L (2006). Socioeconomic aspects of fractures within universal public healthcare: A nationwide case-control study from Denmark. Scandinavia Journal of Public Health, 34: 371-377.
- Wildner M, Casper W, Bergmann KE (1999). A secular trend in hip fracture incidence in East Germany. Osteoporosis International, 9 (2): 144-150.
- Williams W (1971). Principles of clustering. Annual Review of Ecology and Systematics, 2: 303-326.
- Woolf A, Dixon A (1988). Osteoporosis – A clinical guide. Lippincott Williams & Wilkins.

- Yan L, Zhou B, Prentice A, Wang X, Golden M (1999). Epidemiological Study Of Hip Fracture In Shenyang, People's Republic Of China. *Bone*, **24** (2): 151–155.
- Yoshimura T, Tohya T, Onoda C, Okamura H (2005). Poor nutrition in prepubertal Japanese children at the end of World War II suppressed bone development. *Maturitas*, **52** (1): 32-34.
- Zebaze R, Seeman E (2003). Epidemiology of hip and wrist fractures in Cameroon, Africa. *Osteoporosis International*, **14** (4): 301–305.
- Zielhuis G, Kiemeny L (2001). Social epidemiology? No way. *International Journal of Epidemiology*, **30**: 43-44.

ANEXO A

Bibliografia

- Arboleva L, Castro MA, Bartolome E, Gervas I, Vega R (1997). Epidemiología de la fractura osteoporótica de cadera en la provincia de Palencia. *Revista Clínica Española*, **197** (9): 611-617.
- Assunção RM, Barreto SM, Gerra HL, Sakurai E (1998). Mapas de taxas epidemiológicas: uma abordagem Bayesiana. *Caderno Saúde Pública*, **14** (4): 713-723.
- Barcelos C, Bastos FI (1996). Geoprocessamento, ambiente e saúde: uma união possível? *Caderno Saúde Pública*, **12** (3): 398-397.
- Barcelos C, Ramalho W (2002). Situação actual do geoprocessamento e da análise de dados espaciais em saúde no Brasil. *Informática Pública*, **4**: 221-230.
- Baudoin C, Fardellone P, Potard V, Sebert J (1993). Fractures of the proximal femur in Picardy, France, in 1987. *Osteoporosis International*, **3** (1): 43-49.
- Castiglione LH, Pina MF (2003). Elaborar mapas: A construção de um discurso sobre a espacialidade dos fenómenos. *Arquivos de Medicina*, **17**: 303-310.
- Castro F, Ribeiro A (2003). Low incidence of hip fractures in an equatorial area. *Osteoporosis International*, **14** (6): 496-499.
- Chevalley T, Herrmann FR, Delmi M, Stern R, Hoffmeyer P, Rapin C-H, Rizzoli R (2002). Evaluation of the age-adjusted of hip fractures between urban and rural areas: The difference is not related to the prevalence of institutions for the elderly. *Osteoporosis International*, **13** (2): 113-118.
- Coster A, Haberkamp M, Allolio B (1994). Hip fractures incidence rates in Germany compared to other countries. *Sozial-Und Praventivmedizin*, **39** (5): 287-292.
- Cravey AJ, Washburn S, Gesler WM, Arcury TA, Skelly AH (2001). Developing social-spatial knowledge networks: A qualitative methodology for chronic disease prevention. *Social Science and Medicine*, **52** (11): 1763-1775.
- Demeter S, Leslie W, Lix L, MacWilliam L, Finlayson G, Reed M (2007). The effect of socioeconomic status on bone density testing in a public health-care system. *Osteoporosis International*, **18** (2): 153-158.
- Dias A, Ferreira F., Quintal A, Afonso C, Vaz CL, Vaz AL (1990). Epidemiologia e custos das fracturas osteoporóticas em Portugal. *Revista Portuguesa de Reumatologia*, **1**: 26-35.
- Díez A, Puig J, Martínez MT, Díez JL, Aubia J, Vivancos J (1989). Epidemiology of fractures of the proximal femur associated with osteoporosis in Barcelona, Spain. *Calcified Tissue International*, **44** (6): 382-386.
- Donaldson L, Reckless I, Scholes S, Mindell J, Shelton N (2008). The epidemiology of fractures in England. *Journal of Epidemiology and Community Health*, **62** (2): 174-180.
- El Maghraoui A, Koumba B, Jroundi I, Achemlal L, Bezza A, Tazi M (2005). Epidemiology of hip fractures in 2002 in Rabat, Morocco. *Osteoporosis International*, **16** (6): 597-602.

- Falch JA, Aho H, Berglund K, D ppe H, Finsen V, Hagstr m I, J rvinen M, Johnell O, Kaastad TS, Lauritzen JB, Lund B, L thje P, L fman O, Mellstr m D, Mowinckel P, Rossvoll I, Sernbo I, Wits  E, Zetterberg C (1995). Hip fractures in Nordic cities: difference in incidence. *Annals Chirurgicae and Gynologicae*, **84**: 286-290.
- Hagino H, Katagiri H, Okano T, Yamamoto K, Teshima R (2005). Increasing incidence of hip fracture in Tottori Prefecture, Japan: Trend from 1986 to 2001. *Osteoporosis International*, **16** (12): 1963-1968.
- Hashefi M, Burke M, Dempster D, Field E, Sebba A (2005) Practical guide to preventing, recognizing, and managing osteoporosis. *George Washington University Medical Center*.
- Heikkinen T, Willig R, H nninen A, Koskinen K, Mannism ki P, Alavaikko A, Von Und Zu Fraunberg M, Jalovaara P (2004) Hip fractures in Finland: a comparison of patient characteristics and outcomes in six hospitals. *Scandinavian Journal of Surgery*, **93**: 234-240.
- Herrera A, Martinez A, Ferrandez L, Gil E, Moreno A (2006) Epidemiology of osteoporotic hip fractures in Spain. *International Orthopaedics*, **30** (1): 11-14.
- Johnell O, Kanis J (2005). Epidemiology of osteoporotic fractures. *Osteoporosis International*, **16** (2): S3-7.
- Johnell O, Gullberg B, Kanis J, Allander E, Elffors L, Dequeker J, Dilsen G, Gennari C, Vaz A., Lyritis G, Mazzuoli G, Miravet L, Passeri M, Cano R, Rapado A, Ribot C (1995). Risk factors for hip fracture in European women: the MEDOS study. *Journal of Bone and Mineral Research*, **10** (11): 1802-1815.
- Kanis J (2002). International variations in hip fracture probabilities: implications for risk assessment. *Journal of Bone and Mineral Research*, **17**: 1237-44.
- Kanis J, Johnell O, Gullberg B, Allander E, Elffors L, Ranstam J, Dequeker J, Dilsen G, Gennari C, Vaz A, Lyritis G, Mazzuoli G, Miravet L, Passeri M, Perez R, Rapado A, Ribot C (1999). Risk factor for hip fracture in men from southern Europe: the MEDOS study. *Osteoporosis International*, **9** (1): 45-54.
- Karagas MR, Lu-Yao G, Barret J, Beach M, Baron A (1996). Heterogeneity of Hip Fracture: Age, Race, Sex, and Geographic Patterns of Femoral Neck and Trochanteric Fractures among the US Elderly. *American Journal of Epidemiology*, **143** (7): 677-682
- Kirke PN, Sutton M., Burke H, Daly L (2002). Outcome of hip fracture in older Irish women: a 2-year follow-up of subjects in a case-control study. *Injury*, **33**(5): 387-391.
- Krieger N (2001). A glossary for social epidemiology. *Journal of Epidemiology and Community Health*, **55**: 693-700.
- Lau E, Lee J, Suriwongpaisal P, Saw S, Das De S , Khir A, Sambrook P (2001). The incidence of hip fracture in four Asian countries: the Asian Osteoporosis Study (AOS). *Osteoporosis International*, **12**: 239-243.

- Ling X, Almin L, Xile Z, Cummings SR (1996). Very low rates of hip fracture in Beijing, People's Republic Of China - the Beijing osteoporosis project. *American Journal of Epidemiology*, **144**: 901-7.
- Leslie WD, Caetano PA, Roe EB (2005). The impact of hip sub region reference data on osteoporosis diagnosis. *Osteoporosis International*, **16**: 1669-1674.
- Levi N (1997). Incidence of hip fractures - At last steady state in the Nordic countries? *European Journal of Orthopaedic Surgery & Traumatology*, **7**: 251-253.
- Levy P, Levy E, Audran M, Cohen-Solal M, Fardellone P, Le Parc JM (2002). The cost of osteoporosis in men: The French situation. *Bone*, **30** (4) 631-636.
- Lüthje P (1991). The incidence of hip fracture in Finland in the year 2000. *Acta Orthopaedica Scandinavica*, **62**: 35-37.
- Lüthje P, Peltonen A, Nurmi I, Kataja M, Santavirta S (1995). No differences in the incidences of old people's hip fractures between urban and rural populations--a comparative study in two Finnish health care regions in 1989. *Gerontology*, **41** (1): 39-44.
- Lüthje P, Sanavirta S, Nurmi I, Honkanen R, Heliövaara M (1993). Increasing incidence of hip fractures in Finland. *Archives of orthopaedic and trauma surgery* **112**: 280-282.
- Maravic M, Le Bihan C, Landais P, Fardellone P (2005). Incidence and cost of osteoporotic fractures in France during 2001. A methodological approach by the national hospital database. *Osteoporosis International*, **16** (12): 1475-1480.
- Mazzouli GF, Gennari C, Passeri M, Acca M, Camporeale A, Piolo G (1993). Hip fracture in Italy – Epidemiology and preventive efficacy of bone-active drugs. *Bone*, **14**: S81-S84.
- Mazzuoli GF, Gennari C, Passeri M, Cali FS, Acca M, Camporeale A, Pioli G, Pedrazzoni M (1993). Incidence of hip fractures: an Italian survey. *Osteoporosis International*, **3**: 58-59.
- McColl A, Roderick P, Cooper C (1998). Hip fracture incidence and mortality in an English region: a study using routine National Health Service data. *Journal of Public Health Medicine*, **20** (2): 196-205.
- McGuigan F, Murray L, Gallagher A, Davey-Smith G, Noville CE, Van't HR, Boreham C, Ralston S (2002). Genetic and environmental determinants of peak bone mass in young men and women. *Journal of Bone and Mineral Research*, **17** (7) 1273-9.
- Melton L, Atkinson E, Madhok R (1996). Downturn in hip fracture incidence. *Public Health Reports*, **111**: 146-150.
- Memon A, Pospula W, Tanatawy A, Abdul-Ghafar S, Suresh A, Al-Rowaih A (1998). Incidence Of Hip Fracture In Kuwait. *International journal of epidemiology*, **27**: 860-865.
- Mikhailov EE, Benevolenskaya LI, Ershova OB, Bobylev VY (1995). Hip fracture epidemiology in age groups of high osteoporosis risk. *Terapevticheskii Arkhiv*, **67** (10): 39-42.

- Moayyeri A, Soltani A, Larijani B, Naghavi M, Alaeddini F, Abolhassani F (2006). Epidemiology of hip fracture in Iran: results from the Iranian multicenter study on accidental injuries. *Osteoporosis International*, **17** (8): 1252-7125.
- Morales-Torres J, Gutierrez-Urena S (2004). The burden of osteoporosis in Latin America. *Osteoporosis International*, **15**: 625-632.
- Morosano M, Masoni A, Sanchez A (2005). Incidence of hip fractures in the city of Rosario, Argentina. *Osteoporosis International*, **16** (11): 1339-1344.
- Nydegger V, Rizzoli R, Rapin C, Vasey H, Bonjour J (1991). Epidemiology of fractures of the proximal femur in Geneva - incidence, clinical and social aspects. *Osteoporosis International*, **2** (1): 42-47.
- Orwoll ES (1999). Osteoporosis in Men - The effects of gender on skeletal health. Academic Press.
- Pedrazzoni M, Alfano FS, Gatti C, Nigro E, Mazzola M, Ulietti V, Passeri M (1994). Indagine epidemiologica sulle fratture del collo del femore nella provincia di Parma (1980-1992). *Minerva Medica*, **85** (7): 379-386.
- Pina MF (2003^b). Representação Cartográfica de Dados Espaciais. *Arquivos de Medicina*, **17**(5).
- Reimers, A, Laflamme L (2007). Hip fractures among the elderly: personal and contextual social factors that matter. *Journal of Trauma-Injury Infection & Critical Care*, **62** (2): 365-369.
- Schwartz AV, Kelsey JL, Maggi S, Tuttleman M, Ho SC, Jonsson PV, Poor G, Castro JA, Xu L, Matkin CC, Nelson LM, Heyse SP (1999). International variation in the incidence of hip fractures: cross-national project on osteoporosis for the World Health Organization Program for Research on Aging. *Osteoporosis International*, **9** (3): 242-253.
- Sosa M, Segarra M, Hernández D, Ganzález A, Liminana J, Betancor P (1983). Epidemiology of proximal femoral fracture in Gran Canaria (Canary Islands). *Age and ageing*, **22** (4): 285-288.
- US Department of Health and Human Services (2004). Bone Health and Osteoporosis: A Report of the Surgeon General. Rockville, cap. **8**: 186-217.
- Vestergaard P, Rejnmark L, Mosekilde L (2005). Osteoporosis is markedly underdiagnosed: a nationwide study from Denmark. *Osteoporosis International*, **16** (2): 1341-1341.
- Wang M, Dixon L (2006). Socioeconomic influences on bone health in postmenopausal women: findings from NHANES III, 1988-1994. *Osteoporosis International*, **17** (1): 91-98.
- Wildner M, Clark DE (2001). Hip fracture incidence in East and West Germany: reassessment ten years after unification. *Osteoporosis International*, **12** (2): 136-139.
- Woolf A (2006). The global perspective of osteoporosis : the ILAR lecture, June 2004, Berlin. *Clinical rheumatology*, **25** (5): 613-618.
- Zingmond D, Melton L, Silverman S (2004). Increasing hip fracture incidence in California Hispanics, 1983 to 2000. *Osteoporosis International*, **15** (8): 603-610.

ANEXO B

Matriz de Correlações das variáveis dos dados demográficos

	t_h65	t_m65	t_vh	t_vm	t_ref_hm	tx_a01	idep_i	idep_j	t_div	t_sd_hm	t_1c	t_3c	t_sc_c	t_sup_c
t_h65	1,000	,700	,695	,673	,659	,679	,667	-,488	-,108	-,124	,177	-,222	-,189	-,247
t_m65	,700	1,000	,522	,687	,592	,505	,578	-,538	-,150	-,131	,151	-,198	-,037	-,059
t_vh	,695	,522	1,000	,896	,931	,833	,930	-,543	,301	-,187	,484	-,510	-,518	-,456
t_vm	,673	,687	,896	1,000	,944	,797	,930	-,654	,213	-,162	,488	-,476	-,440	-,387
t_ref_hm	,659	,592	,931	,944	1,000	,846	,961	-,613	,258	-,145	,468	-,501	-,483	-,430
tx_a01	,679	,505	,833	,797	,846	1,000	,832	-,431	,084	-,115	,423	-,555	-,584	-,586
idep_i	,667	,578	,930	,930	,961	,832	1,000	-,623	,226	-,172	,430	-,447	-,428	-,395
idep_j	-,488	-,538	-,543	-,654	-,613	-,431	-,623	1,000	,058	,059	-,104	,045	-,107	-,078
t_div	-,108	-,150	,301	,213	,258	,084	,226	,058	1,000	-,344	,407	-,480	-,544	-,239
t_sd_hm	-,124	-,131	-,187	-,162	-,145	-,115	-,172	,059	-,344	1,000	-,147	,307	,256	,084
t_1c	,177	,151	,484	,488	,468	,423	,430	-,104	,407	-,147	1,000	-,616	-,748	-,721
t_3c	-,222	-,198	-,510	-,476	-,501	-,555	-,447	,045	-,480	,307	-,616	1,000	,798	,506
t_sc_c	-,189	-,037	-,518	-,440	-,483	-,584	-,428	-,107	-,544	,256	-,748	,798	1,000	,712
t_sup_c	-,247	-,059	-,456	-,387	-,430	-,586	-,395	-,078	-,239	,084	-,721	,506	,712	1,000
cnp_123	-,294	-,098	-,490	-,423	-,458	-,635	-,426	-,077	-,198	,089	-,703	,533	,733	,961
cnp_45	,128	,177	-,165	-,098	-,135	-,246	-,097	-,314	-,480	,194	-,618	,643	,797	,581
cnp_6_9	,116	-,044	,379	,300	,341	,508	,304	,215	,368	-,152	,746	-,654	-,863	-,878
iperc	-,272	-,085	-,552	-,477	-,527	-,644	-,473	-,035	-,437	,192	-,746	,688	,853	,914
niv_a	-,056	-,044	-,319	-,286	-,323	-,356	-,258	-,045	-,337	,062	-,617	,445	,598	,724
t_res	-,733	-,806	-,672	-,770	-,734	-,578	-,762	,817	,144	,087	-,087	,071	-,088	-,019
t_res_d	-,570	-,603	-,780	-,808	-,807	-,565	-,809	,665	-,519	,297	-,332	,376	,284	,141
t_crpa	-,439	-,638	-,300	-,454	-,386	-,230	-,353	,570	,148	-,033	,140	-,126	-,291	-,217
t_cfa	,086	-,007	,064	,092	,114	,220	,131	-,063	-,095	-,058	,168	-,119	-,180	-,259
t_scfb	,235	,022	,246	,170	,202	,350	,195	,033	,056	-,228	,230	-,399	-,422	-,281
t_scp	,227	,023	,073	,068	,067	,256	,100	-,034	-,163	,067	,004	-,020	-,102	-,141
t_scot	,247	,075	-,004	,094	,037	,205	,054	-,030	-,296	-,010	-,041	,061	-,028	-,133
t_ibd	-,405	-,175	-,494	-,440	-,486	-,642	-,456	,048	-,137	,189	-,453	,601	,665	,550
t_ac	-,225	-,243	-,100	-,167	-,147	-,238	-,188	,211	,550	-,298	,091	-,221	-,241	,047
t_ancl	,286	,210	,488	,458	,493	,478	,460	-,114	,588	-,359	,501	-,590	-,660	-,477
t_ancm	-,088	,064	-,278	-,239	-,265	-,253	-,241	-,060	-,600	,400	-,523	,550	,672	,481
tx_hm0_49	-,698	-,654	-,898	-,939	-,945	-,793	-,951	,786	-,172	,166	-,394	,409	,335	,293
tx_hm50_	,698	,654	,898	,939	,945	,793	,951	-,786	,172	-,166	,394	-,409	-,335	-,293

	niv_a	t_res	t_res_d	t_crpa	t_cfa	t_scfb	t_scp	t_scot	t_ibd	t_sup_c	cnp_123	cnp_45	cnp_6_9	iperc
t_h65	-,056	-,733	-,570	-,439	,086	,235	,227	,247	-,405	-,247	-,294	,128	,116	-,272
t_m65	-,044	-,806	-,603	-,638	-,007	,022	,023	,075	-,175	-,059	-,098	,177	-,044	-,085
t_vh	-,319	-,672	-,780	-,300	,064	,246	,073	-,004	-,494	-,456	-,490	-,165	,379	-,552
t_vm	-,286	-,770	-,808	-,454	,092	,170	,068	,094	-,440	-,387	-,423	-,098	,300	-,477
t_ref_hm	-,323	-,734	-,807	-,386	,114	,202	,067	,037	-,486	-,430	-,458	-,135	,341	-,527
tx_a01	-,356	-,578	-,565	-,230	,220	,350	,256	,205	-,642	-,586	-,635	-,246	,508	-,644
idep_i	-,258	-,762	-,809	-,353	,131	,195	,100	,054	-,456	-,395	-,426	-,097	,304	-,473
idep_j	-,045	,817	,665	,570	-,063	,033	-,034	-,030	,048	-,078	-,077	-,314	,215	-,035
t_div	-,337	,144	-,519	,148	-,095	,056	-,163	-,296	-,137	-,239	-,198	-,480	,368	-,437
t_sd_hm	,062	,087	,297	-,033	-,058	-,228	,067	-,010	,189	,084	,089	,194	-,152	,192
t_1c	-,617	-,087	-,332	,140	,168	,230	,004	-,041	-,453	-,721	-,703	-,618	,746	-,746
t_3c	,445	,071	,376	-,126	-,119	-,399	-,020	,061	,601	,506	,533	,643	-,654	,688
t_sc_c	,598	-,088	,284	-,291	-,180	-,422	-,102	-,028	,665	,712	,733	,797	-,863	,853
t_sup_c	,724	-,019	,141	-,217	-,259	-,281	-,141	-,133	,550	1,000	,961	,581	-,878	,914
cnp_123	,644	,003	,134	-,204	-,283	-,333	-,152	-,167	,612	,961	1,000	,594	-,905	,901
cnp_45	,511	-,356	,012	-,495	-,199	-,266	-,060	,085	,459	,581	,594	1,000	-,873	,683
cnp_6_9	-,640	,189	-,081	,391	,270	,339	,131	,070	-,606	-,878	-,905	-,873	1,000	-,891
iperc	,711	-,003	,287	-,197	-,253	-,368	-,115	-,073	,655	,914	,901	,683	-,891	1,000
niv_a	1,000	-,066	,163	-,151	-,151	-,204	-,020	,063	,384	,724	,644	,511	-,640	,711
t_res	-,066	1,000	,768	,698	-,026	,046	-,054	-,068	,087	-,019	,003	-,356	,189	-,003
t_res_d	,163	,768	1,000	,504	,031	-,004	,048	,125	,175	,141	,134	,012	-,081	,287
t_crpa	-,151	,698	,504	1,000	,201	,295	,183	,036	-,188	-,217	-,204	-,495	,391	-,197
t_cfa	-,151	-,026	,031	,201	1,000	,224	,227	,323	-,421	-,259	-,283	-,199	,270	-,253
t_scfb	-,204	,046	-,004	,295	,224	1,000	,217	,207	-,779	-,281	-,333	-,266	,339	-,368
t_scp	-,020	-,054	,048	,183	,227	,217	1,000	,356	-,353	-,141	-,152	-,060	,131	-,115
t_scot	,063	-,068	,125	,036	,323	,207	,356	1,000	-,296	-,133	-,167	,085	,070	-,073
t_ibd	,384	,087	,175	-,188	-,421	-,779	-,353	-,296	1,000	,550	,612	,459	-,606	,655
t_ac	-,123	,315	-,082	,223	-,121	,071	-,228	-,262	,034	,047	,105	-,280	,090	-,097
t_ancl	-,383	-,165	-,528	,001	,089	,252	,017	-,002	-,443	-,477	-,448	-,536	,544	-,622
t_ancm	,354	-,118	,287	-,278	-,112	-,345	,025	-,049	,417	,481	,450	,604	-,593	,589
tx_hm0_49	,199	,851	,849	,472	-,121	-,165	-,088	-,060	,387	,293	,321	,003	-,190	,380
tx_hm50_	-,199	-,851	-,849	-,472	,121	,165	,088	,060	-,387	-,293	-,321	-,003	,190	-,380

	t_ac	t_ancl	t_ancm	tx_hm0_49	tx_hm50
t_h65	-,225	,286	-,088	-,698	,698
t_m65	-,243	,210	,064	-,654	,654
t_vh	-,100	,488	-,278	-,898	,898
t_vm	-,167	,458	-,239	-,939	,939
t_ref_hm	-,147	,493	-,265	-,945	,945
tx_a01	-,238	,478	-,253	-,793	,793
idep_i	-,188	,460	-,241	-,951	,951
idep_j	,211	-,114	-,060	,786	-,786
t_div	,550	,588	-,600	-,172	,172
t_sd_hm	-,298	-,359	,400	,166	-,166
t_1c	,091	,501	-,523	-,394	,394
t_3c	-,221	-,590	,550	,409	-,409
t_sc_c	-,241	-,660	,672	,335	-,335
t_sup_c	,047	-,477	,481	,293	-,293
cnp_123	,105	-,448	,450	,321	-,321
cnp_45	-,280	-,536	,604	,003	-,003
cnp_6_9	,090	,544	-,593	-,190	,190
iperc	-,097	-,622	,589	,380	-,380
niv_a	-,123	-,383	,354	,199	-,199
t_res	,315	-,165	-,118	,851	-,851
t_res_d	-,082	-,528	,287	,849	-,849
t_crpa	,223	,001	-,278	,472	-,472
t_cfa	-,121	,089	-,112	-,121	,121
t_scfb	,071	,252	-,345	-,165	,165
t_scp	-,228	,017	,025	-,088	,088
t_scot	-,262	-,002	-,049	-,060	,060
t_ibd	,034	-,443	,417	,387	-,387
t_ac	1,000	,196	-,468	,185	-,185
t_ancl	,196	1,000	-,808	-,425	,425
t_ancm	-,468	-,808	1,000	,199	-,199
tx_hm0_49	,185	-,425	,199	1,000	-,1,000
tx_hm50_	-,185	,425	-,199	-,1,000	1,000

ANEXO C

Quadro de Aglomeração

Stage	Cluster Combined		Coefficients	Stage Cluster First Appears		Next Stage
	Cluster 1	Cluster 2		Cluster 1	Cluster 2	
1	253	260	,009	0	0	21
2	217	219	,019	0	0	41
3	254	262	,030	0	0	21
4	52	55	,044	0	0	51
5	127	150	,060	0	0	78
6	111	126	,076	0	0	18
7	103	138	,097	0	0	120
8	232	235	,122	0	0	24
9	79	94	,151	0	0	52
10	97	154	,180	0	0	80
11	189	190	,210	0	0	77
12	155	158	,241	0	0	55
13	147	148	,276	0	0	70
14	172	173	,316	0	0	117
15	57	84	,357	0	0	124
16	119	156	,398	0	0	138
17	47	77	,439	0	0	69
18	111	128	,481	6	0	89
19	165	166	,524	0	0	86
20	33	67	,567	0	0	137
21	253	254	,611	1	3	156
22	51	72	,657	0	0	134
23	131	157	,703	0	0	90
24	232	238	,749	8	0	63
25	40	48	,798	0	0	97
26	86	89	,850	0	0	51
27	256	268	,901	0	0	161
28	195	252	,953	0	0	219
29	29	71	1,006	0	0	58
30	54	61	1,058	0	0	137
31	120	125	1,113	0	0	79
32	181	188	1,170	0	0	50
33	198	247	1,231	0	0	66
34	83	91	1,292	0	0	67
35	151	160	1,356	0	0	123
36	205	234	1,419	0	0	131
37	257	267	1,484	0	0	91
38	18	24	1,551	0	0	164
39	113	145	1,621	0	0	105
40	122	152	1,691	0	0	132
41	210	217	1,765	0	2	189
42	56	63	1,839	0	0	73
43	6	7	1,915	0	0	101
44	107	121	1,993	0	0	225
45	220	237	2,070	0	0	154
46	28	59	2,148	0	0	152

47	261	266	2,228	0	0	200
48	226	249	2,310	0	0	129
49	161	170	2,392	0	0	117
50	181	223	2,477	32	0	196
51	52	86	2,562	4	26	152
52	69	79	2,648	0	9	94
53	178	229	2,734	0	0	82
54	27	38	2,821	0	0	136
55	153	155	2,908	0	12	207
56	264	274	2,995	0	0	200
57	265	278	3,082	0	0	198
58	29	45	3,172	29	0	175
59	185	201	3,262	0	0	118
60	270	275	3,353	0	0	156
61	202	211	3,443	0	0	139
62	102	136	3,534	0	0	185
63	203	232	3,625	0	24	196
64	137	159	3,718	0	0	80
65	187	194	3,812	0	0	128
66	198	222	3,910	33	0	169
67	37	83	4,016	0	34	193
68	200	215	4,124	0	0	142
69	42	47	4,231	0	17	141
70	106	147	4,339	0	13	132
71	183	230	4,449	0	0	163
72	191	193	4,559	0	0	146
73	56	74	4,672	42	0	134
74	271	272	4,786	0	0	111
75	98	109	4,900	0	0	168
76	225	228	5,015	0	0	211
77	189	207	5,132	11	0	154
78	123	127	5,253	0	5	158
79	112	120	5,374	0	31	158
80	97	137	5,495	10	64	190
81	81	96	5,617	0	0	122
82	178	250	5,742	53	0	153
83	70	90	5,875	0	0	121
84	239	240	6,009	0	0	150
85	206	216	6,143	0	0	142
86	165	177	6,278	19	0	187
87	162	176	6,413	0	0	144
88	163	174	6,550	0	0	144
89	111	135	6,689	18	0	186
90	130	131	6,832	0	23	127
91	257	276	6,977	37	0	145
92	204	233	7,124	0	0	139
93	258	263	7,279	0	0	112
94	62	69	7,436	0	52	124
95	186	199	7,594	0	0	226
96	231	242	7,754	0	0	163
97	35	40	7,915	0	25	175

98	115	129	8,076	0	0	186
99	218	245	8,238	0	0	167
100	108	110	8,403	0	0	194
101	6	20	8,572	43	0	147
102	53	58	8,742	0	0	178
103	31	34	8,913	0	0	166
104	78	92	9,088	0	0	183
105	105	113	9,266	0	39	140
106	244	251	9,450	0	0	167
107	66	75	9,634	0	0	197
108	50	68	9,821	0	0	149
109	26	36	10,008	0	0	136
110	1	12	10,205	0	0	235
111	271	273	10,403	74	0	161
112	258	277	10,602	93	0	212
113	243	248	10,802	0	0	131
114	182	241	11,017	0	0	202
115	9	10	11,235	0	0	176
116	32	73	11,453	0	0	195
117	161	172	11,675	49	14	222
118	185	212	11,899	59	0	209
119	169	171	12,125	0	0	159
120	103	146	12,352	7	0	201
121	70	93	12,583	83	0	170
122	81	88	12,814	81	0	170
123	133	151	13,047	0	35	138
124	57	62	13,282	15	94	193
125	41	65	13,519	0	0	237
126	60	80	13,757	0	0	157
127	101	130	13,997	0	90	185
128	180	187	14,240	0	65	191
129	214	226	14,487	0	48	189
130	14	15	14,735	0	0	213
131	205	243	14,994	36	113	192
132	106	122	15,262	70	40	173
133	114	118	15,531	0	0	215
134	51	56	15,803	22	73	206
135	44	64	16,077	0	0	205
136	26	27	16,355	109	54	223
137	33	54	16,637	20	30	195
138	119	133	16,921	16	123	190
139	202	204	17,205	61	92	210
140	104	105	17,492	0	105	182
141	42	87	17,782	69	0	220
142	200	206	18,083	68	85	192
143	116	139	18,386	0	0	201
144	162	163	18,693	87	88	221
145	257	259	19,007	91	0	198
146	191	196	19,326	72	0	219
147	6	21	19,648	101	0	227
148	208	213	19,971	0	0	204

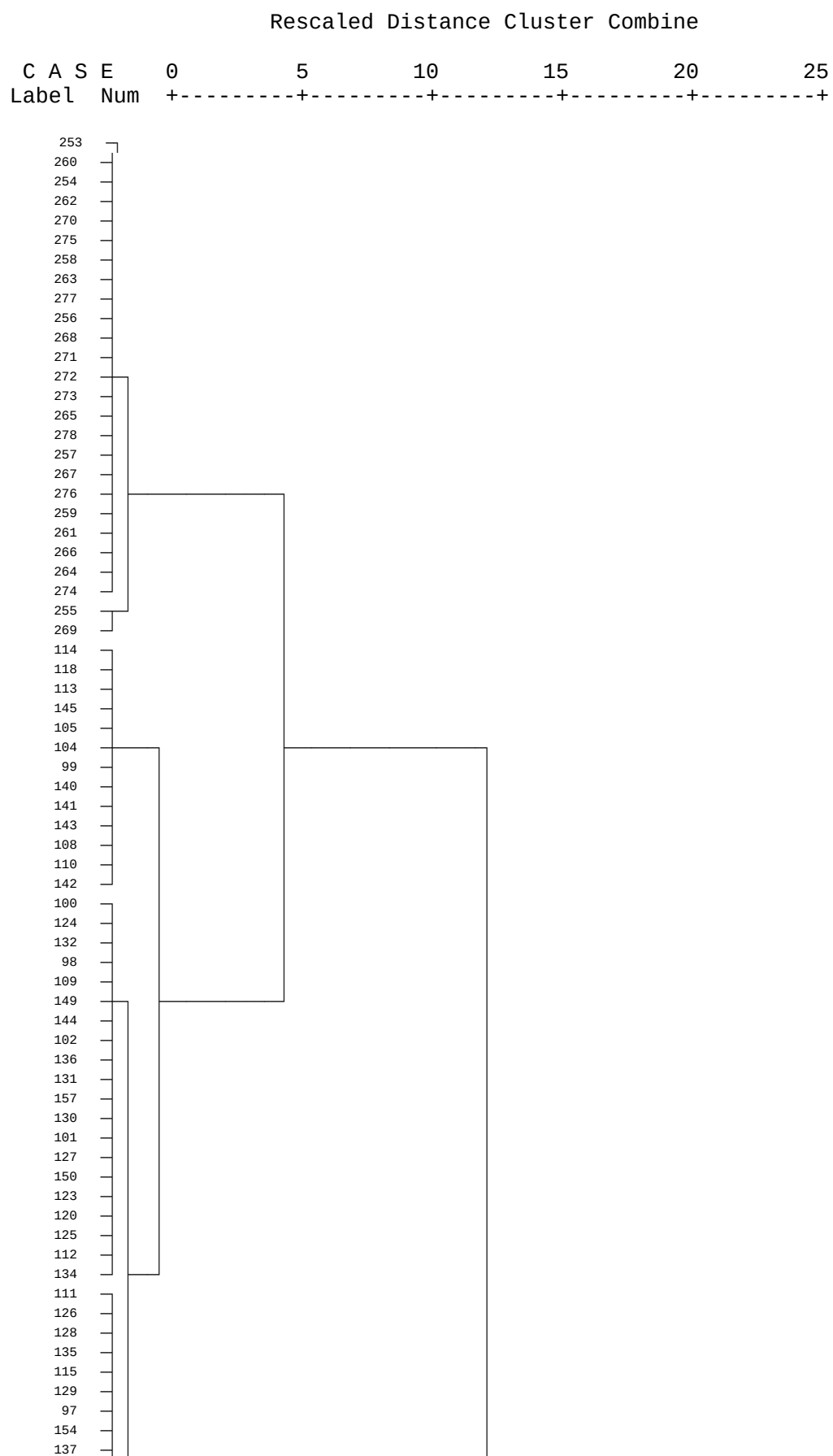
149	50	85	20,297	108	0	228
150	224	239	20,627	0	84	191
151	140	141	20,971	0	0	177
152	28	52	21,345	46	51	206
153	178	179	21,723	82	0	211
154	189	220	22,103	77	45	210
155	46	76	22,491	0	0	220
156	253	270	22,882	21	60	212
157	30	60	23,276	0	126	199
158	112	123	23,672	79	78	208
159	168	169	24,069	0	119	222
160	4	16	24,487	0	0	188
161	256	271	24,913	27	111	239
162	100	124	25,341	0	0	179
163	183	231	25,770	71	96	202
164	5	18	26,201	0	38	253
165	11	13	26,636	0	0	213
166	31	82	27,076	103	0	171
167	218	244	27,520	99	106	217
168	98	149	27,966	75	0	184
169	198	246	28,435	66	0	230
170	70	81	28,904	121	122	183
171	31	39	29,403	166	0	232
172	2	3	29,901	0	0	251
173	106	117	30,418	132	0	207
174	227	236	30,950	0	0	224
175	29	35	31,486	58	97	205
176	9	17	32,022	115	0	227
177	140	143	32,561	151	0	234
178	49	53	33,115	0	102	197
179	100	132	33,684	162	0	246
180	167	175	34,260	0	0	248
181	197	221	34,837	0	0	236
182	99	104	35,419	0	140	215
183	70	78	36,004	170	104	247
184	98	144	36,632	168	0	233
185	101	102	37,262	127	62	218
186	111	115	37,893	89	98	238
187	164	165	38,529	0	86	221
188	4	19	39,193	160	0	235
189	210	214	39,858	41	129	217
190	97	119	40,525	80	138	238
191	180	224	41,193	128	150	209
192	200	205	41,870	142	131	230
193	37	57	42,570	67	124	247
194	108	142	43,280	100	0	234
195	32	33	43,996	116	137	237
196	181	203	44,719	50	63	226
197	49	66	45,447	178	107	231
198	257	265	46,183	145	57	229
199	30	95	46,950	157	0	231

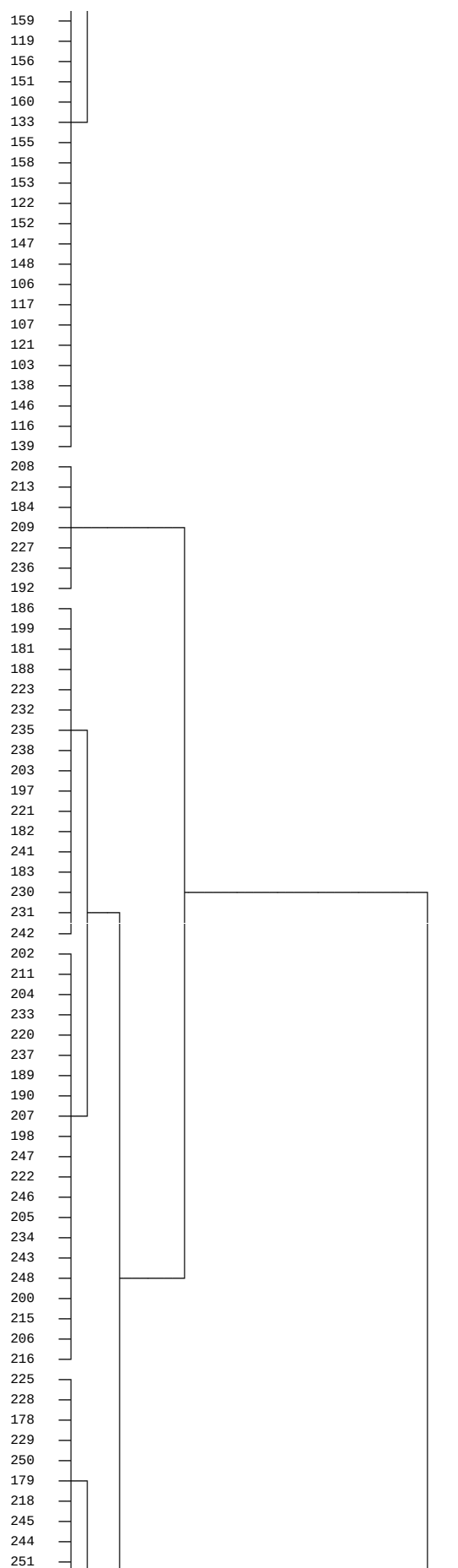
200	261	264	47,729	47	56	229
201	103	116	48,523	120	143	225
202	182	183	49,318	114	163	236
203	22	25	50,140	0	0	243
204	184	208	50,963	0	148	214
205	29	44	51,799	175	135	244
206	28	51	52,651	152	134	228
207	106	153	53,533	173	55	249
208	112	134	54,451	158	0	218
209	180	185	55,415	191	118	256
210	189	202	56,469	154	139	241
211	178	225	57,544	153	76	255
212	253	258	58,627	156	112	239
213	11	14	59,715	165	130	242
214	184	209	60,841	204	0	245
215	99	114	61,981	182	133	250
216	255	269	63,134	0	0	262
217	210	218	64,303	189	167	255
218	101	112	65,543	185	208	233
219	191	195	66,819	146	28	256
220	42	46	68,103	141	155	232
221	162	164	69,429	144	187	248
222	161	168	70,765	117	159	258
223	26	43	72,121	136	0	252
224	192	227	73,692	0	174	245
225	103	107	75,287	201	44	249
226	181	186	76,911	196	95	240
227	6	9	78,552	147	176	242
228	28	50	80,244	206	149	260
229	257	261	82,006	198	200	257
230	198	200	83,774	169	192	241
231	30	49	85,592	199	197	265
232	31	42	87,411	171	220	244
233	98	101	89,244	184	218	246
234	108	140	91,118	194	177	250
235	1	4	93,141	110	188	259
236	182	197	95,451	202	181	240
237	32	41	97,844	195	125	260
238	97	111	100,359	190	186	254
239	253	256	102,916	212	161	257
240	181	182	105,764	226	236	261
241	189	198	108,612	210	230	261
242	6	11	111,590	227	213	259
243	22	23	114,696	203	0	253
244	29	31	117,856	205	232	252
245	184	192	121,123	214	224	272
246	98	100	124,482	233	179	264
247	37	70	127,896	193	183	266
248	162	167	131,458	221	180	258
249	103	106	135,262	225	207	254
250	99	108	139,186	215	234	269

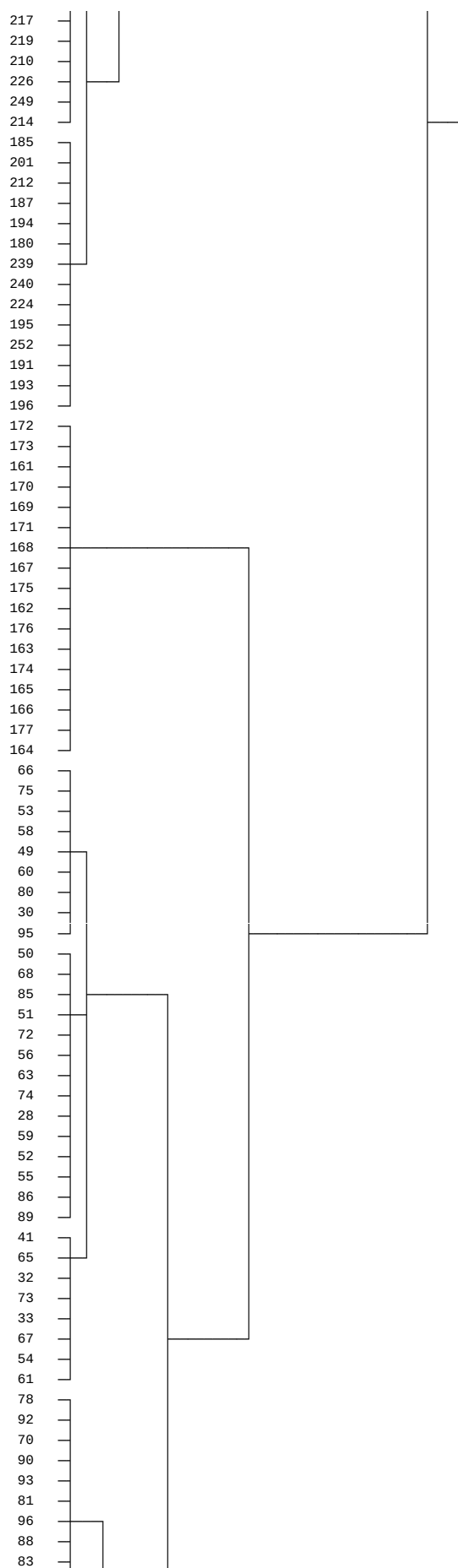
251	2	8	143,194	172	0	267
252	26	29	148,118	223	244	266
253	5	22	153,460	164	243	267
254	97	103	158,824	238	249	264
255	178	210	164,287	211	217	263
256	180	191	169,835	209	219	263
257	253	257	175,852	239	229	262
258	161	162	181,911	222	248	274
259	1	6	188,338	235	242	270
260	28	32	195,436	228	237	265
261	181	189	203,220	240	241	268
262	253	255	212,957	257	216	273
263	178	180	222,811	255	256	268
264	97	98	233,529	254	246	269
265	28	30	244,887	260	231	271
266	26	37	260,701	252	247	271
267	2	5	278,350	251	253	270
268	178	181	301,553	263	261	272
269	97	99	326,966	264	250	273
270	1	2	356,048	259	267	277
271	26	28	399,417	266	265	274
272	178	184	450,824	268	245	275
273	97	253	531,268	269	262	276
274	26	161	612,694	271	258	275
275	26	178	768,161	274	272	276
276	26	97	937,391	275	273	277
277	1	26	1108,000	270	276	0

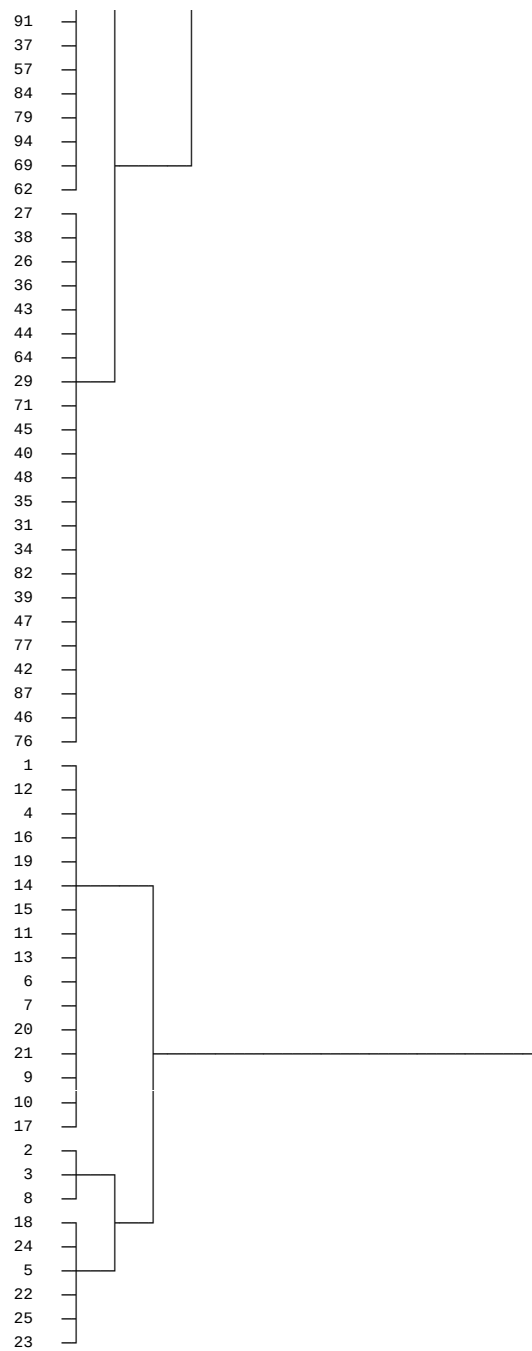
ANEXO D

Dendograma









ANEXO E

Resultados Da Análise Dos Dados Da Mortalidade Infantil

Tabela 1. Mortalidade infantil entre 1930 - 1939, por distrito

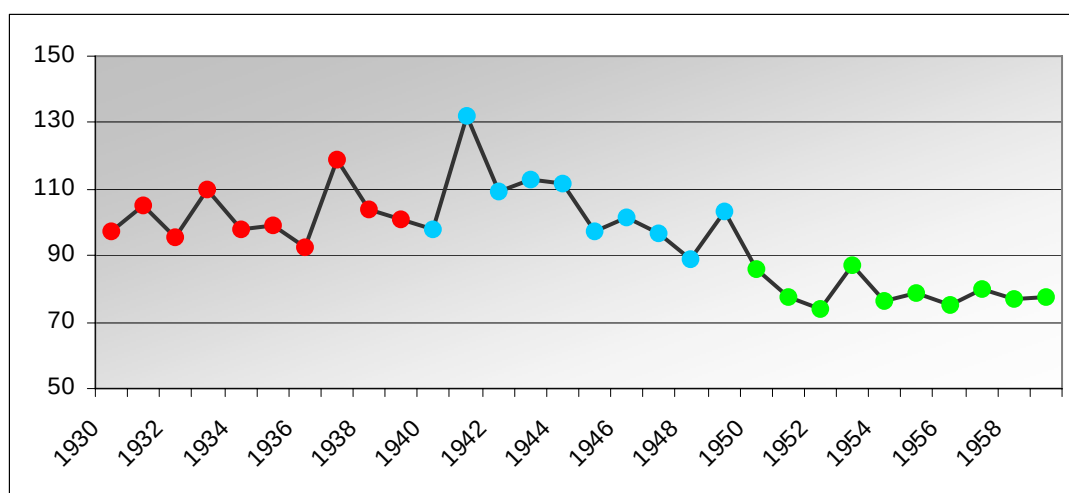
Distritos	Mortalidade infantil									
	1930	1931	1932	1933	1934	1935	1936	1937	1938	1939
AVEIRO	97,51	104,81	95,39	109,84	97,80	99,03	92,24	118,88	104,14	100,71
BEJA	131,78	123,37	106,33	126,74	129,24	115,67	115,37	130,05	127,38	118,95
BRAGA	121,35	122,03	115,30	136,43	117,74	138,21	114,25	139,94	122,97	128,69
BRAGANCA	136,67	139,76	141,30	148,59	149,65	148,19	142,06	154,88	116,74	126,35
CASTELO BRANCO	110,93	101,02	117,76	122,64	115,56	115,65	113,14	109,90	107,32	104,10
COIMBRA	93,17	90,64	90,46	98,32	86,61	98,37	81,93	103,61	84,23	88,42
EVORA	129,17	119,40	129,08	116,50	137,45	120,09	132,73	114,30	118,91	112,61
FARO	125,92	125,78	112,49	111,26	109,17	120,88	116,52	106,17	120,56	97,45
GUARDA	140,90	124,33	140,92	140,15	138,63	130,62	125,37	149,86	124,15	116,78
LEIRIA	105,92	108,63	105,33	93,63	102,23	103,00	105,08	100,43	106,20	97,26
LISBOA	167,47	159,94	165,00	151,96	156,70	151,84	158,42	207,48	151,77	129,44
PORTALEGRE	123,30	111,16	111,54	124,79	113,93	115,29	110,70	118,67	108,59	104,03
PORTO	142,19	158,18	157,56	162,68	164,79	163,25	155,01	182,01	153,72	157,66
SANTAREM	106,67	97,37	98,64	98,08	102,10	102,19	104,93	101,42	92,22	90,58
SETUBAL	163,50	151,75	160,15	160,22	150,12	154,10	161,71	142,36	163,42	129,17
VIANA DO CASTELO	113,17	104,55	99,73	109,38	104,85	107,10	91,84	102,96	103,92	101,92
VILA REAL	109,57	121,89	113,30	132,90	113,42	119,16	110,62	130,24	113,82	111,22
VISEU	102,87	96,69	97,50	102,64	93,52	98,85	96,50	113,46	96,58	100,76

Tabela 2. Mortalidade infantil entre 1940- 1949, por distrito

Distritos	Mortalidade infantil									
	1940	1941	1942	1943	1944	1945	1946	1947	1948	1949
AVEIRO	98,03	131,93	109,17	113,02	111,39	97,26	101,73	96,78	88,68	103,00
BEJA	102,02	151,68	125,50	107,63	116,96	140,03	114,83	97,06	89,86	107,04
BRAGA	122,61	177,54	144,98	159,57	126,62	122,14	137,09	114,75	105,65	125,97
BRAGANCA	130,03	165,02	153,60	141,98	120,20	125,17	129,46	103,36	111,61	136,33
CASTELO BRANCO	103,19	129,53	126,50	118,98	106,09	97,95	115,23	99,71	77,54	89,91
COIMBRA	84,83	102,67	88,16	92,52	86,83	84,73	80,78	71,07	60,10	76,04
EVORA	118,26	168,22	130,41	110,48	120,79	112,35	116,87	92,73	93,65	103,33
FARO	108,34	122,56	110,70	98,10	97,78	124,78	93,38	80,01	77,65	88,98
GUARDA	121,36	158,25	156,81	169,28	133,94	124,90	131,81	131,80	104,62	134,59
LEIRIA	98,42	94,91	86,65	97,01	100,44	87,38	86,57	77,13	73,05	71,31
LISBOA	164,75	155,21	136,73	123,25	127,88	119,68	109,90	100,73	95,66	105,79
PORTALEGRE	119,92	147,60	121,67	109,50	118,64	99,08	118,58	95,28	85,68	101,20
PORTO	155,68	207,07	156,94	156,27	153,02	126,87	152,47	140,33	133,89	154,42
SANTAREM	93,47	109,83	86,50	89,72	93,92	86,09	88,26	77,39	71,43	79,09
SETUBAL	151,33	179,26	148,34	134,20	134,37	128,43	114,01	110,43	101,07	108,24
VIANA DO CASTELO	109,00	138,12	88,55	109,28	109,69	101,09	101,16	97,98	85,68	105,24
VILA REAL	114,13	133,59	136,42	140,46	119,97	107,53	133,77	94,91	98,57	112,02
VISEU	94,02	112,33	115,16	130,41	111,80	90,59	101,86	92,89	83,67	93,31

Tabela 3. Mortalidade infantil entre 1950- 1959, por distrito

Distritos	Mortalidade infantil									
	1950	1951	1952	1953	1954	1955	1956	1957	1958	1959
AVEIRO	86,14	77,76	74,01	87,33	76,25	78,85	75,32	80,08	77,20	77,25
BEJA	87,39	82,35	78,45	79,68	76,20	75,66	92,49	75,04	73,03	84,62
BRAGA	108,60	104,30	117,59	114,91	100,67	104,74	104,25	113,62	100,17	107,85
BRAGANCA	199,43	78,05	100,52	89,91	73,10	87,36	91,25	94,37	89,08	93,41
CASTELO BRANCO	80,29	89,07	93,02	96,14	74,30	80,38	86,62	80,86	88,56	79,62
COIMBRA	124,82	67,77	76,50	72,30	69,90	65,42	64,58	73,48	69,24	73,03
EVORA	102,14	77,99	84,92	82,95	80,62	76,59	75,05	69,85	76,31	83,52
FARO	97,98	83,62	86,51	84,77	77,59	78,35	81,68	76,87	75,10	83,26
GUARDA	87,47	87,33	88,28	90,36	80,27	87,34	75,55	85,16	80,12	93,36
LEIRIA	97,69	84,78	89,00	90,08	80,70	81,14	77,41	75,90	72,85	74,53
LISBOA	95,56	33,91	72,85	75,37	66,66	68,88	73,58	62,00	61,13	62,97
PORTALEGRE	80,03	71,89	72,32	72,25	68,89	76,93	68,18	72,96	59,21	77,27
PORTO	94,83	124,05	135,27	138,85	126,08	133,18	114,64	126,78	118,33	125,17
SANTAREM	85,73	83,86	94,92	86,43	83,75	85,18	82,49	79,83	75,14	76,22
SETUBAL	112,58	89,09	81,18	82,54	76,08	80,10	104,72	74,35	75,47	88,95
VIANA DO CASTELO	99,17	100,74	110,22	101,65	92,31	133,16	89,80	105,21	101,93	104,56
VILA REAL	103,55	87,42	105,74	99,91	91,44	93,05	85,65	92,26	100,46	107,09
UIXEU	87,57	79,67	85,22	90,36	73,71	82,17	85,89	84,99	81,17	83,84

**Figura 1.** Gráfico da mortalidade infantil entre 1930 e 1959, referente ao distrito de Aveiro.

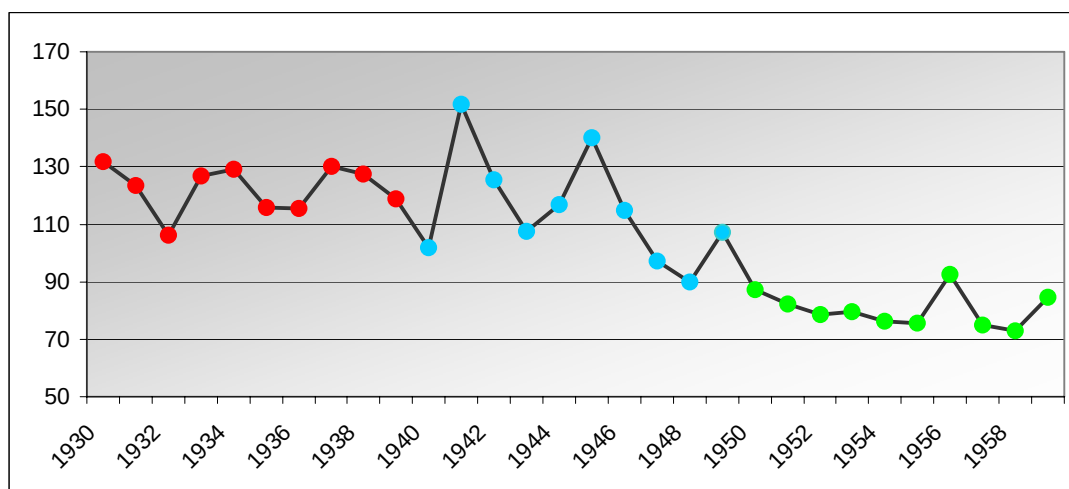


Figura 2. Gráfico da mortalidade infantil entre 1930 e 1959, referente ao distrito de Beja.

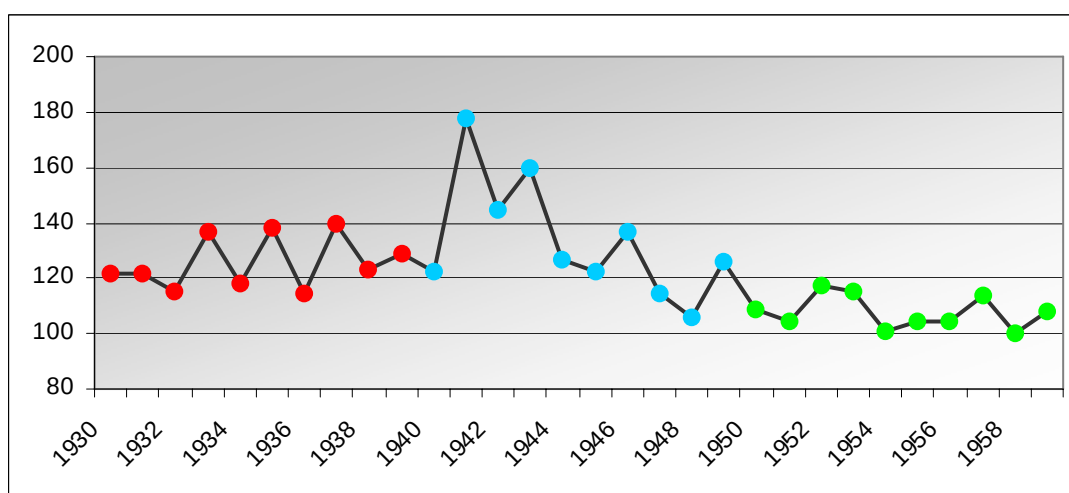


Figura 3. Gráfico da mortalidade infantil entre 1930 e 1959, referente ao distrito de Braga.

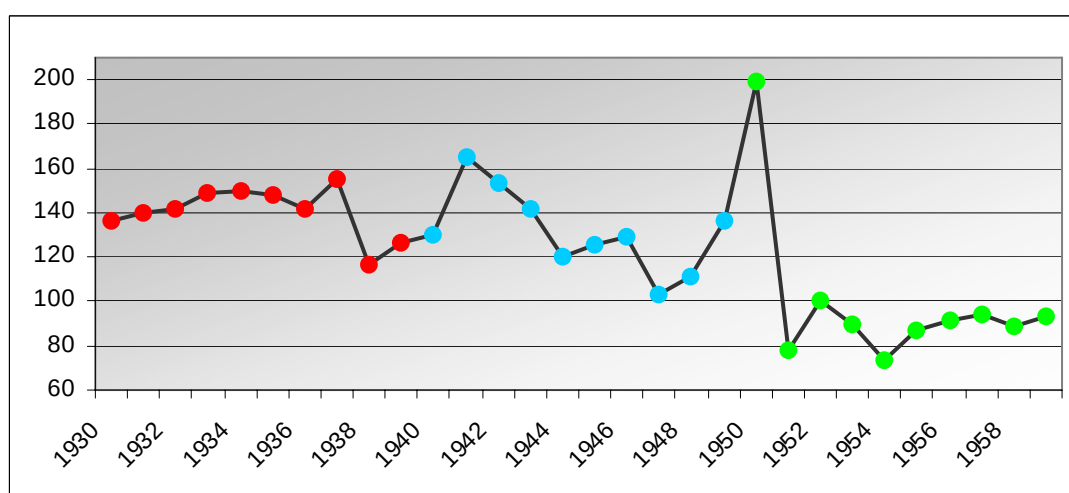


Figura 4. Gráfico da mortalidade infantil entre 1930 e 1959, referente ao distrito de Bragança.

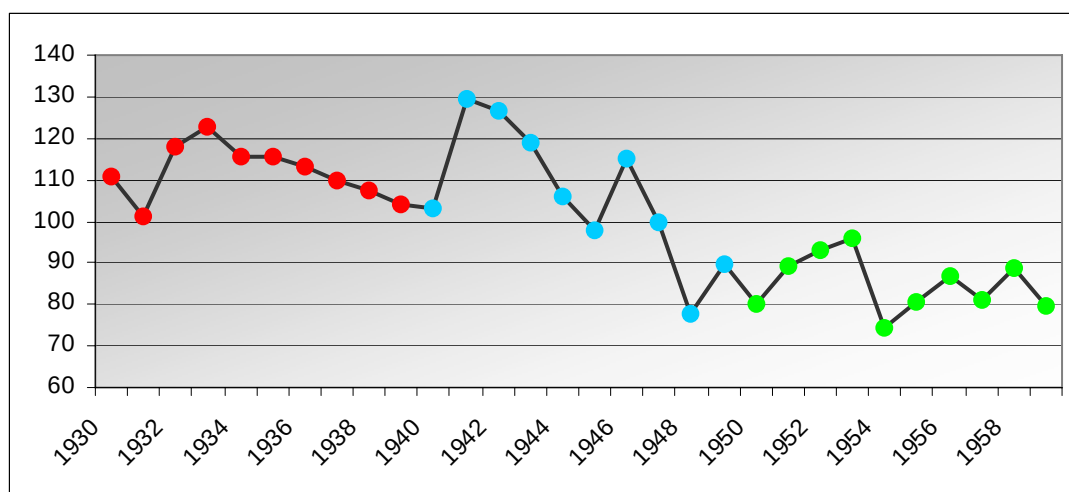


Figura 5. Gráfico da mortalidade infantil entre 1930 e 1959, referente ao distrito de Castelo Branco.

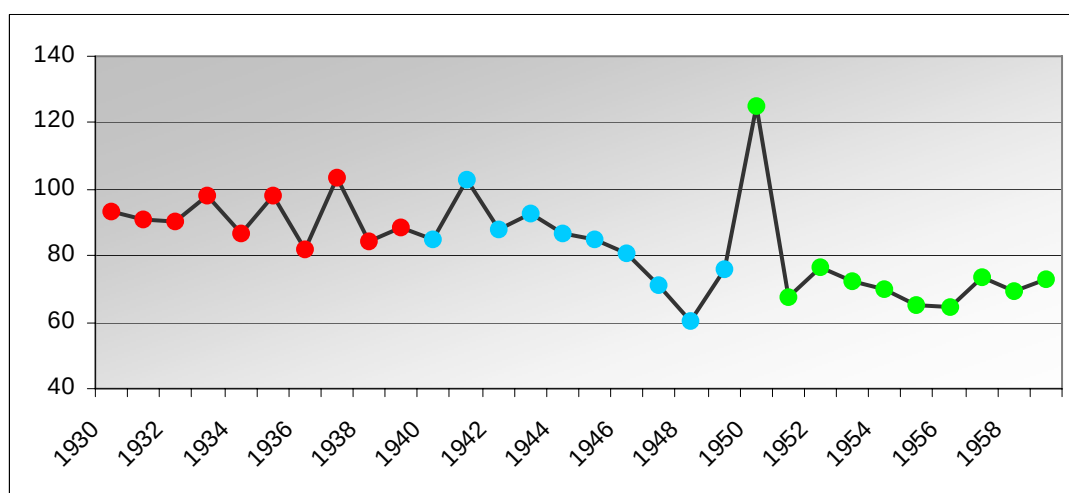


Figura 6. Gráfico da mortalidade infantil entre 1930 e 1959, referente ao distrito de Coimbra.

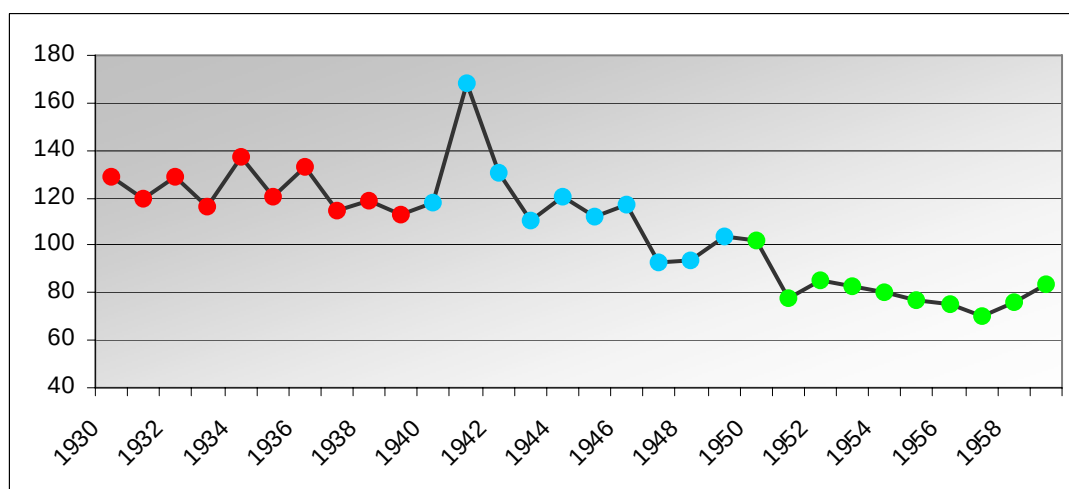


Figura 7. Gráfico da mortalidade infantil entre 1930 e 1959, referente ao distrito de Évora

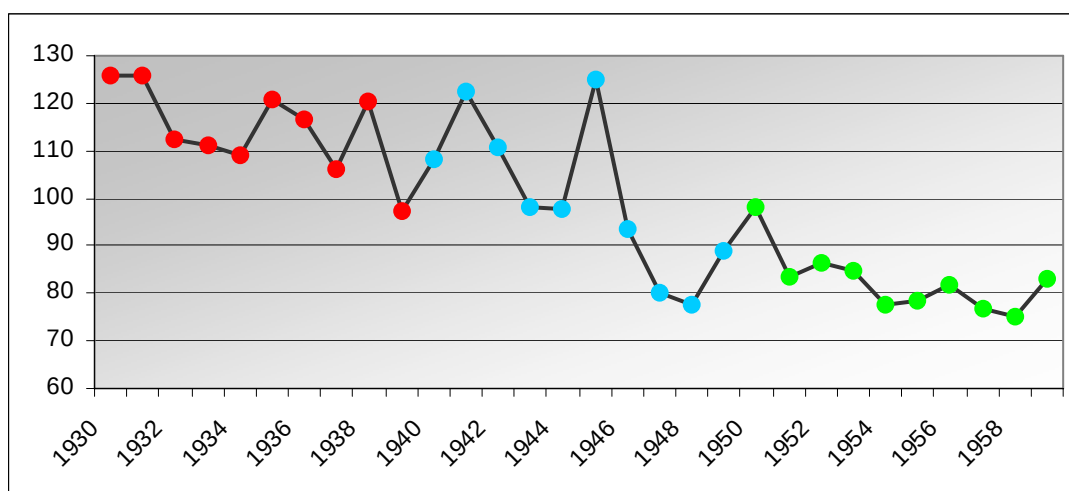


Figura 8. Gráfico da mortalidade infantil entre 1930 e 1959, referente ao distrito de Faro.

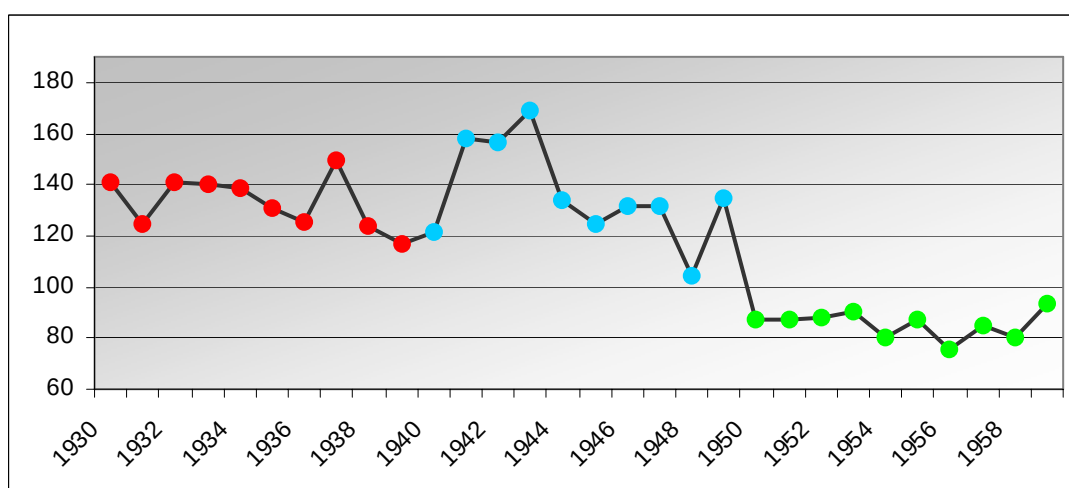


Figura 9. Gráfico da mortalidade infantil entre 1930 e 1959, referente ao distrito de Guarda.

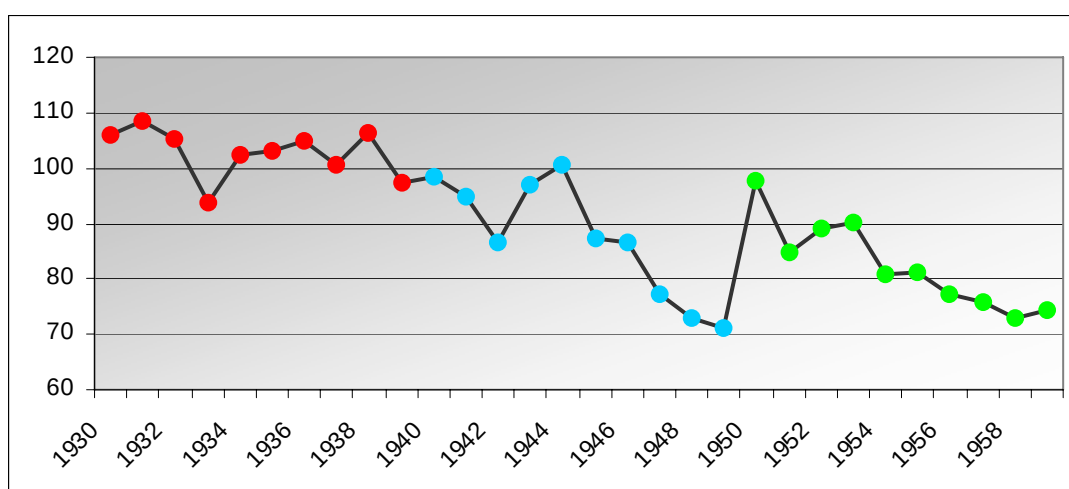


Figura 10. Gráfico da mortalidade infantil entre 1930 e 1959, referente ao distrito de Leiria.

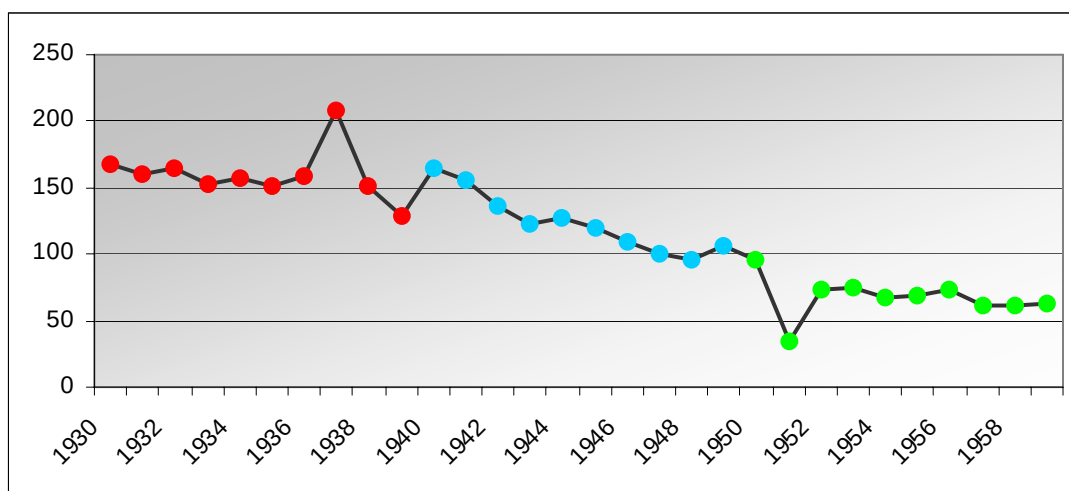


Figura 11. Gráfico da mortalidade infantil entre 1930 e 1959, referente ao distrito de Lisboa.

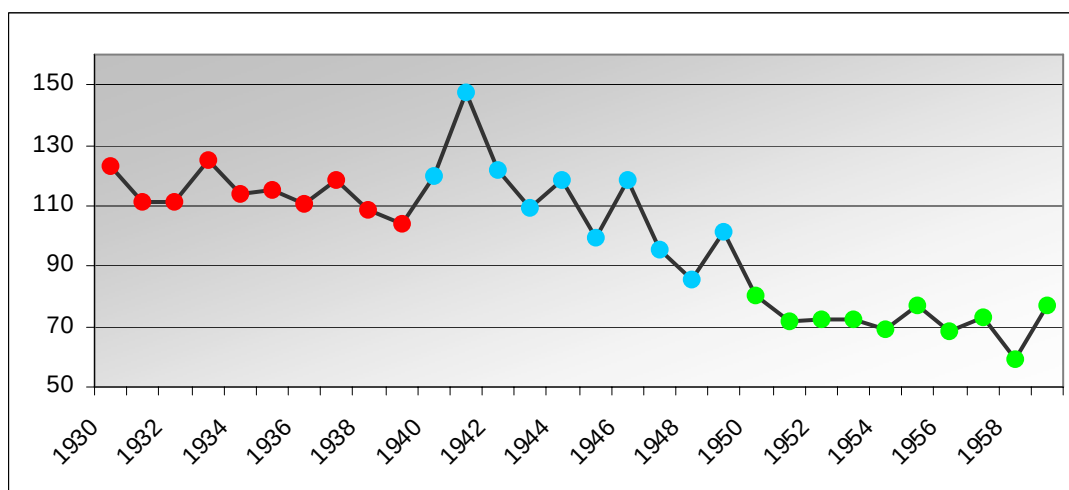


Figura 12. Gráfico da mortalidade infantil entre 1930 e 1959, referente ao distrito de Portalegre.

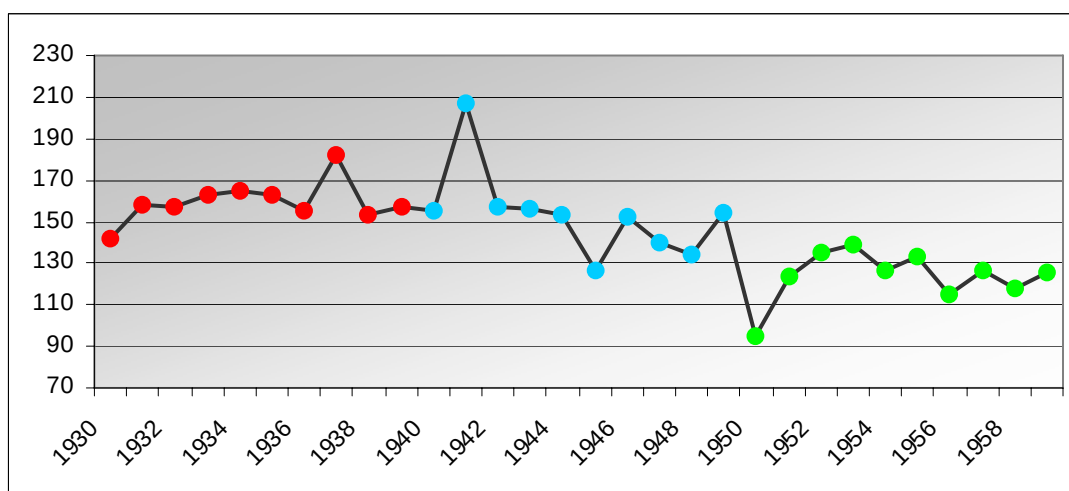


Figura 13. Gráfico da mortalidade infantil entre 1930 e 1959, referente ao distrito de Porto.

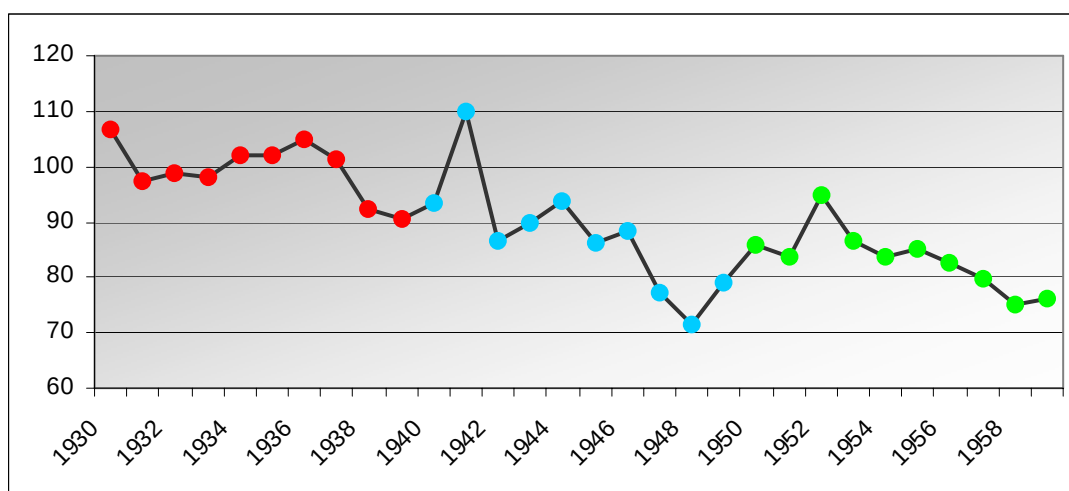


Figura 14. Gráfico da mortalidade infantil entre 1930 e 1959, referente ao distrito de Santarém.

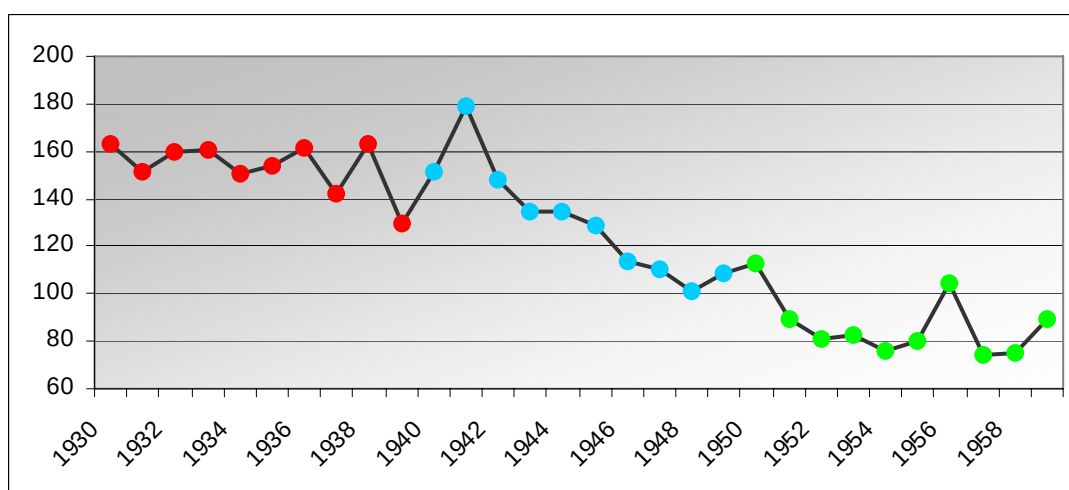


Figura 15. Gráfico da mortalidade infantil entre 1930 e 1959, referente ao distrito de Setúbal.

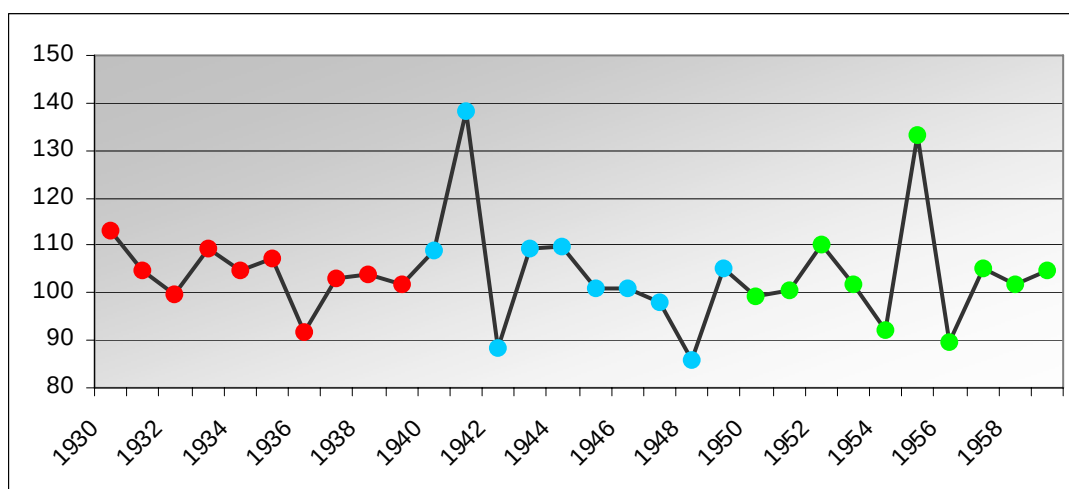


Figura 16. Gráfico da mortalidade infantil entre 1930 e 1959, referente ao distrito de Viana do Castelo.

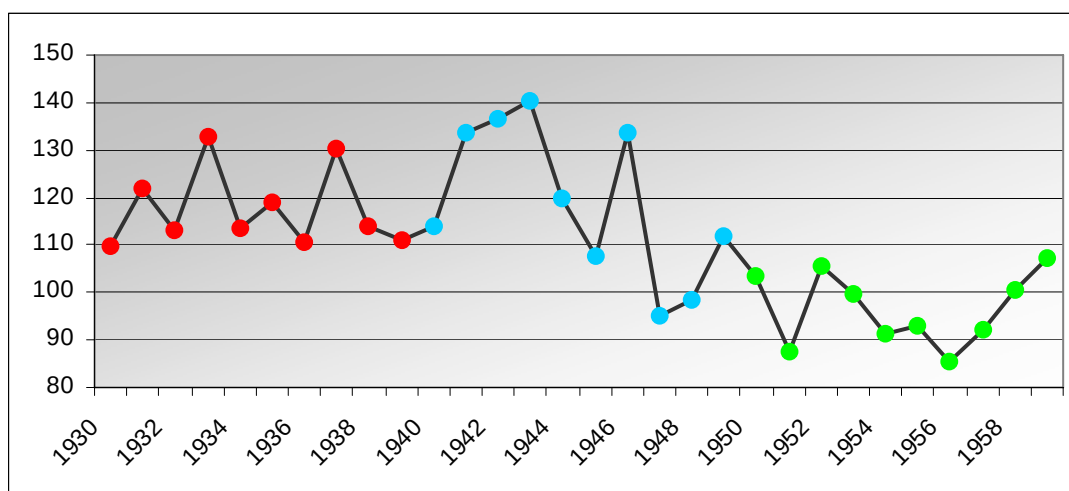


Figura 17. Gráfico da mortalidade infantil entre 1930 e 1959, referente ao distrito de Vila Real.

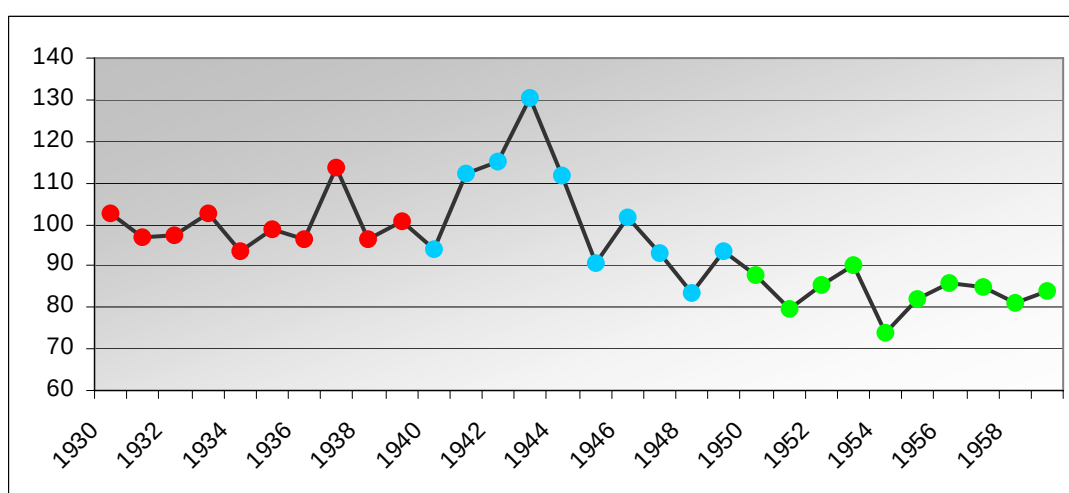


Figura 18. Gráfico da mortalidade infantil entre 1930 e 1959, referente ao distrito de Viseu.

