

**UNIVERSIDADE DE ÉVORA**

**Departamento de Matemática**

**MESTRADO EM MODELAÇÃO ESTATÍSTICA E ANÁLISE DE DADOS**

## **Caracterização do doente com VIH/ SIDA**

**Estudo estatístico ao impacto da infecção no Baixo Alentejo**

Dissertação de Mestrado sob a orientação do Professor Doutor Paulo Infante

Departamento de Matemática – Universidade de Évora

Ana Sofia Teixeira Marques Dionísio

Évora, 2010

## **Agradecimentos**

Embora as tarefas que constituem qualquer trabalho de investigação tenham uma forte componente de solidão, a presente tese de mestrado nunca teria sido possível sem o apoio das pessoas e entidades que em seguida refiro.

Ao Professor Doutor Paulo Infante, meu orientador, não só por nas suas aulas conseguir de uma forma tão rigorosa e simultaneamente descontraída cativar o interesse dos alunos nas matérias que lecciona, mas também, e principalmente, pela sua amizade, disponibilidade, partilha do saber e valiosas contribuições para a prossecução deste trabalho.

Um agradecimento muito especial à Professora Doutora Andreia Dionísio, por todo o apoio.

À Unidade Local de Saúde do Baixo Alentejo (ULSBA) pela disponibilização dos dados e incentivo à investigação.

Ao Dr. Telo Faria, médico responsável da equipa que acompanha os doentes infectados por VIH/SIDA no Hospital de Beja, pela disponibilidade que de imediato demonstrou assim que lhe falei dos objectivos desta tese, pela sua partilha de saber, bem como dos dados de que dispunha nessa altura, e ainda por acreditar.

Aos enfermeiros Carlos Albino e José Carlos Neves pelo importante contributo que também tiveram neste trabalho, pela amizade e apoio sempre demonstrados e por me esclarecem as imensas dúvidas iniciais sobre a doença e a forma como estes doentes são tratados.

À Dr.<sup>a</sup> Rosa Bento, Directora do Serviço de Patologia Clínica da ULSBA, por todo o apoio e valiosas contribuições em momentos chave do trabalho.

Uma palavra de agradecimento também à Dr.<sup>a</sup> Elizabete Silva, psicóloga na ULSBA, pela disponibilização dos dados outrora tão difíceis de obter sobre a infecção na região do Baixo Alentejo, e um agradecimento também à Dr.<sup>a</sup> Ana Moleiro, farmacêutica na ULSBA, pelos seus esclarecimentos sobre os medicamentos antiretrovirais.

À minha família. Aos meus pais, irmãs, cunhados, sobrinhos, e tia, pelo amor, apoio constante e por sempre estarem presentes.

A todos os amigos e amigas, pela sua inestimável amizade, algo tão importante para conseguir manter a sanidade mental necessária para continuar e continuar.

Por último, mas não menos importante, ao João. Pelo amor, que é tudo.

*“...uma doença transforma-se no mal do século porque cristaliza/simboliza a própria maneira como uma sociedade vive colectivamente o medo e a morte.*

*Nesse sentido, a doença é tão importante pelos seus efeitos imaginários como pelos seus efeitos reais.*

*A SIDA não escapa a esta regra: não tardou a sair do mundo médico para questionar os próprios fundamentos da nossa sociedade.*

*Presente na vida quotidiana, obriga-nos a reflectir e, eventualmente a modificar os nossos comportamentos.*

*Nenhuma doença, na época contemporânea, nos incitou tanto a interrogar-nos sobre a nossa identidade, os nossos valores, o nosso conceito de tolerância e de responsabilidade”*

(Montagnier, 1994)

# Índice

<b>Resumo</b>	<b>7</b>
<b>Abstract</b>	<b>8</b>
<b>1. Introdução</b>	<b>9</b>
<b>2. Enquadramento teórico sobre a infecção VIH/SIDA</b>	<b>13</b>
<b>2.1 – O que é o VIH/SIDA</b>	<b>13</b>
2.1.1 – Sistema Imunitário e Imunodeficiência	13
2.1.2 – Infecção por VIH e fase SIDA	14
2.1.3 - Transmissão do VIH	17
2.1.4 - Diagnóstico do VIH	17
2.1.5 - Fases da doença	18
2.1.6 - Tratamento	20
<b>2.2 – Dados actuais sobre a doença</b>	<b>22</b>
2.2.1 - Dados Globais	22
2.2.2 - Dados Europa e União Europeia	24
2.2.3 – Dados Portugal	25
2.2.4 – Dados Alentejo, Baixo Alentejo, Beja	30
<b>3. Metodologia e Alguns Aspectos Teóricos</b>	<b>32</b>
<b>3.1 – Obtenção dos dados e construção da base de dados</b>	<b>32</b>
<b>3.2 – Métodos de análise dos dados</b>	<b>33</b>
3.2.1 – Estatística descritiva e inferencial	33
3.2.2 – Análise de sobrevivência	35
3.2.2.1 – Conceitos básicos	36
3.2.2.2 – Alguns conceitos importantes	37
3.2.2.3 – Tipos de modelos	39
3.2.2.3.1– Modelos não-paramétricos	40
3.2.2.3.2– Modelos semi-paramétricos	42
3.2.2.3.3– Modelos paramétricos	44
3.2.3 – Modelos de fragilidade	50
3.2.4 – Análise dos resíduos e adequabilidade do modelo	50
<b>4. Análise exploratória dos dados</b>	<b>54</b>
<b>4.1 – Caracterização da base de dados</b>	<b>54</b>
<b>4.2 – Análise Descritiva e inferencial</b>	<b>56</b>
4.2.1 – Sexo, Idade e Estado Civil	56
4.2.2 – Sector de Actividade e Profissão	61
4.2.3 – Distrito e Concelho de Residência	62
4.2.4 – Estado Actual, Tipo de Vírus e Grupos/ Comportamentos de Risco	65
4.2.5 – Situação da infecção aquando da data de entrada no acompanhamento hospitalar	67
4.2.6 – Falecidos	70
4.2.7 – Terapêutica Antiretroviral (TAR): Início, Adesão versus Interrupção e Antiretrovirais Prescritos	75
4.2.8 – Carga Viral após TAR	83
4.2.9 – Patologias Verificadas e Comportamentos de Adicção	83
4.2.10 – Frequência das consultas de acompanhamento – Enfermagem e Psicologia	86
<b>4.3 – Análise Descritiva por Anos</b>	<b>88</b>
4.3.1 – Ano de Nascimento	88

4.3.2 – Ano de Diagnóstico	89
4.3.3 – Ano de entrada no acompanhamento hospitalar – Hospital de Beja	90
4.3.4 – Ano de Falecimento	91
4.3.5 – Ano de início da TAR	91
<b>5. Análise de Sobrevivência</b>	<b>93</b>
<b>5.1 – Dados Censurados</b>	<b>93</b>
<b>5.2 – Análise Univariada</b>	<b>94</b>
5.2.1 – Análise univariada a variáveis qualitativas – Estimativas de Kaplan-Meier	94
5.2.1.1- Sexo	95
5.2.1.2- Estado Civil	96
5.2.1.3- Recluso	97
5.2.1.4– Grupo/ Comportamento de Risco: Heterossexual	98
5.2.1.5 – Grupo/ Comportamento de Risco: Toxicodependente	99
5.2.1.6– Grupo/Comportamento de Risco: Transusão de Sangue	100
5.2.1.7 - Categoria Clínica	101
5.2.1.8 - Categoria de Contagem de Células CD4	102
5.2.1.9 - Tipo de Vírus	103
5.2.1.10 - Início de TAR (Terapêutica Antiretroviral)	104
5.2.1.11 - Interrupção da TAR	105
5.2.1.12 - Grupo Antiretrovirais	106
5.2.1.13 - Carga Viral	106
5.2.1.14 - Patologia: Tuberculose	108
5.2.1.15 - Patologia: Hepatite B	109
5.2.1.16 - Patologia: Hepatite C	110
5.2.1.17 - Patologia: Depressão Nervosa	111
5.2.1.18 - Alcoolismo	112
5.2.1.19 - Tabagismo	113
5.2.1.20 - Sector de Actividade	114
5.2.1.21 - Concelho de Residência	115
5.2.1.22 - Distrito de Residência: Beja	117
5.2.1.23 - País de Origem	117
5.2.1.24- País de Origem: Portugal	118
5.2.1.25 - Consulta de Enfermagem	119
5.2.1.26 - Consulta de Psicologia	120
5.2.2 – Análise Univariada – Variáveis categóricas – Modelo de regressão de Cox	121
5.2.2.1- Categoria de Contagem de Células CD4	121
5.2.2.2 - Concelho de Residência	122
5.2.2.3 - País de Origem	123
5.2.3 – Análise Univariada – Variáveis Contínuas – Modelo de regressão de Cox	125
5.2.3.1 - Idade de entrada	125
5.2.3.2 - Idade de diagnóstico	127
5.2.3.3- Idade no início da TAR	127
5.2.4 – Resumo Análise Univariada	128
<b>5.3 – Análise Multivariada</b>	<b>129</b>
5.3.1– Inclusão variáveis significativas	129
5.3.2 – Inclusão variáveis não significativas na análise univariada	134
<b>5.4 – Análise da forma funcional das variáveis contínuas</b>	<b>136</b>
5.4.1 - Idade de Entrada – IdadeEnt	136
5.4.2 - Idade no Diagnóstico – IdadeDiag	137
<b>5.5 – Avaliação de interações</b>	<b>138</b>
<b>5.6 – Validação de pressupostos e verificação da adequabilidade do modelo</b>	<b>139</b>

5.6.1 – Proporcionalidade de riscos	139
5.6.2 – Análise de Resíduos e Adequabilidade do modelo	141
<b>5.7 – Análise paramétrica</b>	<b>150</b>
<b>5.8 – Interpretação do modelo final</b>	<b>152</b>
<b>6. Conclusão</b>	<b>154</b>
<b>7. Referências Bibliográficas</b>	<b>159</b>
<b>8. Abreviaturas e siglas</b>	<b>163</b>
<b>9. Glossário</b>	<b>164</b>
<b>10. Anexos</b>	<b>167</b>

## **Caracterização do doente com VIH/ SIDA**

### **Estudo estatístico ao impacto da infecção no Baixo Alentejo**

#### **Resumo**

O VIH/SIDA tem provocado um enorme impacto na sociedade em geral. Não apenas ao nível da saúde e conseqüente redefinição de estratégias de prevenção e tratamento, mas também pelas questões éticas, sociais e culturais que veio levantar.

Não existindo quaisquer estudos sobre o tema na região assume particular interesse uma análise que caracterize o doente com VIH/SIDA no Baixo Alentejo.

Pretende-se caracterizar a população infectada na região, não apenas descritivamente como também inferir acerca destes indivíduos e verificar que aspectos específicos da sua conduta ou natureza possam estar a afectar o seu tempo de sobrevivência desde a data de diagnóstico da infecção até à ocorrência de morte.

A base de dados em análise é constituída por todos os doentes seguidos no Hospital de Beja entre os anos 2000 a 2009. Aos dados foram aplicadas diversas técnicas de estatística descritiva e inferencial e ainda técnicas específicas de Análise de Sobrevivência.

Da análise descritiva aos dados verifica-se que a população infectada e seguida na região do Baixo Alentejo em muito se assemelha à população infectada a nível nacional, na sua distribuição por sexo e idades, tipo de vírus, estadio da doença, comportamentos de risco e patologias apresentadas.

No estudo das variáveis que influenciam o tempo de sobrevivência dos indivíduos resultaram estatisticamente significativas e como factores de risco de morte: um nível de células CD4 resultante da contagem pelo valor mais baixo que o indivíduo obteve até à data inferior a 200 células por microlitro de sangue, a idade com que foi diagnosticado com a doença (indivíduos mais velhos apresentam maior risco), a existência das co-infecções Tuberculose e Hepatite C, e ainda o indivíduo sofrer de alcoolismo. Por outro lado, e como factores protectores resultaram a idade com que o indivíduo iniciou o acompanhamento hospitalar (quanto mais novo menor o risco), e ainda a frequência da consulta de acompanhamento de enfermagem no Hospital.

## **Characterization of patients with HIV/AIDS**

### **Statistical Study to the impact of infection on Baixo Alentejo**

#### **Abstract**

HIV/AIDS has caused a huge impact in society. Not only in the field of health and respective development of prevention and treatment strategies, but also for the ethical, social and cultural issues the disease had brought out.

Once the inexistence of any type of studies applied to this specific region of Alentejo in Portugal, the present study assumes particular interest in characterising HIV/AIDS patients in Baixo Alentejo.

The intend of this study is to characterize the infected population in Baixo Alentejo, not only in a descriptive way but also infer about these individuals and analyse which specific aspects on their behaviour or nature are affecting their survival time since the diagnostic date of the infection to the occurrence of death.

The database in analysis contains all patients followed/monitored in Hospital de Beja between years 2000 to 2009. To treat the data were considered descriptive and inferential statistical techniques and Survival Analysis.

From the descriptive analysis to data we can verify that the infected population monitored in Baixo Alentejos's region is very similar when compared to the infected population at a national level, especially in what concerns to its distribution by sex, age, type of virus, disease stage, risk behaviour and the presence of other infections or pathologies.

In the survival analysis, the variables that seem to influence individuals survival time in a significant statistically way were, as risk factors: having a level of CD4 cells less than 200 cells/*ul* resulting from the lowest value measured, the age with which the patient had been diagnosed with HIV/AIDS (in this case, older individuals present a higher risk), the existence of co-infections as Tuberculosis and Hepatitis C, and also alcoholism. On the other side resulted as statistically significant protective factors: the age with which the patient initiated hospital monitoring (in this case, the younger the individual lower the risk of death), and the attendance/frequency of nursing monitor sessions held at the hospital.



## 1. Introdução

Decorridos quase 30 anos após a publicação dos primeiros casos de morte devido a imunossupressão, em consequência de uma infecção que ainda hoje não está controlada, não existe actualmente qualquer perspectiva a curto ou médio prazo de um tratamento ou vacina que erradiquem a epidemia.

Dados recentes (OMS, 2007) referem que o Vírus da Imunodeficiência Humana (VIH) e Síndrome da Imunodeficiência Adquirida (SIDA) matam por dia cerca de 5700 pessoas no mundo, e Portugal tem das mais elevadas taxas de incidência de VIH/SIDA na Europa, sendo a causa de morte mais representativa no País entre os 30 e 39 anos (ACSS, 2007). Estima-se que existam cerca de 34.000 pessoas actualmente infectadas no nosso País (ONUSIDA, 2008), valor para o qual se considera uma taxa de infectados não diagnosticados de 30%, de acordo com a média da União Europeia (CNLCS, 2007).

Para além dos efeitos directos da infecção na saúde do indivíduo, o vírus tem um elevado impacto na sociedade em geral. Mais concretamente, pelo facto de afectar principalmente a população activa compromete o desenvolvimento da comunidade, pois aquele é o sector da população que assegura simultaneamente a produção de riqueza e o equilíbrio demográfico. Por outro lado, a despesa nacional com os tratamentos antiretrovíricos tem sido crescente e actualmente estima-se em 150 milhões de euros (ACSS, 2007). Para além desse elevado encargo financeiro, o alastramento da doença veio levantar outras questões não menos importantes tais como a falta de informação da população, o preconceito, a exclusão e discriminação social, étnica e sexual, num desrespeito pelos direitos humanos universais. Por tudo isto a infecção por VIH é doença de declaração obrigatória, consta até do actual Plano Nacional de Saúde e é considerada uma área de actuação prioritária do Governo.

A relevância de um estudo que caracterize o doente com VIH/SIDA parece óbvia, e ainda que este seja um contributo limitado, espera-se que ajude pelo menos a compreender melhor uma realidade que está a afectar a nossa sociedade.

A maior parte dos trabalhos de investigação publicados acerca do estudo do tempo de sobrevivência em indivíduos infectados com VIH/SIDA provém dos E.U.A., alguns mais recentes do Brasil, e por norma tendem a cingir-se à análise dos efeitos da terapia antiretroviral. Quanto aos métodos estatísticos mais utilizados são efectivamente a estimação via Kaplan-Meier e Modelo Clássico de Cox.

Ainda assim, existem alguns estudos que tentam avaliar o efeito de mais variáveis sobre o tempo de sobrevivência. É o caso de Lemp G. *et al* (1990) cuja aplicação da regressão de Cox permitiu concluir que as variáveis: diagnóstico inicial, idade e ano do diagnóstico eram significativamente explicativas do tempo de sobrevivência dos indivíduos estudados. O evento de interesse neste estudo era a ocorrência de morte.

Num outro estudo, Frieland *et al*, 1991, tentaram avaliar, entre outros, o efeito de algumas copatologias tendo concluído que os doentes com Sarcoma de Kaposi tendem a sobreviver mais tempo do que os que apresentem pneumonia *Pneumocystis carinii* ou neoplasias. Neste estudo resultaram também como factores explicativos na sobrevivência dos indivíduos o sexo, idade, raça e comportamentos de risco.

Morgan *et al* (2002) tentaram analisar a diferença no tempo mediano de sobrevivência entre indivíduos infectados com VIH-1 provenientes da África rural face a indivíduos infectados provenientes de países industrializados. Foram utilizados os métodos de Kaplan-Meier e concluiu-se que no período anterior ao desenvolvimento da terapia antiretroviral o tempo de sobrevivência verificado nos indivíduos infectados em África era similar ao verificado nos países industrializados. Esta conclusão espelha a desigualdade existente no acesso à terapia antiretroviral nos países desenvolvidos face ao resto do mundo. Neste âmbito, um outro estudo, Jahn *et al* (2008) veio a concluir, no caso específico do Malawi (país do sudeste africano), que a distribuição gratuita da terapia antiretroviral veio provocar um decréscimo na taxa de mortalidade relacionada com a SIDA de 33% entre os anos 2004 a 2006. No Brasil, Marins *et al* (2003) concluíram que desde que o tratamento antiretroviral triplo passou a ser de acesso gratuito e universal o tempo mediano de sobrevivência dos indivíduos subiu drasticamente.

Signorini *et al* (2005) investigaram a influência de factores sócio demográficos, clínico-profiláticos e terapêuticos na sobrevivência de indivíduos depois do diagnóstico de SIDA. De entre as conclusões apresentadas, resultaram como factores de risco o baixo nível de escolaridade, a admissão hospitalar e a falta de acompanhamento. Pneumonia *Pneumocystis*

*carinii* e *Toxoplasmosis prophylaxis* resultaram como factores protectores. A nível de terapêutica, o sistema triplo foi considerado o melhor regime.

Outra perspectiva na análise do tempo de vida do doente com VIH/SIDA foi apresentada por Kitahata *et al* (1996). Os autores tentaram analisar a relação entre o nível de experiência dos médicos no que respeita à SIDA e o tempo de sobrevivência dos seus pacientes com a doença. Concluíram que os doentes tratados pelos médicos com mais experiência apresentavam um risco de morte inferior em 31% face aos doentes tratados por médicos com menos experiência em casos de SIDA. Desta forma, o nível de experiência do médico que trata o doente resultou como uma variável significativa na explicação da sobrevivência dos doentes.

Em Portugal existem alguns estudos publicados sobre aplicação da Análise de Sobrevivência em indivíduos com VIH/SIDA, mas como já referido, tendem a tentar analisar maioritariamente o efeito da terapêutica antiretroviral.

Nesta tese assume um particular interesse o estudo da população com VIH/SIDA no Baixo Alentejo, pois o trabalho no Hospital de Beja, cuja área de influência é precisamente o Baixo Alentejo, possibilita um contacto mais directo com a forma como o doente é seguido nesta instituição da Unidade Local de Saúde permitindo o acesso a informação que se considera ser muito importante analisar, evidentemente com os devidos cuidados no respeito da confidencialidade dos dados.

Pretende-se acima de tudo caracterizar a população infectada na região, não só em termos descritivos através de uma análise inicial e exploratória dos dados, mas também inferir acerca destes indivíduos e detectar possíveis factores de risco e /ou protectores, e verificar que aspectos específicos da sua conduta ou natureza possam estar a afectar o seu tempo de sobrevivência, para além dos tratamentos a que são sujeitos. Considerar-se-á então para o efeito uma amostra constituída por todos os utentes seguidos no Hospital de Beja entre os anos 2000 a 2009, cujos dados se distribuem por variáveis do tipo qualitativo e quantitativo.

Para concretizar os objectivos de investigação anteriormente referidos, começar-se-á por apresentar uma breve contextualização da doença e sua problemática, bem como os dados actualmente existentes sobre o tema.

Seguidamente, e no que respeita às metodologias estatísticas utilizadas no tratamento e análise de dados, apresentar-se-á uma revisão da literatura publicada sobre as técnicas que se consideram ser as mais adequadas nesta investigação, nomeadamente a Análise de Sobrevida.

Estas serão aplicadas à amostra, donde se registarão os resultados obtidos e respectivas conclusões, apresentadas no final deste trabalho.

## **2. Enquadramento teórico sobre a infecção VIH/SIDA**

Neste capítulo será abordada a temática inerente à infecção VIH/ SIDA, como se define e manifesta a doença, como é transmitido o vírus, a sua evolução histórica, e o seu impacto na sociedade em geral. Serão ainda apresentados os dados actuais, a nível mundial, nacional e regional.

### **2.1 – O que é o VIH/SIDA**

#### **2.1.1 – Sistema Imunitário e Imunodeficiência**

O sistema imunitário é constituído por células e substâncias solúveis e é o responsável pela defesa do organismo de todos os agentes estranhos, contribuindo por isso de forma essencial para a condição de sobrevivência. A base do sistema imunitário é o complexo major de histocompatibilidade, único de um indivíduo para outro, e é este que ajuda a identificar o que é próprio e o que é alheio no organismo.

Os componentes major do sistema imunitário são os linfócitos B, células fagocitárias, linfócitos T e sistema complemento. As suas células mais importantes são os glóbulos brancos (dos quais os macrófagos, neutrófilos e linfócitos – B, T e células *natural killers* - são exemplo) e as substâncias solúveis que mais se destacam são os anticorpos ou imunoglobinas, as proteínas do sistema complemento e as citocinas ou mensageiras [1].

As três principais funções do sistema imunitário são o reconhecimento, a mobilização e o ataque. A resposta imune desencadeia-se com a correcta identificação e processamento de invasores ou ameaças ao organismo. Esses agentes estranhos podem ser antigénios (proteínas, hidratos de carbono, células cancerosas e toxinas) ou micróbios (bactérias, vírus, fungos, gérmes ou microorganismos), capazes de estimular aquela resposta [1].

Após o reconhecimento da ameaça (pelos macrófagos, e linfócitos B e T), os mecanismos de imunidade, inata ou aprendida, são accionados. Assim que o antigénio detectado é absorvido, são libertadas citocinas que ajudam na mobilização de mais células e mais potentes capazes de combatê-lo e eliminá-lo (processo que se denomina por quimiotaxia), assistindo-se então ao ataque propriamente dito e conseqüente eliminação ou isolamento dos agentes invasores.

Posteriormente, se o ataque tiver sido bem sucedido e estivermos perante um organismo saudável, a informação referente ao antigénio combatido fica registada e memorizada de forma a ser accionada em caso de invasão futura (e assim, se já existir memória nos anticorpos serão maioritariamente accionados os mecanismos de resposta secundária). No caso desse antigénio não ser já conhecido, serão accionados os mecanismos de resposta primária e informados os de resposta secundária, que começam a segregar os respectivos anticorpos [1].

Algumas pessoas apresentam um sistema imunitário deficiente na medida em que este não consegue desempenhar correctamente as suas funções de defesa, pelo que essas pessoas são consideradas imunodeprimidas.

A Imunodeficiência designa um conjunto de perturbações em que o sistema imunitário não é suficientemente forte para proteger o organismo contra agentes estranhos (Mata, 2006). Estes indivíduos apresentam por isso uma maior susceptibilidade à infecção, pelo que a sua condição de sobrevivência é frágil e encontra-se constantemente ameaçada.

A Imunodeficiência pode ainda ser do tipo congénita (primária), se apresentada logo à nascença e devido ao funcionamento anormal da medula óssea, ou adquirida (secundária), se o indivíduo a adquirir mais tarde. Este é o tipo mais frequente de imunodeficiência [2] e as suas causas a nível geral são inúmeras tais como deficiência nutricional, insuficiência renal, diabetes, transplante de órgãos, terapia cancerígena, terapia contra doenças autoimunes, infecções virais, doenças do sangue e cancro, queimaduras severas, cirrose hepática (que pode decorrer de alcoolismo), hepatite crónica, e envelhecimento normal, entre outros (Mata, 2006) [2].

### **2.1.2 – Infecção por VIH e fase SIDA**

Segundo investigações realizadas nesta área, a infecção causada pelo vírus da imunodeficiência humana (VIH) é uma doença provocada por um ou dois tipos de vírus que progressivamente destroem os glóbulos brancos a que chamamos linfócitos, causando a síndrome da imunodeficiência adquirida (SIDA) e outras doenças derivadas de uma imunidade deficiente [3].

A SIDA é o conjunto de sintomas e sinais resultantes de um dano específico do sistema imunológico, e constitui o estadio mais avançado da infecção por VIH. Nesta situação, o seu portador encontra-se sujeito a doenças oportunistas, isto é, doenças que em indivíduos com um

sistema imunitário eficiente não têm qualquer efeito ou (que) são pouco eficazes, mas no caso em que o indivíduo se encontra com um sistema imunitário debilitado mostram-se muito eficazes e potencialmente mortais (Nájera, 1997). São exemplo as infecções por vírus (ex. Citomegalovírus, Herpes simples), infecções por bactérias (ex. *Mycobacterium avium-intracelulare*, *Mycobacterium tuberculosis* complex, Salmonella), infecções por fungos (ex. Candidíase da boca e esófago, Pneumonia por *Pneumocystis carinii*, Criptococose, Histoplasmose, Coccidiomicose), infecções por parasitas (ex. Toxoplasmose, Criptosporidiose, Isosporidiose), neoplasias (ex. Linfomas, Sarcoma de Kaposi), e manifestações neurológicas (ex. demência). Estas, entre outras, são consideradas doenças infecciosas definidoras de SIDA, de acordo com os critérios do CDC (Miranda, 2003).

O VIH é um retrovírus, que infecta células do sistema imunitário de tal forma que este não o consegue eliminar. Ao penetrar na célula, o VIH transforma o seu código genético de ARN em ADN, o que é possível através de uma enzima chamada transcriptase reversa, que lhe permite replicar-se e destruir estas células. Para completar o seu ciclo de reprodução, o vírus utiliza ainda outras duas enzimas, a protease e a integrase [4].

O vírus reconhece uma proteína receptora, citocina chamada CD4 (essencial no sistema imunitário pois tem como função informar outras células sobre a necessidade de combater vírus) [4], presente nos linfócitos T4 e macrófagos, podendo ainda ter outros receptores como o CCR5, presente nos macrófagos e o CXCR4, presente em ambos os linfócitos T4 e macrófagos. Ao acoplar-se àquelas células através dos receptores, o vírus entra nelas e funde a sua membrana com a da célula. Nas primeiras semanas, a resposta imunitária efectuada pelos linfócitos T8 ao VIH é vigorosa. No entanto, os T8 estão sob controlo de citocinas produzidas pelos T4 que foram infectados, e as concentrações de linfócitos nos gânglios linfáticos devido àquela resposta vigorosa levam a que os virions sobreviventes infectem gradualmente mais e mais linfócitos, até que a resposta imune seja revertida. Além disso, as constantes mutações do ADN do vírus mudam a conformação das proteínas de superfície, o que dificulta continuamente o seu reconhecimento pelo sistema de defesa. Crê-se que o vírus cria diariamente dez milhões de novos virions, destruindo outro tanto de células CD4, que o sistema imunitário tenta repor, mas acaba por não aguentar o ritmo e esgotar-se (Balows *et al*, 1991). [4,5].

O sistema imunitário encontra-se por isso completamente comprometido com o VIH, desde que este ataca o organismo até ao momento da morte do indivíduo. De facto, o VIH é actualmente um dos vírus que causa mais mortes no mundo, o que em parte muito se deve ao seu elevado grau de transmissibilidade.

Cada vírion de VIH só tem um dos receptores, existindo por isso os tipos virion T-trópico, para o CCR5, ou o M-trópico, para o CXCR4. Normalmente são estes últimos que transmitem a infecção a outros indivíduos pois os macrófagos estão presentes em maior quantidade nas secreções genitais e no sangue e sobrevivem melhor. Por outro lado, os T-trópicos são menos infecciosos mas são aqueles que invadem os linfócitos, causando a sua diminuição no sistema imunitário e conduzindo o organismo à imunodeficiência, e portanto susceptível a outras infecções. É também sabido que os M-trópicos se multiplicam e rapidamente surgem virions mutantes do tipo T-trópico [5]. Um aspecto curioso é que os indivíduos que, por deficiência genética, não expressam CCR5, não adquirem o vírus mesmo se continuamente em risco. Para além destes, existem também casos em que os portadores de VIH não desenvolvem SIDA, não deixando no entanto de poder transmitir o vírus aos restantes. Pensa-se que estes indivíduos têm um sistema imunitário fortalecido e que terão contraído uma forma “debilitada” do vírus [6].

Segundo a teoria actualmente aceite, pensa-se que o VIH seja derivado do vírus da imunodeficiência símia. Existem dois tipos do vírus, que só se reproduzem nos humanos, o denominado VIH-1, o mais agressivo, cuja progressão da infecção assintomática para a SIDA é mais rápida, causando a SIDA típica e está presente em todo o mundo sendo por isso o mais predominante, e o VIH-2, que se transmite com menos facilidade e que causa uma doença semelhante, mais frequente na África Ocidental e também existente em Portugal (possivelmente devido à presença nos países outrora colonizados) (Miranda, 2003). O VIH-2 apresenta também um período mais prolongado entre a infecção e o aparecimento da doença (20 a 30 anos) (Pereira, 2008) [4].

Existem, pelo menos, nove subtipos do VIH-1, geneticamente diferentes, sendo que na América do Norte e Europa é mais predominante o subtipo B, na África predomina o subtipo C e na Ásia a forma combinada AE [4].



### 2.1.3 - Transmissão do VIH

O VIH transmite-se através dos fluidos corporais, tais como o sangue, sémen, secreções vaginais, líquido do cérebro e da espinal-medula, que contenham células infectadas ou partículas do vírus. Este também está presente nas lágrimas, na urina e na saliva, mas em concentrações ínfimas e portanto não é transmitido por essas vias, facto que não invalida o normal convívio social entre os indivíduos [3].

Pode dizer-se que se trata de um vírus relativamente frágil no aspecto em que não sobrevive fora do organismo, à temperatura ambiente, mais do que uma hora, salvo em condições excepcionais [4]. A sua transmissão entre indivíduos verifica-se na sequência de relações sexuais (anais ou vaginais) desprotegidas, transfusão de sangue, partilha de agulhas contaminadas, e entre mãe-filho durante a gravidez, parto ou na amamentação [5], constituindo estes os principais grupos de risco.

### 2.1.4 - Diagnóstico do VIH

Pressupondo elevado índice de suspeita clínica, o diagnóstico e confirmação da infecção pelo VIH são realizados por Serologia, efectuados precocemente através da pesquisa do antigénio p<sub>24</sub> (proteína de 24 kilodaltons do nucleocapsídeo do vírus da imunodeficiência humana tipo I) ou da carga viral (ARN VIH-1 por PCR com quantificação ou do ADN VIH-1), ou, se após a seroconversão, pesquisa de anticorpos através de duas provas imunoenzimáticas – no teste ELISA (Balows *et al*, 1991). No entanto, estes podem conduzir a falsos positivos, tal que nos resultados positivos aplica-se de seguida um outro teste, mais específico e caro, como por exemplo o teste de *Western Blot* (muito utilizado nos Estados Unidos) ou de testes ELISA's mais específicos (no caso do Reino Unido), de forma a confirmar a suspeita. Nenhum destes testes detecta a presença do vírus em indivíduos recentemente infectados [6], mas fora do contexto da primo-infecção, se o teste ELISA resultar negativo é suficiente para excluir a infecção (Miranda, 2003).

Posteriormente, a determinação da carga viral do VIH e a contagem de linfócitos T4 (células CD4) são utilizados no acompanhamento laboratorial dos indivíduos infectados.

### 2.1.5 - Fases da doença

Os indivíduos infectados com VIH perdem os linfócitos T4 em três fases [3]:

1. **Assintomática:** entre 2 a 4 semanas após a infecção, 50 a 89% dos indivíduos podem experienciar um síndrome viral agudo (SVA) durante 1 a 3 semanas, com sintomas semelhantes aos de uma gripe, não lhes dando por isso a devida importância (Oliveira).

Uma pessoa saudável tem um número de linfócitos T4 de aproximadamente 800 a 1300 células por microlitro de sangue. Nos primeiros meses posteriores à infecção pelo VIH, o número pode reduzir-se de 40 a 50 %. Por esta altura ocorre também o chamado “período de janela”, que antecede a seroconversão, pois não existem ainda níveis detectáveis de anticorpos e portanto os testes ao vírus acabam por resultar em falsos-negativos [6]. A seroconversão só ocorre na convalescença da SVA (Le SIDA parlons-en, 2001).

Durante estes primeiros meses, o doente pode transmitir o VIH a outros porque no seu sangue circulam muitas partículas do vírus, e apesar do seu organismo lutar contra este, é incapaz de eliminar a infecção;

2. **Sintomática Não SIDA:** depois de aproximadamente 6 meses, o número de partículas de vírus no sangue atinge um valor estável, que varia de pessoa para pessoa. No entanto, continuam em número suficiente para prosseguir a destruição de linfócitos e transmitir a doença a outros indivíduos. Nesta fase “silenciosa” do vírus, o seu portador já é considerado seropositivo por ter um nível de anticorpos detectável, não tendo ainda apenas desenvolvido SIDA, e continua a poder transmitir a infecção. Podem passar muitos anos, em média 8 a 11 (Pereira, 2008), nos quais se verifica uma diminuição lenta mas progressiva dos valores dos referidos linfócitos até níveis abaixo do normal (cerca de 50 a 100/mm<sup>3</sup> ao ano (Miranda, 2003). Os elevados valores de partículas virais e os baixos valores de linfócitos ajudam o médico a identificar os doentes com maior risco de desenvolver SIDA;
3. **SIDA:** durante um ano ou dois anteriores ao desenvolvimento da SIDA, o número de linfócitos T4 costuma descer mais rapidamente. A vulnerabilidade à infecção aumenta à medida que o número daqueles linfócitos baixa para menos de 200 células por microlitro

de sangue. O nível de vírus no corpo e o aparecimento de certas infecções oportunistas indicam que a infecção VIH progrediu para a SIDA. Deixa de haver resposta imunitária eficaz a invasores. O indivíduo começa a ter os sintomas que advêm das doenças oportunistas, acabando por falecer. A sobrevivência média após a ocorrência de uma infecção oportunista definidora de SIDA é de 1,3 anos (Miranda, 2003);

Existem no entanto cerca de 5 a 15% dos indivíduos (Miranda, 2003), denominados “não progressivos a longo prazo”, que permanecem com contagens de T4 acima de 600 células por microlitro de sangue durante longos períodos (10 a 15 anos após a infecção), tal que o tempo que decorre até ao estadio da SIDA é muito mais alargado, e ainda os que simplesmente não desenvolvem SIDA (cerca de 1 a 2% daqueles). Destes últimos, crê-se que um grupo de proteínas identificadas como anti-VIH (defensivas alfa-1, alfa-2 e alfa-3) podem estar na origem dessa imunidade natural contra o vírus [7];

De acordo com a etapa ou situação clínica em que indivíduo se encontra, e o seu nível de células CD4, a tabela de classificação do CDC divide os indivíduos adultos e adolescentes em três categorias clínicas (A, B e C) e três categorias para as células CD4. Assim:

<b>CD4 (cel/mm<sup>3</sup>)</b>	<b>Assintomático, Infecção Aguda, LGP</b>	<b>Sintomático</b>	<b>Indicadores de SIDA</b>
>=500	A1	B1	C1
200 - 499	A2	B2	C2
< 200	A3	B3	C3

Tabela 1 – Classificação do CDC para a infecção VIH

Da interpretação da tabela verifica-se que os estádios A3, B3, C1, C2 e C3 são definidos como SIDA pois esses indivíduos apresentam doença definidora de SIDA e/ou células CD4 <200 por microlitro de sangue. Este valor deve ser determinado pelo valor mais baixo e não pelo mais recente (Washington *et al*, 2006; Neves, 2008).

### **2.1.6 - Tratamento**

Não existe cura nem vacina para o VIH/ SIDA. Existem apenas esquemas terapêuticos para controlar a doença e por conseguinte a qualidade e esperança de vida do seu portador, transformando a SIDA numa doença crónica.

Durante os últimos anos têm sido desenvolvidos vários tipos de medicamentos antiretrovirais, cuja administração conjunta durante prolongados períodos de tempo proporciona a travagem da reprodução do vírus. De facto, estes medicamentos tentam agir directamente ao nível das funções da biologia do vírus que sejam suficientemente diferentes das funções de células humanas, e o seu objectivo é o controlo do número de virions no organismo, evitando dessa forma a descida do número de linfócitos T4, e aumentando a longevidade do doente. Não eliminam a infecção, impedem apenas a progressão do vírus no corpo.

Estudos defendem que na ausência de qualquer tratamento, a média estimada de sobrevivência após a infecção pelo VIH é de 11 anos (OMS, 2007).

Um aspecto fundamental do tratamento da SIDA consiste no tratamento das infecções e tumores oportunistas, que costuma necessitar de hospitalização.

Existem medicamentos que combinam os seguintes grupos: inibidores da enzima transcriptase reversa (NITR e NNITR) através da qual o vírus se reproduz, inibidores da protease (IP) que cliva as proteínas do vírus após a transcrição, inibidores da enzima integrase (II), responsável pela entrada do material genético do vírus no núcleo da célula humana, inibidores do CCR5, impedindo assim um dos meios de ligação do vírus à célula-alvo, e inibidores de fusão (IF) cujo objectivo é dificultarem a forma como o vírus se funde à membrana plasmática da célula. A terapêutica combinada permite uma melhor estabilização do doente, supressão da carga viral e previne resistência à medicação (Nájera, 1997).

O mais comum é o doente só começar a ser medicado quando se encontra na fase sintomática com doença definidora de SIDA. Em Portugal, para dar efectivo início ao tratamento, o médico é obrigado a notificar o caso às autoridades (Centro de Vigilância Epidemiológica das Doenças Transmissíveis – Instituto Nacional de Saúde Dr. Ricardo Jorge), e o doente tem de apresentar uma situação clínica individual justificável (ex. gravidez) e/ou estar já na fase sintomática com

doença definidora de SIDA, ou com uma contagem de células CD4 inferior a 350 células por microlitro de sangue, e/ou carga viral superior a 100.000 cópias VIH-1 RNA por mililitro de sangue.

Sabe-se que quanto mais cedo os medicamentos começarem a ser administrados ao doente, mais facilmente a infecção é controlada, mas por outro lado este tipo de medicamentos acarreta efeitos secundários adversos, e não mantêm a sua eficácia a longo prazo, o que justifica o facto de serem regularmente substituídos, e são também muito dispendiosos, tal que a maior parte dos médicos opta por monitorizar o doente através da contagem de linfócitos T4, e só inicia o tratamento quando se considera que o doente atingiu a fase pré-crítica (pré-SIDA) e este se encontra devidamente motivado para aderir à terapêutica (ACSS, 2007). De facto, a eficácia do tratamento depende muito das características biológicas do doente e da fase da doença em que se encontra, mas principalmente do seu nível de adesão (respeito pela prescrição), da forma como cumpre a medicação, e da sua responsabilidade e motivação.

Existe pela parte da comunidade em geral, e do doente em particular, um grande estigma associado à doença. Por um lado, o não querer tomar a medicação porque isso é um aceitar que se está doente e na verdade não se sente como tal, e por outro o desgaste associado à contínua toma de medicação para evitar ficar doente (Albino, 2008).

Como já referido, a considerar é também o facto dos efeitos colaterais e complicações que a terapia antiretroviral normalmente acarreta. Em grande parte dos doentes surgem novas patologias associadas a complicações hepáticas e tumores, entre outros (Oliveira). Além disso, uma percentagem muito significativa dos doentes com infecção VIH tem infecção activa de vírus da hepatite B e/ou C (VHB e/ou VHC), talvez pelo facto de estas compartilharem as mesmas vias de transmissão, e o que se verifica é que a terapêutica antiretroviral agrava a hepatite B e, por outro lado, a hepatite C afecta a evolução da infecção VIH. Por conseguinte, é comum que a maior parte dos doentes tenha também de seguir tratamento para a hepatite.

Em Portugal, o utente seropositivo ou com SIDA está isento de taxas moderadoras nos serviços de saúde [8], e a terapêutica antiretroviral é universal, gratuita e de distribuição hospitalar [8]. Mas tal não é suficiente para controlar a infecção. A cobertura de tratamento antiretroviral para adultos e crianças com VIH/SIDA em estado avançado é superior a 75%, mas a proporção actual

de adesão à terapêutica é de aproximadamente 40% [9], muito aquém dos 95% defendidos como taxa ideal para o sucesso do tratamento.

## 2.2 – Dados actuais sobre a doença

Estima-se que mais de 15 mil pessoas sejam infectadas por dia no mundo, 33 milhões estão actualmente infectadas e, em média, 3 milhões morrem a cada ano. A maioria dos casos ocorre em África, onde a principal forma de transmissão é o sexo heterossexual (ONUSIDA, 2008).

As regiões em risco que apresentam crescimento elevado no número de novas infecções são o Leste da Europa, Índia e Sudeste Asiático (ONUSIDA, 2008).

A taxa de infecção em toxicod dependentes é de aproximadamente 80% nas cidades europeias e americanas.

Embora seja considerado e punido como crime o acto de propagação de doença contagiosa, estudos revelam que 3 em cada 4 pessoas com infecção VIH escondem este facto dos seus parceiros sexuais casuais, o que agrava claramente a situação epidémica da doença [12].

### 2.2.1 - Dados Globais

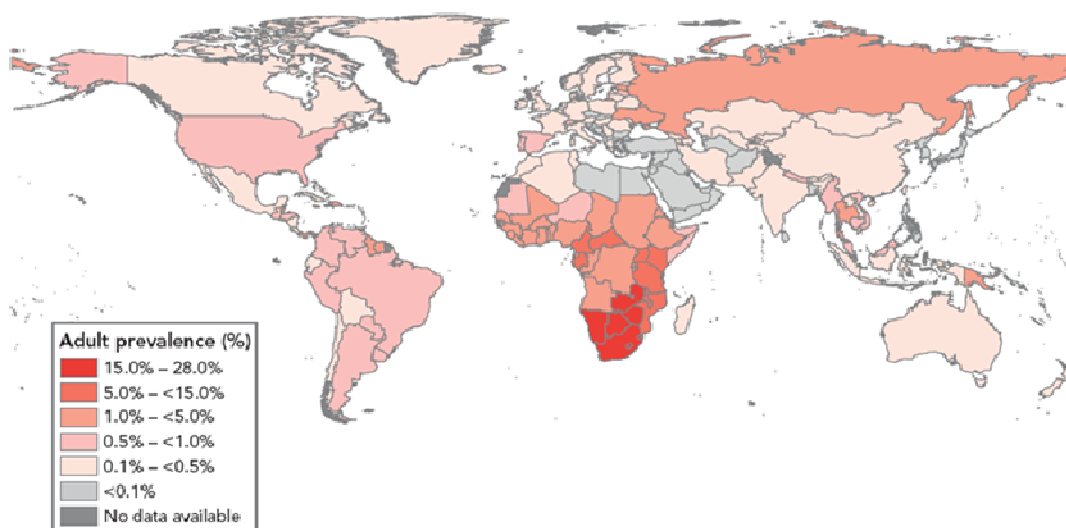


Figura 1 – Mapa-mundo representativo da taxa de prevalência da infecção VIH/ SIDA (ONU. 2008)

De acordo com os dados da ONUSIDA, referentes ao ano 2007, nos países mais afectados pelo VIH, a esperança de vida diminuiu em aproximadamente 20 anos, e desde que a epidemia começou, 25 milhões de pessoas, a nível mundial, já morreram devido ao vírus. Tem por isso provocado profundas alterações demográficas nesses países.

Em 2007, o número estimado de novas infecções pelo VIH era 2,5 vezes superior ao número de doentes submetidos à terapêutica antiretroviral nesse ano, indiciando que o maior investimento deverá ser feito na prevenção de novos casos, não só através da sensibilização das populações mas também pelo combate à pobreza e desigualdades económicas. Crê-se que neste ano, cerca de 33 milhões de pessoas estavam infectadas, 2,7 milhões adquiriram a infecção nesse período (novos casos), e 2 milhões morreram por causa relacionada com o VIH (OMS, 2007). Cerca de 30,8 milhões dos infectados são indivíduos com mais de 15 anos. A taxa de prevalência nos adultos (15-49 anos) é de 0,8%. E as mulheres com mais de 15 anos representam 47% do total de infectados.

A taxa de prevalência nos jovens adultos (15-24 anos) é de 0,6% e 0,4% para as mulheres e homens, respectivamente.

A região mais afectada é a África Subsariana, com cerca de 67% do total mundial de infectados e 75% das mortes por SIDA.

A Indonésia e Federação Russa constituem as preocupações mais emergentes, uma vez que o número de casos nestes territórios está a aumentar rapidamente.

Por sexo, verifica-se que a percentagem de cada um dos sexos no total de infectados tem permanecido estável (50%), no entanto, a perspectiva é que aumente o número de mulheres infectadas.

O grupo etário entre os 15 e 24 anos representa cerca de 45% do total de novas infecções, a nível mundial. E no que respeita aos dados sobre crianças, crê-se que cerca de 370.000 indivíduos com menos de 15 anos adquiriram a infecção em 2007, representando cerca de 4,8% do total de pessoas infectadas. E cerca de 270.000 morreram devido à SIDA. De facto, a progressão da doença nas crianças é mais agressiva e por isso a sua taxa de mortalidade é elevada. Em 2007, estimou-se em 2 milhões o número total de crianças a viver com o vírus.

Os grupos de risco e sua representatividade no total de infectados são muito variáveis de região para região. Mas a nível global, pode afirmar-se que a transmissão heterossexual permanece a mais significativa, seguida da transmissão entre utilizadores de drogas injectáveis e comportamento homossexual.

### **2.2.2 - Dados Europa e União Europeia**

A infecção por VIH é considerada um importante problema de saúde pública na Europa.

Os últimos dados disponíveis do EuroHIV reportam-se a 2006, e referem um total de 86.912 novos casos diagnosticados de infecção por VIH (69% da Europa de Leste, 29% da Europa Ocidental e 2% da Europa Central), bem como 13.199 casos de SIDA (49% da Europa Ocidental, 47% da Europa de Leste, e 4% da Europa Central) diagnosticados nos 50 dos 53 países da região europeia da OMS. A tendência é crescente para os casos de nova infecção, mas o número de casos de SIDA tem vindo a decrescer, excepto na Europa de Leste (HIV/AIDS EuroHIV, 2007).

A taxa de incidência da infecção por VIH na Europa é de 111,1 por milhão de habitante.

Se considerarmos apenas a União Europeia, foram reportados 26.220 novos casos de infecção (cerca de 30% do total). A sua taxa de incidência é de 67,2 por milhão de habitante. E foram diagnosticados 6.720 casos de SIDA, cuja taxa de incidência é de 13,7 por milhão de habitantes.

Verifica-se que a taxa de incidência da infecção por VIH varia muito entre as três sub-regiões da Europa. De facto, o maior número de casos reportados provém da Europa de Leste (com uma taxa de incidência de 211 por milhão de habitantes, bastante superior aos verificados na Europa Ocidental - 83 por milhão de habitantes - e Europa Central - 9 por milhão de habitantes).

Na Europa Ocidental, apenas 10% dos novos casos diagnosticados são jovens (dos 15 aos 24 anos), e 35% do sexo feminino.

Se a análise for efectuada por países, tem-se a Estónia, Ucrânia, Federação Russa e Portugal com taxas acima dos 200 por milhão de habitante (em Portugal aquele valor é de 205), seguidos do



Reino Unido, República da Moldávia, Letónia, Luxemburgo e Suíça, com valores acima dos 100 por milhão de habitante.

Quanto ao número de casos de SIDA diagnosticados, embora o seu valor tenha sido, em 2006, superior na Europa Ocidental, a sua taxa de incidência é inferior à verificada no Este (15,9 e 25 por milhão de habitantes, respectivamente). Na Europa Central o número de novos casos de SIDA diagnosticados no ano é de 2,8 por milhão de habitantes.

Na análise por países, verifica-se que a Ucrânia, Portugal e Espanha apresentam taxa de incidência superior a 20 por milhão de habitantes. Em Portugal o seu valor é igual a 66, por milhão de habitantes.

Na Europa, as doenças oportunistas definidoras de SIDA mais diagnosticadas são a tuberculose pulmonar (35%), síndrome de *wasting* de VIH (20%), e pneumonia por *Pneumocystis carini* (13%), sem diferença entre sexos.

No que respeita à mortalidade, 4386 pessoas morreram no ano 2006, um valor inferior ao reportado no ano anterior (em cerca de 12%).

De facto, até 2006 foram diagnosticados um total de 327.068 casos de SIDA, sendo que 54% destes morreram até ao final do ano em análise. Por conseguinte, cerca de 150.000 pessoas estão a viver com a doença (e possibilidade de transmiti-la).

Na Europa de Leste, a principal via de transmissão do vírus VIH é a partilha de seringas pelos toxicodependentes (IDU), embora se registe um aumento na transmissão via heterossexual. Na Europa Ocidental e Central a principal categoria são os heterossexuais (com um peso de 54% e 52%, respectivamente), embora se tenham verificado aumentos no número de casos reportados de transmissão homossexual (MSM). A taxa de transmissão mãe-filho, transfusão de sangue ou nosocomial varia entre os 1 e 5%.

### **2.2.3 – Dados Portugal**

Ao contrário dos restantes países desenvolvidos, em Portugal a incidência de SIDA e a mortalidade associada à infecção têm permanecido estáveis, apesar do acesso generalizado aos

esquemas terapêuticos combinados (ACSS, 2007). O nosso país apresenta a segunda maior incidência da infecção VIH e a maior incidência de SIDA na União Europeia. Além disso, a infecção VIH/SIDA é a causa de morte mais representativa entre os 30 e 39 anos (ACSS, 2007).

A ONUSIDA estima que existam cerca de 34.000 pessoas a viver com o VIH no nosso País. A taxa de prevalência para os adultos infectados entre os 15 e os 49 anos é de 0,5% (OMS, 2007). Se analisarmos o grupo etário dos 15 aos 24 anos, têm-se taxas de prevalência de 0,3% e 0,5% para as mulheres e homens, respectivamente. De facto, 28% dos indivíduos infectados são mulheres com mais do que 15 anos (OMS, 2007).

“Permanecem escassos os dados disponíveis sobre a infecção e seus determinantes, estão dispersos pelos diversos níveis do sistema de saúde, a sua recolha nem sempre obedece a critérios uniformes e concordantes com os padrões internacionalmente definidos, resultando pouco fiáveis” (CNLCS, 2007). O sistema de declaração obrigatória não responde às necessidades nacionais e a subnotificação é a norma, embora nos últimos anos se tenha verificado mais rapidez na notificação após o diagnóstico. Tal induz-nos a aceitar que a realidade do impacto da doença está a ser subestimada pelos dados disponíveis.

De acordo com a OMS/ ONUSIDA, a epidemia portuguesa é do tipo concentrado. A prevalência na população geral portuguesa é inferior a 1% mas pelo menos em três grupos de risco (utilizadores de drogas injectáveis, homossexuais e reclusos) é superior a 5%. De facto, os reclusos são uma população preocupante, pela sua vulnerabilidade, com elevadas taxas de prevalência de VIH (cerca de 10%), e incidência (superior a 80 por milhão de pessoas ao ano) e outras infecções transmitidas sexualmente como hepatite B e C, e ainda tuberculose (Mauch, 2009).

Portugal é o terceiro país da União Europeia com mais casos de co-infecção VIH/SIDA e tuberculose, tendo-se verificado a presença de VIH/SIDA em 15% dos casos de tuberculose avaliados (CLNCS, 2007).

Cerca de 46 unidades hospitalares fornecem tratamento antiretrovívico (ACSS, 2007). Por região de saúde, cerca de 60% dos doentes são acompanhados em Lisboa e Vale do Tejo, 28% no Norte, 7% no Centro, 4% no Algarve e 1% no Alentejo.

O peso relativo das vias de transmissão da infecção tem alterado ao longo do tempo no nosso País. Os utilizadores de drogas injectáveis têm sido, desde o início e até ao ano 2003 a maior proporção de infectados, mas a transmissão heterossexual está a sobrepor-se à transmissão parentérica. Crê-se que 16,5% dos casos de transmissão heterossexual ocorreram em indivíduos provenientes de países com epidemia generalizada (CNLCS, 2007).

A transmissão homossexual está a manter-se relativamente estável, a transmissão mãe-filho tem diminuído (situando-se a respectiva taxa abaixo dos 2%), e a transmissão associada à dávida de sangue e órgãos é praticamente inexistente (INSA, 2009).

A publicação mais recente do Instituto Nacional de Saúde indica que durante o ano 2008 foram notificados 2668 casos de infecção pelo VIH, nos vários estadios (31% casos de SIDA, 12% casos sintomáticos não-SIDA e 57% casos de portadores assintomáticos). Daquele total de casos notificados, 1201 (45%) foram diagnosticados nesse mesmo período. As categorias mais relevantes são: heterossexuais (57,6%) toxicodependentes ou utilizadores de drogas injectáveis (21,9%) e homo/bissexuais (16,8%). Refira-se ainda que, para todas as categorias, é maior o peso dos portadores assintomáticos, seguidos pelos casos de SIDA e por último os casos sintomáticos não-SIDA (INSA, 2009).

A 31 de Dezembro encontravam-se notificados 34.888 casos de infecção VIH/SIDA (15.020 casos de SIDA, 16.494 casos de portadores assintomáticos – PA e 3374 casos sintomáticos não-SIDA - CRS), constituindo os Toxicodependentes cerca de 42,5% de todas as notificações, evidenciando a tendência inicial da epidemia no País. A categoria Heterossexual tem um peso crescente de 40% e os Homo/Bissexuais cerca de 12,3%. As restantes formas de transmissão perfazem 5,2% do total (INSA, 2009).

Quanto ao sexo, cerca de 74% do total de casos notificados são homens (INSA, 2009).

Cerca de 25% do total geral de casos notificados correspondem a óbitos, sendo que os casos de SIDA (85%) são os que mais contribuem para esse valor. (INSA, 2009)

Ainda de acordo com os dados do INSA (2009), o total acumulado de casos de SIDA naquela data era de 15.020, dos quais 81,7% são indivíduos do sexo masculino, e o grupo etário mais significativos para ambos os sexos é entre os 20 e os 49 anos (e neste pesa mais o grupo dos 30

aos 34 anos). Nos homens o grupo mais representativo é dos 30 aos 34 anos, e nas mulheres é dos 25 aos 29 anos. As categorias de transmissão mais representativas são, por ordem, Toxicodependentes (47%), Heterossexuais (36%) e Homo/Bissexuais (13%).

O peso das infecções pelo VIH 1 é de 95%, seguido do VIH 2 e de infecções associadas a ambos.

Os dados referentes à mortalidade não estão completos pela usual falta de notificação quando ocorre o óbito, mas os dados existentes permitem afirmar que 48% dos casos de SIDA correspondem a óbitos ocorridos desde a sua data de diagnóstico e posterior notificação.

Outro aspecto relevante é o decréscimo na taxa de mortalidade que se tem verificado após a introdução da TAR. Isso por um lado parece ser um bom sinal mas não esqueçamos que o número de pessoas capazes de transmitir o vírus é superior. De facto, “o efeito da terapêutica antiretrovívica no decréscimo do número de mortes nos indivíduos infectados pelo VIH, associado aos novos casos de infecção (incidência), levou a um aumento do número total de indivíduos infectados na população” (INSA, 2009).

No que respeita à patologia associada aos casos de SIDA notificados até 2008, 88% apresentam infecção oportunista (sendo a mais frequente e mortífera a Tuberculose), 4% Sarcoma de Kaposi, igual peso para infecção oportunista e Sarcoma de Kaposi, seguidos de Linfoma, encefalopatia, Síndrome de emaciação por VIH, Pneumonia Intersticial Linfóide e Carcinoma invasivo do colo do útero.

Por categoria de transmissão, as infecções oportunistas representam 93% das patologias observadas nos Toxicodependentes, 88% nos Heterossexuais e 72% nos Homo/Bissexuais.

O distrito com maior número de casos notificados de SIDA é Lisboa (com 40%), seguida do Porto (23%) e Setúbal (13%). Os distritos com menos casos de SIDA são Portalegre (0,2%), Guarda e Bragança (ambos com 0,3%).

Quanto aos casos sintomáticos não-SIDA, 75% são homens, e o grupo etário mais representativo para ambos os sexos é dos 20 aos 49 anos. As categorias de transmissão aqui mais representativas são: Heterossexuais (cerca de 43%), seguidos dos Toxicodependentes (36%), e Homo/Bissexuais (15%). A maior parte das infecções é do tipo VIH 1 (93%).

Cerca de 14% dos casos notificados de portadores nesta fase do vírus constituem óbitos, a maior parte com patologia caracterizada por sintomas constitucionais.

A tendência na forma de transmissão do vírus é diferente da registada nos casos de SIDA. De facto, embora a Heterossexual seja crescente e a Homo/Bissexual relativamente estável, verifica-se nestes portadores um decréscimo acentuado da infecção via endovenosa para os utilizadores de drogas.

Quanto às patologias verificadas nestes doentes, incluem-se em infecções oportunistas não incluídas na classificação epidemiológica para os casos de SIDA.

Considerando agora o grupo mais representativo (47%) no total acumulado de casos notificados até final de 2008 – Portadores Assintomáticos (dos quais 1510 foram notificados durante o último ano), este é constituído por 68% de homens. O grupo etário mais representativo continua a ser dos 20 aos 49 anos.

Cerca de 5% dos casos notificados correspondem a óbitos.

As principais categorias de transmissão são o grupo dos Heterossexuais (43%), Toxicodependentes (40%) e Homo/Bissexuais (11%). A tendência no peso de cada uma das categorias no total de infectados assintomáticos é ligeiramente decrescente para os Heterossexuais, decrescente acentuado para os Toxicodependentes e crescente para os Homossexuais.

Quanto à sensibilização da população em geral para a doença, um estudo elaborado por Amaro *et al.* (2004), indicam que 68% das pessoas (foram inquiridos 1000 indivíduos entre os 18-69 anos seleccionados aleatoriamente na população portuguesa) percepciona o seu risco de infecção elevado, 48% considera-se em maior risco que a generalidade da população, e 37% refere que já sentiram medo de terem sido infectados. Quanto ao comportamento sexual, apenas 22% dos inquiridos referiram o uso consistente de preservativo.

Em termos de discriminação, cerca de 38% dos inquiridos consideraram que os infectados deveriam ser isolados (total ou parcialmente), e segundo o Eurobarómetro, Portugal apresenta

níveis desfavoráveis face aos outros países europeus no que respeita ao conhecimento e compreensão da doença. De facto, 30% continuam a considerar que a infecção se transmite pelo beijo, 18% pela comida e talheres e 5% pelo aperto de mão. Entre os jovens, dos 15 aos 19 anos, apenas 15% referiram conhecimentos correctos sobre o VIH e 39% referiram não ter usado preservativo na última relação sexual (Amaro *et al.*, 2004).

#### **2.2.4 – Dados Alentejo, Baixo Alentejo, Beja**

Não existem muitos dados disponíveis acerca do impacto da infecção na região. De acordo com o INSA (2009) o Alentejo tem um peso de 1,5 %, sendo que o distrito de Beja representa 0,5% do total de casos notificados de SIDA em Portugal. Quanto aos portadores sintomáticos não-SIDA ou CRS, os indivíduos cujo distrito de residência é Beja constituem 1,1%. (Alentejo – 1,9%)

No caso dos portadores assintomáticos – PA, cerca de 0,8% residem no distrito de Beja. (Alentejo – 1,9%).

Os dados do DDI-URVE (2009) sobre o número de casos com residência no distrito de Beja à data da notificação identificam 261 indivíduos (dos quais 140 PA, 83 SIDA e 38 CRS). Os homens parecem ser o sexo mais afectado pois representam cerca de 74% do total.

A maioria dos indivíduos tem entre 25 e 34 anos.

Toxicodependente é o grupo de risco com maior peso, seguido do heterossexual e homo/bissexual. O tipo de vírus mais frequente é o VIH-1 (cerca de 98% dos casos).

Em relação à sua situação, 47 dos casos notificados são de indivíduos já falecidos e 214 de indivíduos em vida. A tuberculose representa a doença indicadora de SIDA com maior número de notificações (DDI-URVE, 2009).

No que respeita ao concelho de residência, a maioria dos casos são provenientes de Beja (cerca de 42% do total). Mas mesmo assim Odemira, Castro Verde e Aljustrel têm também algum peso. O concelho com menos casos notificados é Cuba (DDI-URVE, 2009).

O facto do caso notificado pertencer ao distrito de Beja não implica necessariamente que esse doente esteja a ser ou tenha sido seguido no hospital distrital de Beja. Na verdade, espera-se que

existam muitos utentes que estão a ser seguidos na região mas cuja residência seja outra que não a do distrito de Beja, e vice-versa. Tal deve-se ao factor discriminação e preconceito que leva os doentes a esconder a doença da comunidade onde vivem e preferirem tratar-se noutra localidade.

Os doentes que estão a ser seguidos no Hospital de Beja provêm de diversas fontes: referenciação do CAD ou CAT, utentes entrados na urgência com elevada suspeita da doença, ou utentes que desconfiam estar infectados e dirigem-se directamente aos Centros de Aconselhamento ou ao Hospital. Após a confirmação da infecção o doente passa a ser seguido por uma equipa de médicos, enfermeiros, psicólogos, assistente social, nutricionista e farmacêutica. Estes tentam transmitir-lhe (ao doente e família) o máximo de informação sobre a doença, monitorizam a sua evolução e proporcionam ainda o acompanhamento multidisciplinar caso o doente inicie a terapêutica antiretrovítica. A frequência dessas consultas de acompanhamento é variável de acordo com o tipo de doente, pois depende não só da sua carga vírica e evolução da doença, mas também do seu nível de adesão e atitude perante a doença. De facto, existem uma série de factores que condicionam o número de idas às consultas de acompanhamento, tais como características do próprio indivíduo e da doença, a relação equipa de saúde/ doente e entre outros, a dificuldade no acesso à consulta periódica de acompanhamento consoante a distância do seu local de habitação e o hospital (Albino, 2008).

## **Síntese – Capítulo 2**

Neste capítulo foram abordados os tópicos mais relevantes no enquadramento teórico sobre a doença. De facto, o VIH/ SIDA tem progredido de forma rápida a nível mundial, e é esta sua característica de elevada transmissibilidade, aliada à falta de informação da população em geral, que vaticinam o alastramento cada vez maior da epidemia. Além disso, a ciência não conseguiu ainda encontrar uma solução terapêutica ideal e, por isso, o vírus continua a ser um dos maiores causadores de morte no mundo.

### **3. Metodologia e Alguns Aspectos Teóricos**

Este capítulo apresenta a metodologia de investigação adoptada, mais concretamente, os métodos e instrumentos utilizados para a obtenção e análise dos dados. Numa primeira secção é apresentada a forma de obtenção dos dados, assim como a construção da base de dados. De seguida são apresentados os métodos estatísticos utilizados para proceder à análise dos dados, nomeadamente a estatística descritiva e inferencial numa fase inicial e a análise de sobrevivência na fase posterior. A explanação dos métodos estatísticos usados no presente trabalho de investigação tem como objectivo clarificar os procedimentos efectuados, sendo por isso sumária.

#### **3.1 – Obtenção dos dados e construção da base de dados**

O processo de recolha dos dados e construção da respectiva base de dados foi moroso e implicou a intervenção junto de diversas fontes.

Sendo o objectivo principal estudar os indivíduos infectados que foram ou continuam a ser monitorizados no Hospital de Beja, considerando que a área de influência deste hospital é precisamente o Baixo Alentejo, sabia-se que os dados estariam contidos nos registos hospitalares. E de facto estavam, mas, contrariamente ao esperado, de uma forma muito dispersa em todo o universo de utentes que recorrem ou são referenciados para o hospital.

As dificuldades iniciais prenderam-se com as situações do indivíduo infectado consciente da sua doença poder ou não identificar-se como portador do vírus aquando da sua chegada aos serviços hospitalares (desconhecendo estes *a priori* a sua infecção) e nesse caso o registo de admissão fica à partida condicionado, mas também porque os utentes podem simplesmente ter recorrido às várias valências hospitalares (nomeadamente Internamento, Consulta Externa, Urgência e Bloco Operatório) e isso implicaria conjugar demasiadas fontes em busca dos utentes com a infecção, o que seria impraticável face ao tempo disponível.



Chegou-se entretanto à conclusão que todos os utentes infectados têm algo em comum: são referenciados para uma determinada especialidade de consulta externa, onde são seguidos. Isso proporcionou evitar a redundância da situação anterior e se antes a pesquisa tinha de ser efectuada de entre todos os utentes admitidos no hospital e respectivas patologias, passou então a ser efectuada de entre todos os utentes que frequentaram aquela consulta, sendo que esta é de cariz geral e portanto contém outros utentes que não apenas os infectados por VIH/SIDA.

Nesta fase, foi a própria equipa responsável pelo acompanhamento dos indivíduos infectados que disponibilizou uma listagem de utilização interna, com os números de processo dos doentes seguidos desde o ano 2000 até final do ano 2009, bem como os dados referentes ao sexo, data de diagnóstico, ano de entrada, data de nascimento, população linfática, carga viral, situação clínica, contagem de células CD4, situação actual no acompanhamento hospitalar, tipos de vírus, terapêutica antiretroviral, patologias identificadas, comportamentos de risco, sector de actividade e concelho de residência. Esta listagem foi vital para a prossecução do estudo, e, embora não estivesse completa, constituiu o efectivo ponto de partida da presente investigação.

Identificados os doentes, recolheu-se a restante informação, utente a utente, tal que a base de dados final tem cerca de 50 variáveis, devidamente apresentadas no próximo capítulo 4 (Análise Exploratória dos Dados).

## **3.2 – Métodos de análise dos dados**

A análise dos dados é realizada, numa primeira fase, através de mapas, histogramas e medidas da estatística descritiva e inferencial. Numa segunda fase, são estimados modelos da análise de sobrevivência com vista a explorar e compreender quais as variáveis que efectivamente podem influenciar o tempo de sobrevivência dos indivíduos em estudo.

Relativamente ao software estatístico utilizado para o tratamento dos dados é efectuado com recurso ao SPSS (v. 17) e R-Project (v. 2.8.0).

### **3.2.1 – Estatística descritiva e inferencial**

Inicialmente, os dados são alvo de análise através de medidas da estatística descritiva, nomeadamente média, desvio-padrão, assimetria, curtose. A análise destas medidas, complementada com histogramas, caixas de bigodes, tabelas de contingência, gráficos de barras

e tabelas de frequências, dão a possibilidade de melhor compreender a realidade em mãos e simultaneamente analisar eventuais relações entre variáveis.

Com vista a generalizações acerca dos estimadores calculados, serão elaborados alguns testes de hipóteses que se considerem adequados em determinadas circunstâncias, nomeadamente testes de hipóteses para a diferença entre proporções, testes de hipóteses para a diferença de médias e testes de independência (Teste Qui-Quadrado e Teste Fisher).

Os testes para proporções são adequados quando os dados sob análise consistem em contagens, frequências ou percentagens. Estes testes visam avaliar uma determinada afirmação sobre a proporção ou percentagem de uma população. Em muitos aspectos são muito semelhantes aos testes para as médias, sendo que nos testes para proporções os dados amostrais apresentam-se em contagens, ao invés de medidas.

Com o teste de independência pretende-se verificar se duas características de uma determinada população podem ser consideradas independentes, sendo analisada para o efeito uma amostra aleatória e representativa dessa população. O teste de independência considerado no presente estudo – Teste Qui-Quadrado – é um teste não paramétrico que visa apurar se existe associação entre duas variáveis do tipo categórico. Os valores observados são dispostos numa tabela de contingência  $2 \times 2$ , para a qual se calculam os respectivos valores esperados. Se as variáveis forem independentes as frequências observadas estarão próximas das esperadas.

O Qui-Quadrado mede a probabilidade de as diferenças encontradas nos dois grupos da amostra serem devidas ao acaso, partindo do pressuposto que, na verdade, não há diferenças entre esses grupos na população donde provêm. Se a probabilidade for alta pode-se concluir que não há diferenças estatisticamente significativas. Se a probabilidade for baixa pode-se concluir que um grupo é diferente do outro grupo, quanto à característica estudada, e de forma estatisticamente significativa. Tem como pressupostos: as observações devem ser independentes; os itens de cada grupo são seleccionados aleatoriamente; as observações são frequências ou contagens; cada observação pertence a uma e uma só categoria; nenhuma frequência esperada pode ser inferior a 1; e não se aplica se 20% das observações ou valores esperados forem inferiores a 5.

O Teste Qui-quadrado deverá ser substituído pelo Teste Exacto de Fisher quando os valores esperados nas células da tabela sejam inferiores a 5.

### 3.2.2 – Análise de sobrevivência

A análise de sobrevivência (também designada de análise de duração) é uma técnica econométrica desenvolvida com o intuito de explicar e prever o tempo até falha ou tempo até ocorrência de um determinado evento. A sua aplicação consiste em modelar e determinar a relação entre a distribuição do tempo de vida e determinadas variáveis explanatórias associadas e que representem factores que se supõe afectarem a sobrevivência, sendo a variável resposta o tempo de sobrevivência.

Na ciência biomédica, o termo “sobrevivência” é especialmente usado, dado o elevado número de aplicações desta técnica, à análise do tempo de sobrevivência de doentes. Tem-se a possibilidade de estudar a sobrevivência num grupo de indivíduos ou compará-la em dois ou mais grupos.

Existem dois tipos de modelos de sobrevivência: contínuos e discretos. A utilização de ambos depende do tipo de variáveis em estudo, mais precisamente, do evento em estudo ocorrer em qualquer momento no tempo (análise contínua) ou em momentos discretos (análise discreta).

Dado que o evento de interesse deste estudo pode ocorrer em qualquer momento, serão apenas descritos modelos de análise de sobrevivência contínuos. Além disso, é assumido que a população é homogénea (salvo se especificamente mencionado o contrário), ou seja, as diferenças entre os indivíduos são descritas pelas covariáveis, todos os indivíduos em estudo são susceptíveis da ocorrência do evento (morte), existe apenas um evento de interesse e este ocorre apenas uma vez e individualmente. Assume-se ainda que o tempo censurado e de sobrevivência são independentes. Neste contexto, este estudo baseia-se essencialmente em modelos de dois estados (*two-stage models*), como descritos na Figura 2.

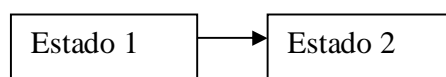


Figura 2 – Estrutura do modelo de dois estados

### 3.2.2.1 – Conceitos básicos

Os dados de sobrevivência consistem no tempo que decorre até à ocorrência de um determinado evento, que normalmente é designado por morte ou falha. Uma característica deste tipo de dados é a possibilidade da não observação do evento de interesse em alguns indivíduos, o que pode ser uma censura ou um truncamento.

Dados truncados são aqueles que para entrar no estudo foram sujeitos a um determinado condicionamento, tal que são desprezados todos os valores ou observações que não cumpram aquela condição.

Os dados censurados podem ser divididos em três tipos: (i) censura à direita, onde tudo que se sabe é que o evento ainda não ocorreu até ao instante observado; (ii) censura à esquerda, onde tudo o que se sabe é que o evento ocorreu em algum instante de tempo antes do início do estudo; (iii) e censura intervalar, é aquela em que se sabe que o evento ocorreu dentro de um intervalo de tempo conhecido (Cleves *et al*, 2004, Collet, 1994, e Hosmer e Lemeshow, 1999).

Neste trabalho de investigação verifica-se a existência de dados censurados à esquerda e dados censurados à direita. De facto, consideraram-se censurados todos os casos de indivíduos em que não se conseguiu confirmar a data de óbito ocorrida durante o período em análise, ou indivíduos cuja ocorrência do evento morte não tenha sido directamente relacionada com o VIH/SIDA, bem como os indivíduos cujo estado actual no final do estudo seja activo, e ainda os que abandonaram o acompanhamento hospitalar ou foram transferidos para outras instituições e como tal não se conseguiu saber se experimentaram (e quando) ou não o evento de interesse.

Além do tempo de sobrevivência e da variável indicadora de falha, os dados de sobrevivência podem conter um conjunto de variáveis observáveis que podem estar relacionadas com estes tempos. Estas variáveis são conhecidas por covariáveis ou variáveis explanatórias/ explicativas. Quando os tempos de sobrevivência estão relacionados com as covariáveis diz-se que a população é heterogénea. Caso contrário a população é tida como homogénea (como já referido anteriormente).

Por outro lado, considera-se que os dados são truncados quando se excluem algumas observações da amostra. Existem dois tipos de truncamento: (i) truncar à direita existe quando a amostra apenas inclui indivíduos que tenham experienciado o evento de interesse num determinado momento; (ii) truncar à esquerda existe quando a amostra exclui os indivíduos que tenham experienciado o evento de interesse antes do período em estudo. Cleves *et al.* (2004) refere ainda a possibilidade de fazer truncamento intervalar, o que inclui indivíduos com faltas de informação.

De acordo com diversos autores, nomeadamente Cleves *et al.* (2004) o método dos mínimos quadrados (OLS) não é adequado para analisar dados de sobrevivência, dado que este método de estimação assume que os resíduos são normalmente distribuídos, pressuposto que não é válido em dados relativos ao tempo de sobrevivência de algo (os valores dos tempos são sempre não-negativos, por exemplo). Apesar de tudo, a regressão linear pode ser robusta face à não-normalidade dos resíduos, mas perde robustez face a outras características dos dados de sobrevivência, nomeadamente a não simetria e o facto de serem não-unimodais.

Alguns autores, nomeadamente Jenkins (2005) consideram que modelos de variável dependente binária podem ser uma alternativa ao OLS dado que permitem o uso de dados censurados e truncados. Contudo a variável dependente refere apenas a existência ou não do evento e não o tempo relativo ao evento, perdendo-se por isso bastante informação (Allison, 2004).

### **3.2.2.2 – Alguns conceitos importantes**

A distribuição do tempo de sobrevivência é usualmente descrita por quatro funções: (i) função de sobrevivência; (ii) função de densidade; (iii) função de risco; e (iv) função de risco acumulada. Estas funções são matematicamente equivalentes, ou seja, a partir de uma poderão ser encontradas as restantes.

Seja  $T$  uma variável aleatória não-negativa e contínua, que representa o tempo de sobrevivência (medida em minutos, horas, dias, meses, anos, etc.), e  $t$  seja um valor específico para a variável  $T$

e os tempos de sobrevivência independentes entre si. A função densidade cumulativa de  $T$  é dada por:

$$F(t) = \int_0^t f(x)dx = P(T \leq t) \quad (1).$$

A função de sobrevivência constitui a probabilidade de um indivíduo sobreviver para além do tempo  $t$  (Collet, 1994) e pode ser definida da seguinte forma:

$$S(t) = \int_t^{\infty} f(x)dx = P(T > t) = 1 - F(t) \quad (2)$$

$S(t)$  é uma função monótona e não crescente no tempo  $t$ , com as seguintes propriedades:

$$S(t) = \begin{cases} 1 & \text{se } t = 0 \\ 0 & \text{se } t = \infty \end{cases}.$$

A representação gráfica de  $S(t)$  é a curva de sobrevivência e pode ser usada para determinar a mediana e outros percentis.

A função densidade, também designada por função densidade probabilidade, do tempo de sobrevivência  $T$  é o limite de probabilidade que um indivíduo falha (morre) num pequeno intervalo (infinitesimal), mais especificamente:

$$f(t) = \lim_{\Delta t \rightarrow 0} \frac{P(t \leq T < t + \Delta t)}{\Delta t} = \frac{dF(t)}{dt} = \frac{d}{dt} [1 - S(t)] = -S'(t) \quad (3).$$

A função densidade é não-negativa e pode assumir valores superiores a 1, uma vez que não se refere a uma probabilidade (Jenkins, 2005).

A função de risco (*hazard function*) é muito utilizada na modelação da distribuição do tempo de vida, constituindo o potencial instantâneo por unidade de tempo da ocorrência do evento, dado que o indivíduo sobreviveu o tempo  $t$ :

$$h(t) = \lim_{\Delta t \rightarrow 0} \frac{P(t \leq T < t + \Delta t | T > t)}{\Delta t} = -\frac{d \ln S(t)}{dt} = \frac{f(t)}{1 - F(t)} = \frac{f(t)}{S(t)} \quad (4).$$

Esta função representa assim a taxa instantânea de ocorrência do evento no instante  $t$ , condicional à sobrevivência até esse instante. A taxa de risco pode variar de zero a infinito e

pode apresentar diversas formas (crescente, decrescente, constante, *bathtub shaped* ou em forma de banheira, entre outras) permitindo de forma flexível retirar algumas conclusões sobre a evolução do padrão de ocorrência ou mortalidade (Lee e Wang, 2003). Convém referir que enquanto que a função de sobrevivência descreve a experiência de sobrevivência, a função de risco descreve a função de morte (ou falha).

A função de risco acumulada, ou Taxa Cumulativa de Risco (*Hazard* cumulativa) é o risco total de falha (ou morte) acumulado até ao tempo  $t$  e é definido da seguinte forma:

$$H(t) = \int_0^t h(x)dx = -\int_0^t \frac{1}{S(x)} \left[ \frac{d}{dx} S(x) \right] dx = -\ln S(t) \quad (5).$$

Esta função tem as seguintes propriedades:

$$H(t) = \begin{cases} 0 & \text{se } t = 0 \\ \infty & \text{se } t = \infty \end{cases}.$$

A relação entre esta função e as anteriormente definidas pode ser descrita por:

$$S(t) = \exp(-H(t)) \quad (6),$$

$$F(t) = 1 - \exp(-H(t)) \quad (7),$$

$$f(t) = h(t) \exp(-H(t)) \quad (8).$$

### 3.2.2.3 – Tipos de modelos

A análise de sobrevivência contínua incorpora diversos tipos de modelos, que podem ser paramétricos, semi-paramétricos ou não-paramétricos (*vide* Figura 3).

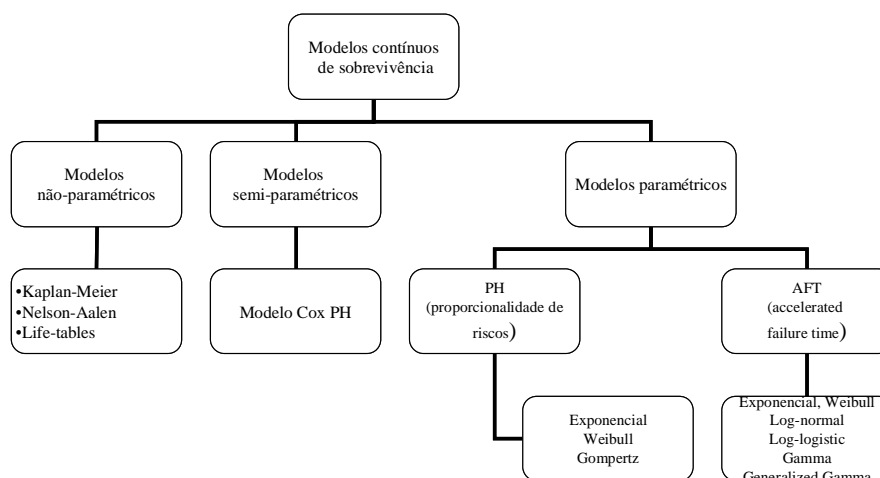


Figura 3 – Tipos de modelos de sobrevivência contínuos

### 3.2.2.3.1– Modelos não-paramétricos

Os modelos não-paramétricos não têm pressupostos acerca da forma das funções de sobrevivência e por isso são muito úteis na fase exploratória da análise dos dados (Cleves *et al*, 2004; Collet, 1994). A função de risco é estimada com base nos dados empíricos de tempo de sobrevivência e no *status* (0 se censurado, 1 se evento ocorreu) do indivíduo. As covariáveis estratificam os dados em grupos e comparam a função de risco de sobrevivência com as funções dos grupos. Uma limitação deste tipo de modelos é que são eficazes apenas para um pequeno número de grupos a apenas se pode analisar uma covariável de cada vez. Além disso, covariáveis contínuas não podem ser analisadas à luz de modelos não-paramétricos, a menos que sejam discretizadas.

Alguns autores, nomeadamente Lee e Wang (2003) consideram que os modelos não-paramétricos são mais eficientes que os modelos paramétricos quando a distribuição teórica adequada não é conhecida, no entanto são menos eficientes que os modelos paramétricos quando o tempo de sobrevivência segue uma distribuição conhecida. Os modelos não-paramétricos podem ser úteis para seleccionar uma distribuição de probabilidade teórica, através da análise das curvas de sobrevivência.

Seja  $t_1 < t_2 < \dots < t_k < \infty$  os tempos observados de falha ou morte e sejam  $t_i$  e  $t_j$  independentes ( $i \neq j$ ). O estimador de Kaplan-Meier (KM) da função de sobrevivência (Kaplan e Meier, 1958) é dado por:

$$\hat{S}(t_j) = \prod_{j|t_j < t} \left( 1 - \frac{d_j}{n_j} \right) \quad (9),$$

Onde  $d_j$  é o número de eventos ocorridos até  $t_j$  e  $n_j$  é o número de indivíduos que apresentam risco de falha (morte) imediatamente a seguir a  $t_j$  (ou seja, o número de indivíduos que sobrevive pelo menos até  $t_j$ ) sendo  $n_j$  dado por:

$$n_j = (m_j + d_j) + (m_{j+1} + d_{j+1}) + \dots + (m_k + d_k) \quad (10).$$

Onde  $m_j$  é o número de indivíduos cujo tempo de sobrevivência é censurado no intervalo  $[t_j + t_{j+1}[$ .



As estimativas de KM para a função de sobrevivência só podem ser estimadas para dados de sobrevivência não censurados (Jenkins 2005, Lee e Wang 2003). Conseqüentemente, a curva de sobrevivência é uma função degrau (ou função *step*) contínua com início no ponto  $t = 0$  (Collet, 1994). Quando existem observações censuradas, a estimativa da função de sobrevivência no momento  $t$  é a proporção de indivíduos vivos até àquele momento (Collet, 1994).

O estimador da função de risco acumulada pode ser derivado através do estimador KM. O estimador da função de risco não pode ser directamente derivado a partir do estimador da função de risco acumulada usando a derivada desta função relativamente a  $t$ . Isto porque a função de risco acumulada é uma função degrau e por isso, o seu declive não é bem definido (Cleves *et al*, 2004, Jenkins, 2005).

Uma alternativa para encontrar o estimador da função de risco é dividir o tempo de sobrevivência em intervalos regulares de tempo e determinar a estimativas da função de risco acumulada (Collet, 1994 e Jenkins, 2005). A função de risco para cada intervalo é então dada por:

$$\hat{h}(t) = \frac{d_j}{n_j \tau_j} \quad (11),$$

onde  $t_j \leq t < t_{j+1}$  e  $\tau_j = t_{j+1} - t_j$  é a amplitude do intervalo  $j$ . De referir que o intervalo não é fechado à direita, por isso esta equação não pode ser usada para determinar o rácio de risco do intervalo que começa no tempo mais longo de falha (morte).

De referir ainda que o estimador de Kaplan-Meier é um estimador de máxima verosimilhança e por isso, sob certas condições, as suas estimativas são consistentes e assintoticamente normais (Lee e Wang, 2003) e pode ter em conta dados censurados e truncados.

No sentido de comparar curvas de sobrevivência, são realizados testes específicos, e não-paramétricos, cuja hipótese nula  $H_0$  invoca que as curvas de sobrevivência são estatisticamente idênticas. Destes, o teste *Log-Rank* é mais potente para alternativas em que as funções de risco (*hazard*) sejam proporcionais. Já o teste de *Breslow* apresenta maior potência para alternativas em que as funções de risco (*hazard*) se cruzam, isto é, quando a mortalidade entre os grupos

estudados não é proporcional. Além disso, a sua potência decresce à medida que o número de casos censurados aumenta.

O estimador de Nelson-Aalen, mais indicado para amostras muito pequenas, estima numa primeira fase a função de risco acumulada e em seguida deriva a função de sobrevivência, processo inverso ao estimador de Kaplan-Meier (Cleves *et al*, 2004). O estimador de Nelson-Aalen para a função de risco acumulada é expresso da seguinte forma:

$$\hat{H}(t_j) = \sum_{j|t_j < t} \left( \frac{d_j}{n_j} \right) \quad (12),$$

e a estimativa derivada da função de sobrevivência (também designada por estimador Fleming-Harrington) é calculada com base na equação (6).

De acordo com Cleves *et al*. (2004) e Jenkins (2005), os estimadores de Kaplan-Meier e de Nelson-Aalen são assintoticamente equivalentes, contudo para amostras de menor dimensão, o segundo estimador produz estimativas mais robustas da função de risco acumulada e o estimador de Kaplan-Meier produz melhores estimativas para a função de sobrevivência.

O estimador *life table* (tabelas de vida) assenta na mesma filosofia do estimador de Kaplan-Meier, no entanto foi desenvolvido essencialmente para resolver problemas onde o número de falhas e o número de indivíduos no grupo de risco estão agrupados em intervalos de tempo. Dado que este trabalho de investigação não utiliza este tipo de dados, este método não será alvo de explanação mais desenvolvida.

### 3.2.2.3.2– Modelos semi-paramétricos

O modelo de Cox PH (ou modelo de riscos proporcionais) é, provavelmente, o modelo de sobrevivência mais popular (Cleves *et al*, 2004). Foi proposto por Cox em 1972 e é um modelo semi-paramétrico, uma vez que a função do risco base é tida como não conhecida e não parametrizada (não se assume qualquer distribuição estatística), o que indica que não é tomado qualquer pressuposto acerca da forma da função de risco (Allison, 2004), que constitui a principal vantagem deste modelo face a modelos paramétricos. Além disso, possibilita estudar a

influência de um alargado número de covariáveis sobre a sobrevivência, as quais agem multiplicativamente sobre o risco, constituindo, esta sim, a parte paramétrica do modelo.

O modelo de Cox PH é robusto uma vez que produz estimativas que são muito similares aos do modelo paramétrico correcto (Cleves *et al*, 2004) e a função de risco deste modelo é definida por:

$$h(t|X) = h_0(t) \exp(\beta' X), \quad h_0(t) \geq 0, \quad X_p \in \mathfrak{R}^p \quad (13),$$

no qual  $\beta = [\beta_1, \dots, \beta_p]$  são coeficientes de regressão das  $p$  covariáveis existentes no modelo e representam o efeito dessas covariáveis na sobrevivência do indivíduo. O valor calculado em  $\exp\{\beta\}$  corresponde à razão de riscos (RR), tal que permite verificar qual risco de morte em que um indivíduo incorre num dado factor, mantendo tudo o resto constante, que influencia o seu tempo de sobrevivência (por exemplo, o risco de morte associado ao facto de fumar face ao facto de não fumar, o risco de morte por cada ano a mais de vida, entre outros).

Este modelo é conhecido por modelo de riscos proporcionais pois a razão das taxas de falha de dois indivíduos é constante no tempo, isto é, a razão das funções de risco para dois indivíduos diferentes,  $i$  e  $j$ , é

$$\frac{h(t|\beta' X_i)}{h(t|\beta' X_j)} = \frac{h_0(t) \exp(\beta' X_i)}{h_0(t) \exp(\beta' X_j)} = \exp(\beta' X_i - \beta' X_j) \quad , \quad i, j = 1, \dots, p \quad (14),$$

e não depende do tempo.

A proporcionalidade de riscos implica que o risco de morte que um indivíduo apresenta no início do estudo manter-se-á igual para todo o período de acompanhamento.

O processo de selecção das variáveis a incluir no modelo tem as seguintes etapas:

1. Ajustar  $p$  modelos com uma covariável;
2. Comparar os modelos obtidos com o modelo nulo;
3. Incluir as covariáveis potencialmente importantes num só modelo;
4. Omitir uma covariável de cada vez e reter no modelo as significativas;
5. As variáveis entretanto excluídas na etapa 2 podem agora, na presença das restantes, revelar-se importantes. Devem incluir-se uma de cada vez e reter as significativas;

6. Verificação final de modo a assegurar que nenhuma covariável possa ser omitida sem levar a um aumento significativo de  $-2LL$ , e nenhuma covariável não incluída leva a uma redução significativa do valor da estatística.

A significância das covariáveis pode ser apurada, entre outros, através do teste Wald (cuja hipótese nula  $H_0: \beta=0$ ).

O modelo de Cox PH apenas pode ser estimado através do método de verosimilhança parcial, o que permite estimar os coeficientes sem assumir uma distribuição de probabilidade teórica para a função de risco de base ( $h_0(t)$ ).

Existem pressupostos-chave no modelo de Cox PH, refiram-se:

- Taxas de risco proporcionais;
- A influência das covariáveis na função de risco (*hazard*) não sofre qualquer alteração durante o período em que os indivíduos estão em observação;
- Podem ser analisadas variáveis contínuas e categóricas;
- Não é necessário conhecer  $h_0(t)$  para estimar o coeficiente de cada covariável, e consequentemente a *hazard ratio*;
- $h_0(t)$  não é especificada – pode estimar-se no caso de ser necessário estimar  $S(t)$ ;
- Não se podem considerar covariáveis que mudem ao longo do tempo (a maior parte das covariáveis (ex: sexo, etnia, ...) são constantes).

### 3.2.2.3.3– Modelos paramétricos

Os modelos paramétricos assumem que a distribuição de probabilidades dos dados é conhecida e por isso pode ser previamente considerada pelo investigador. Existem várias distribuições que podem ser utilizadas neste tipo de modelos, nomeadamente: exponencial, Weibull, Gompertz, log-normal, log-logistic, gamma e gamma generalizada.

Os modelos paramétricos facilmente “acomodam” dados censurados dos três tipos, enquanto que o modelo de Cox apresenta dificuldades nos dados censurados à esquerda e censura intervalar.

De acordo com Box-Steffensmeier e Jones (2004) os modelos paramétricos apresentam algumas vantagens face aos modelos semi-paramétricos, mais concretamente: (i) permitem analisar o efeito da dependência da sobrevivência; (ii) a realização de predição é facilitada; (iii) produzem desvios-padrão pequenos para os coeficientes das covariáveis dada a eficiência obtida com o uso do estimador de máxima verosimilhança que utiliza toda a informação acerca do tempo de sobrevivência, ao contrário da verosimilhança parcial que usa apenas a ordem do tempo de verosimilhança.

De seguida são apresentados sumariamente os modelos paramétricos mais utilizados.

### **Modelo Exponencial**

O modelo exponencial pode acomodar as formas PH (proporcionalidade dos riscos) e AFT (*accelerated failure time*). Estes modelos são similares, ainda que escritos de forma diferente. O modelo exponencial é o modelo paramétrico mais simples porque a função risco da distribuição exponencial é constante ao longo do tempo (Collet, 1994), por isso não reflecte a dependência da sobrevivência. Esta dependência ocorre quando as taxas de risco (*hazard rates*) variam com o tempo de sobrevivência e todas as outras distribuições de probabilidade têm em conta este facto, pelo que apresentam maior flexibilidade que a exponencial.

Neste sentido, o modelo exponencial é algo restritivo e por vezes pouco apropriado em muitas situações concretas (Box-Steffensmeier e Jones, 2004, Lee e Wang, 2003). A função de risco do modelo exponencial é dada por:

$$h(t|X) = \lambda, \quad \lambda > 0, t > 0 \quad (15).$$

Quando  $\lambda = 1$ , é obtido um modelo exponencial unitário.

Nos modelos de sobrevivência paramétricos, o parâmetro  $\lambda$  é reparametrizado com vista a incorporar as covariáveis. Considera-se que  $\lambda = \exp(\beta'X)$  na forma PH e  $\lambda = \exp(-\beta'X)$  na forma AFT – *Accelerated failure time* (Cleves *et al*, 2004).

A função sobrevivência do modelo exponencial é:

$$S(t|X) = \exp(-\lambda X), t > 0 \quad (16)$$

e a função densidade é:

$$f(t|X) = \lambda \exp(-\lambda X), t > 0 \quad (17),$$

sendo a função de risco acumulada:

$$H(t|X) = \lambda t, t > 0 \quad (18).$$

### Modelo de Weibull

A distribuição de Weibull é uma das mais utilizadas nos modelos paramétricos da análise de sobrevivência, sendo de evidenciar exemplos de mortalidade humana e desemprego (Lee e Wang, 2003). Tal como o modelo anterior, este modelo pode ser apresentado nas formas PH e AFT. A distribuição de Weibull tem dois parâmetros,  $\lambda$  e  $p = 1/\sigma$ , que são respectivamente os parâmetros de escala e de forma (Box-Steffensmeier e Jones, 2004).

A função de risco do modelo PH Weibull é dada por<sup>1</sup>:

$$h(t|X) = \lambda p t^{p-1}, \lambda > 0, p > 0 \quad (19)$$

e a respectiva AFT é

$$h(t|X) = \lambda p (\lambda t)^{p-1}, \lambda > 0, p > 0 \quad (20).$$

A função de risco do modelo de Weibull é monotonamente crescente (decrecente) caso  $p > 1$  ( $p < 1$ ). Caso  $p = 1$  trata-se da distribuição exponencial, o que significa que o modelo exponencial é um caso especial do modelo de Weibull.

A função de sobrevivência do modelo PH Weibull é dada por:

$$S(t|X) = \exp(-\lambda t^p) \quad (21),$$

a respectiva função densidade é designada da seguinte forma:

---

<sup>1</sup> Por questão de simplicidade de apresentação da matéria, são apresentadas apenas as funções da forma PH.

$$f(t|X) = \lambda p t^{p-1} \exp(-\lambda t^p) \quad (22).$$

Por fim, a função de risco acumulada do modelo PH de Weibull pode ser definido por:

$$H(t|X) = \lambda t^p \quad (23).$$

O modelo de Weibull não é considerado em diversas situações, uma vez que permite que as taxas de risco (*hazard rate*) mudem apenas numa direcção ao longo do tempo, o que constitui uma limitação se se comparar com os modelos log-normal e log-logístico.

### Modelo de Gompertz

Embora não tenha sido considerado no presente trabalho de investigação, refira-se que distribuição de Gompertz é muito utilizada na investigação médica e biológica em estudo de mortalidade dos indivíduos (Cleves *et al.*, 2004) e também em estudos politico-demográficos (Box-Steffensmeier e Jones 2004). Este tipo modelo assume proporcionalidade dos riscos.

### Modelo Log-normal

O modelo de sobrevivência log-normal tem sido usado em inúmeros estudos de economia e medicina (Lee e Wang, 2003). A função de risco deste modelo é definida da seguinte forma:

$$h(t|X) = \frac{\frac{1}{\sigma\sqrt{2\pi}} t^{-1} \exp\left\{-\frac{1}{2}\left[\frac{\ln(t\lambda)}{\sigma}\right]^2\right\}}{1 - \Phi\left[\frac{\ln(t\lambda)}{\sigma}\right]}, \quad \sigma > 0, t > 0 \quad (24),$$

onde  $\Phi$  é a função distribuição da distribuição normal padrão e  $\sigma = p^{-1}$ . Esta função de risco é caracterizada por dois parâmetros de escala (média e desvio-padrão). Uma das vantagens deste modelo é que o risco não é monótono. O parâmetro  $\sigma$  indica a rapidez com que a taxa de risco chega ao pico. A função de sobrevivência do modelo log-normal é dada por:

$$S(t|X) = 1 - \Phi\left[\frac{\ln(t\lambda)}{\sigma}\right] \quad (25),$$

sendo a função densidade definida da seguinte forma:

$$f(t|X) = \frac{1}{\sigma\sqrt{2\pi}} t^{-1} \exp\left\{-\frac{1}{2}\left[\frac{\ln(t\lambda)}{\sigma}\right]^2\right\} \quad (26)$$

e a função de risco acumulada:

$$H(t|X) = -\log\left\{1 - \Phi\left[\frac{\ln(t\lambda)}{\sigma}\right]\right\} \quad (27).$$

### Modelo Log-logístico

Embora não tenha sido considerado no presente trabalho de investigação refira-se que o modelo log-logístico de sobrevivência é muito popular e a sua forma é relativamente similar à log-normal, excepto na cauda da distribuição (log-logística tem cauda mais pesada que a log-normal).

### Modelo Gamma

Não tendo sido considerado para o presente trabalho, refira-se que a distribuição gamma é muito utilizada em estudos sobre a sobrevivência humana (Lee e Wang, 2003).

Tanto a distribuição de Weibull como a gamma são generalizações da distribuição exponencial, por isso os modelos de sobrevivência são relativamente similares. A distribuição gamma tem dois parâmetros ( $\lambda$  e  $k$ ), que são respectivamente os parâmetros de escala e de forma.

### Modelo Gamma Generalizado

Este modelo é talvez o mais flexível, no contexto dos modelos paramétricos, uma vez que é caracterizado por três parâmetros livres (um de escala,  $\lambda$ , e dois de forma,  $p$  e  $k$ ). O termo “generalizado” tem a ver com o facto de se reduzir a outras distribuições perante valores concretos dos parâmetros, nomeadamente se  $k = 1$  converge para a distribuição de Weibull. Caso  $k = p = 1$  reduz para a distribuição exponencial, se  $k = 0$  converge para a log-normal e caso  $p = 1$  obtém-se a distribuição gamma estandardizada.

### Estimação e selecção dos modelos paramétricos

Todos os modelos paramétricos podem ser estimados através de método da máxima verosimilhança (MLE) (Box-Steffensmeier e Jones, 2004), uma vez que este método tem em



consideração dados censurados e sob certos pressupostos, as suas estimativas são assintoticamente não-enviezadas e eficientes.

A selecção da distribuição paramétrica mais adequada é de extrema importância, uma vez que, se seleccionada uma distribuição que não a mais correcta, as estimativas obtidas incorrem em erro (Box-Steffensmeier e Jones, 2004). Caso o investigador tenha uma ideia clara acerca da distribuição dos dados, então a estimação paramétrica afigura-se adequada, contudo, se a distribuição lhe é desconhecida, será preferível utilizar um modelo semi-paramétrico para avaliar a forma da distribuição e verificar se o pressuposto da proporcionalidade dos riscos é satisfeito (Cleves *et al*, 2004).

De acordo com Cleves *et al*. (2004), podem ser testadas as seguintes hipóteses utilizando como base a distribuição gamma generalizada para os modelos acoplados (*nested models*):

$H_0 : k = 1 \wedge p = 1$  (tempo de sobrevivência segue uma distribuição exponencial)

$H_0 : k = 1$  (tempo de sobrevivência segue uma distribuição de Weibull)

$H_0 : k = 0$  (tempo de sobrevivência segue uma distribuição log-normal)

$H_0 : p = 1$  (tempo de sobrevivência segue uma distribuição gamma).

A hipótese nula pode ser testada através de dois testes equivalentes: razão de verosimilhanças (*likelihood-ratio*) e teste de Wald (Cleves *et al*, 2004, Lee e Wang, 2003). Ainda assim, há que ter em conta que nenhum destes testes é válido para dados não acoplados (*non-nested*) (Box-Steffensmeier e Jones, 2004). A alternativa é utilizar os critérios AIC (Akaike Information Criterion) e BIC (Bayesian Information Criterion).

Graficamente também se consegue detectar que distribuição melhor adequa aos dados em estudo. Se representada a curva de Kaplan-Meier as distribuições mais adequadas são as que mais se aproximem dessa curva.

A análise da bondade de ajustamento dos modelos baseado nos resíduos de Cox-Snell também permite seleccionar entre modelos paramétricos (Box-Steffensmeier e Jones, 2004). Os resíduos

de Cox-Snell comportam-se como uma exponencial unitária se o modelo estiver bem ajustado, tal que se deve confirmar se aqueles se encontram aproximadamente sobre uma recta.

### 3.2.3 – Modelos de fragilidade

Os modelos de sobrevivência focados até aqui, assumem que as diferenças entre indivíduos são perfeitamente incluídas na covariáveis, ou seja, em todos os restantes aspectos os indivíduos em estudo são similares (é assumida a independência entre os indivíduos). No entanto, este pressuposto é extremamente forte e simultaneamente pouco realista, pois a possibilidade de existirem variáveis omitidas é muito elevada (Hosmer e Lemeshow, 1999). Tal pode acontecer porque, por vezes, é impossível incluir todas as variáveis que caracterizam os indivíduos, porque algumas variáveis não são conhecidas e porque outras podem simplesmente ser não-mensuráveis.

Neste contexto, a heterogeneidade não-observada deve ser incluída no modelo. Esta heterogeneidade não-observada constitui não mais que uma fragilidade para os indivíduos ou grupos em estudo.

Os modelos de fragilidade são, assim, caracterizados pela introdução de um efeito aleatório na função de risco (heterogeneidade não-observada). São então incorporados os chamados componentes de fragilidade. A forma mais usual de se introduzir uma fragilidade no modelo de Cox é:

$$h(t|X) = h_0(t) \mu \exp(\beta' X) \quad (28),$$

onde  $\mu$  é a fragilidade, que tem média 1 e variância  $\delta^2$  desconhecida (para a qual se assume a distribuição gamma). Note-se que  $\mu = 0$  se o modelo de fragilidade se reduzir ao modelo de riscos proporcionais.

Neste trabalho de investigação não serão utilizados empiricamente modelos de fragilidade, pelo que não se desenvolve mais a sua apresentação teórica.

### 3.2.4 – Análise dos resíduos e adequabilidade do modelo

O diagnóstico dos modelos de sobrevivência deve ter em consideração a forma funcional das covariáveis, a validação do pressuposto da proporcionalidade dos riscos (confirmado pelo Teste

de Rho, ou graficamente no caso das covariáveis categóricas), a análise da bondade do ajustamento do modelo (*goodness-of-fit*), onde se pode considerar, para além dos resíduos, o Índice de Prognóstico, e a identificação de *outliers* e influência de observações.

O Índice de Prognóstico (IP) constitui o preditor linear do modelo de Cox,  $x\beta$ , calculado para cada indivíduo utilizando as covariáveis observadas e as estimativas dos coeficientes de regressão do modelo ajustado. Assim, os indivíduos são estratificados em grupos de tamanhos aproximadamente iguais (com alto, médio e baixo IP), e os valores médios de cada uma das covariáveis dentro de cada grupo originam curvas de sobrevivência sob o modelo ajustado. Se o modelo for razoável, no gráfico, as curvas ajustadas pelo modelo em cada estrato estarão próximas das estimadas por Kaplan-Meier (Carvalho, 2008) [14].

A análise do ajustamento dos modelos de regressão é baseada na comparação entre os valores observados ( $y_i$ ) e os estimados da variável dependente  $\left(\hat{y}_i\right)$ , em que os resíduos ( $\varepsilon_i$ ) são dados por:

$$\varepsilon_i = y_i - \hat{y}_i \quad (29).$$

Apesar da maioria das metodologias de diagnóstico terem sido desenvolvidos no contexto do Modelo de Cox PH, podem ser aplicadas para modelos semi-paramétricos e paramétricos (Box-Steffensmeier e Jones, 2004). Os principais métodos de análise da adequabilidade de modelos baseados em resíduos são: resíduos de Cox-Snell, resíduos de Schoenfeld, resíduos Scaled Schoenfeld, *martingale residuals* e resíduos *deviance*.

### Resíduos de Cox-Snell

Os resíduos de Cox-Snell para cada indivíduo  $i$  ( $r_{cs_i}$ ) são definidos da seguinte forma:

$$r_{cs_i} = -\log \hat{S}_i(t | X) = \hat{H}_i(t | X) \quad (30).$$

Podem ser usados para analisar a bondade do ajustamento de qualquer modelo paramétrico ou semi-paramétrico (Box-Steffeensmeier e Jones, 2004) (Box-Steffeensmeier e Jones, 2004, Cleves *et al*, 2004, Lee e Wang, 2003). Estes resíduos não estão distribuídos de forma simétrica em torno de zero e não podem assumir valores negativos (Collet, 1994).

### Resíduos de Schoenfeld

Estes resíduos constituem a diferença entre o valor observado e o valor esperado de cada covariável para cada indivíduo em cada momento de falha (morte) (Box-Steffeensmeier e Jones, 2004, Lee e Wang, 2003). Neste contexto, não faz sentido a sua utilização em dados censurados. A equação com a qual se obtém o resíduo do individuo  $i$  da covariável  $k$  é dada por (Hosmer e Lemeshow, 1999):

$$r_{S_{ik}} = \frac{\partial L}{\partial \beta_k} = \delta_i (X_{ik} - \bar{X}_{w_i,k}) = \delta_i \left( X_{ik} - \frac{\sum_{j \in R(t_j)} X_{jk} \exp(\beta' X)}{\sum_{j \in R(t_j)} \exp(\beta' X)} \right) \quad (31),$$

onde  $X_{ik}$  é o valor da covariável  $k$  para o indivíduo  $i$  e  $R(t_j)$  é o conjunto de risco no momento  $t_j$ .

Os resíduos de Schoenfeld não exibem autocorrelação e assintoticamente têm média nula (Lee e Wang, 2003). O pressuposto de proporcionalidade dos riscos (PH) pode ser verificado através destes resíduos, uma vez que a dependência temporal dos coeficientes das covariáveis pode ser examinado graficamente, o que torna a sua análise relativamente subjectiva (Cleves *et al*, 2004).

### Resíduos Scaled Schoenfeld

Estes resíduos são especialmente úteis para avaliar o pressuposto de proporcionalidade dos riscos e são obtidos da seguinte forma (Hosmer e Lemeshow, 1999):

$$r_{S_{ik}}^* = \left[ V(r_{S_{ik}}) \right]^{-1} r_{S_{ik}} \quad (32).$$

Na equação (2.43)  $V(r_{S_{ik}})$  é o estimador da matriz de covariância do vector dos resíduos de Schoenfeld.

### Resíduos Martingala

Os resíduos martingala constituem a diferença entre o observado e o valor esperado de eventos para um indivíduo baseado no modelo estimado. São expressos da seguinte forma:

$$M_i = N_i(t) - \hat{H}_i(t | X) \quad (33),$$

onde  $N_i(t)$  é o número observado de eventos para cada indivíduo  $i$  no momento  $t$  e  $\hat{H}_i(t | X)$  é o respectivo valor esperado baseado no modelo previamente estimado. Caso o evento de interesse apenas ocorra uma vez (por exemplo a morte) então a equação (44) é equivalente a:

$$M_i = \delta_i - \hat{H}_i(t | X) = \delta_i - r_{cs_i} \quad (34).$$

Assintoticamente estes resíduos apresentam valor esperado nulo e ausência de autocorrelação, sendo inclusivamente o seu somatório nulo em amostras de grande dimensão (Collet, 1994) e apresentam especial utilidade para avaliar e examinar a forma das covariáveis do modelo (Box-Steffensmeier e Jones, 2004, Cleves *et al*, 2004, Collet, 1994, Hosmer e Lemeshow, 1999).

Graficamente, a representação dos resíduos martingale permite a detecção de eventual padrão nos dados, a detecção de possíveis *outliers*, e, se representados versus covariável permitem avaliar a forma funcional dessa covariável.

### Resíduos Deviance

Os resíduos *deviance* foram introduzidos e explorados por Therneau et al. (1990) e constituem uma transformação dos resíduos martingala com vista a obter uma distribuição mais normalizada quando o modelo adequado é especificado (Collet, 1994).

Quando menos de um quarto das observações é censurada, os resíduos deviance assumem uma distribuição muito semelhante à normal, mas nos casos em que a percentagem de censurados é elevada (> 40%) existem distorções à distribuição normal que podem afectar os resultados obtidos. Estes desvios são nulos quando os resíduos martingala são nulos também e são sempre negativos para observações censuradas. O gráfico com a representação destes resíduos versus preditor linear é usualmente utilizado para a detecção de observações afastadas que constituam por isso candidatos a *outliers*.

## 4. Análise exploratória dos dados

Neste capítulo serão analisados os dados recolhidos sobre os utentes que foram seguidos no Hospital de Beja entre os anos 2000 a 2009, no âmbito da infecção VIH/ SIDA.

Começar-se-á por caracterizar a base de dados em estudo, não apenas quanto à sua dimensão e as variáveis que a constituem, mas também, e descritivamente, o que revelam da realidade que visam representar.

### 4.1 – Caracterização da base de dados

A amostra é constituída por 302 indivíduos infectados com VIH/ SIDA, seguidos ao longo de 10 anos no Hospital de Beja.

A Tabela 2 apresenta as variáveis em estudo e respectiva codificação.

Nome	Descrição
<b>id</b>	Identificação do indivíduo
<b>Sexo</b>	1= Feminino; 2= Masculino
<b>AnoNasc</b>	Ano de nascimento do indivíduo (1919 a 1988)
<b>IdadeActAct</b>	Idade actual do indivíduo activo (dos 23 aos 90 anos)
<b>AnoEnt</b>	Ano de entrada no acompanhamento (2000 a 2009)
<b>IdadeEnt</b>	Idade do indivíduo aquando da entrada no acompanhamento do Hospital de Beja (dos 17 aos 86 anos)
<b>AnoDiag</b>	Ano do diagnóstico (1987 a 2009)
<b>IdadeDiag</b>	Idade do indivíduo aquando do diagnóstico da infecção (dos 16 aos 86 anos)
<b>TipoVirus</b>	Tipo de Vírus (1= VIH 1; 2= VIH 2; 3= VIH 1/2)
<b>EstCivil</b>	Estado Civil do indivíduo (1= Casado; 2= Solteiro; 3= Divorciado; 4= Viúvo)
<b>Profiss</b>	Profissão do indivíduo
<b>SectAct</b>	Sector de Actividade do indivíduo (1 = Sector Primário e/ou Secundário; 3= Sector Terciário)
<b>ConcRes_</b>	Concelho de Residência do indivíduo (1 – Aljustrel; 2 – Almodôvar; 3 – Alvito; 4 – Barrancos; 5 – Beja; 6 – Castro Verde; 7 – Cuba; 8 – Ferreira do Alentejo; 9 – Mértola; 10 – Moura; 11 – Odemira; 12 – Ourique; 13 – Serpa; 14 – Vidigueira; 15 – Outros)
<b>Distrito</b>	Distrito de Residência do indivíduo (1- Beja; 2 – Coimbra; 3 – Évora; 4 – Faro; 5 – Lisboa; 6 – Portalegre; 7 – Setúbal)
<b>DistritoBeja</b>	Distrito de Residência do indivíduo (1- Beja; 2 – Outros – Não Beja)
<b>PaísOr</b>	País de origem do indivíduo (1 – Portugal; 2 – Angola; 3 – Cabo Verde; 4 – Cuba; 5 – E.U.A.; 6 – Espanha; 7 – Guiné-Bissau; 8 – Moçambique; 9 – República Dominicana)
<b>PaísOrPort</b>	País de Origem do indivíduo (1- Portugal; 2 – Outros – Não Portugal)

<b>Toxicodep</b>	Toxicodependente (1= Sim; 2= Não)
<b>Heteross</b>	Heterossexual (1= Sim; 2= Não)
<b>HomoBissex</b>	Homossexual/Bissexual (1= Sim; 2= Não)
<b>TransfSang</b>	Transfusão de Sangue (1= Sim; 2= Não)
<b>GrupRisco</b>	Grupo de Risco (1= Heterossexual; 2= Heterossexual e Transfusão de Sangue; 3= Homo/Bissexual; 4= Toxicodependente Heterossexual; 5= Toxicodependente Homo/Bissexual)
<b>Recluso</b>	1= Sim; 2= Não
<b>Tubercul</b>	Tuberculose – 1= Sim; 2= Não
<b>Hep B</b>	Hepatite B – 1= Sim; 2= Não
<b>Hep C</b>	Hepatite C – 1= Sim; 2= Não
<b>DepNerv</b>	Depressão Nervosa – 1= Sim; 2= Não
<b>Álcool</b>	Alcoolismo – 1= Sim; 2= Não
<b>Tabag</b>	Tabagismo – 1= Sim; 2= Não
<b>CatClinic</b>	Categoria Clínica aquando da data de entrada (1= A/B; 2= C)
<b>ContCD4</b>	Contagem de células CD4 pelo valor mais baixo até à data de entrada (4 a 1368)
<b>CatCD4</b>	Categoria para contagem de células CD4 (1= “>=500”; 2= “200-499”; 3= “<200”)
<b>EstadoAct</b>	Estado Actual do indivíduo (1= Abandono; 2= Activo; 3= Falecido 4= Transferido)
<b>Status</b>	0= censura; 1= óbito (na sequência da infecção VIH/SIDA)
<b>AnoFalec</b>	Ano do falecimento (2002 a 2009)
<b>CausaFalec</b>	Causa do falecimento
<b>CausaFalecSida</b>	Falecimento na sequência da infecção VIH/ SIDA (1= Sim; 2= Não)
<b>TempSobrTF</b>	Tempo de Sobrevivência Total desde nascimento até óbito – só Falecidos (25 a 83 anos)
<b>TempSobrDF</b>	Tempo de Sobrevivência desde diagnóstico até óbito – só Falecidos (0 a 17 anos)
<b>TempSobrTD</b>	Tempo de Sobrevivência, em dias, do total de indivíduos desde a data de diagnóstico até à data de abandono, transferência, óbito ou data actual consoante o estado actual do indivíduo (até 8220 dias)
<b>TAR</b>	Início de Terapêutica Antiretroviral (1= Sim; 2= Não)
<b>AnoTAR</b>	Ano de início da TAR (1991 a 2009)
<b>IdadeTAR</b>	Idade com que o indivíduo iniciou TAR (20 aos 83 anos)
<b>GupoAR</b>	Grupo/ Classe dos antiretrovirais (1= NITR; 2= NITR, IP; 3= NITR, IP, II; 4= NITR, NNITR)
<b>CargaV</b>	Carga Viral após início da TAR (1= Indetectável; 2= Detectável)
<b>InterrpTAR</b>	Interrupção da TAR por variados motivos (1= Sim; 2= Não)
<b>ConsEnf</b>	Frequência das consultas de Enfermagem (1= Sim; 2= Não)
<b>FaltaConsEnf</b>	Uma ou mais faltas a consultas de Enfermagem programadas por variados motivos (1= Sim; 2= Não)
<b>ConsPsic</b>	Frequência das consultas de Psicologia (1= Sim; 2= Não)
<b>FaltaConsPsic</b>	Uma ou mais faltas a consultas de Enfermagem programadas por variados motivos (1= Sim; 2= Não)

Tabela 2 – Variáveis contituíntes da base de dados

De referir é também o facto de os dados não estarem completos em todas as variáveis, isto é, existem campos em branco nos casos em que não se conseguiu extrair a informação desejada relativamente a um determinado indivíduo. Esta existência de dados omissos ou dados *missing* deveu-se à falta de registo nas fontes de dados utilizadas e/ou registos incongruentes entre as várias fontes, donde se optou por deixar o campo em branco. Ainda assim, os casos de dados em falta estão devidamente identificados e além disso não são em número tão elevado que impossibilitem o estudo dos restantes dados disponíveis.

Considera-se que os dados que constituem a amostra são representativos do impacto da infecção VIH/ SIDA no Baixo Alentejo, como é objectivo do presente estudo, no sentido em que constam de todos os utentes que estão (ou estiveram) a ser seguidos na região. Por conseguinte, e assumindo que existirão muitos outros indivíduos não diagnosticados e que deveriam estar a ser seguidos no Hospital, aceita-se que os resultados da análise aos dados disponíveis possam inferir, genericamente, acerca de qualquer indivíduo com a doença a residir na região. Para além disso, apurar-se-ão também as semelhanças com os dados nacionais sobre a realidade da infecção, de forma a confrontar resultados e apurar o efectivo grau de representatividade da amostra, não só ao nível regional mas também ao nível nacional.

## **4.2 – Análise Descritiva e inferencial**

Como não existem quaisquer estudos sobre as características do doente com VIH/ SIDA no Baixo Alentejo, esta primeira análise descritiva dos dados é tão importante quanto a análise inferencial que a sucede. De facto, para quem quer conhecer estes doentes, e tomar decisões acerca do seu acesso aos serviços de saúde, os esquemas de tratamento, ou até mesmo desenvolver programas de sensibilização, informação e prevenção, importa acima de tudo conhecer as suas características gerais. Quem são, que idade têm, onde vivem e o que fazem são informações essenciais.

### **4.2.1 – Sexo, Idade e Estado Civil**

A Figura 4 evidencia a forma como os indivíduos em estudo se distribuem por sexo.



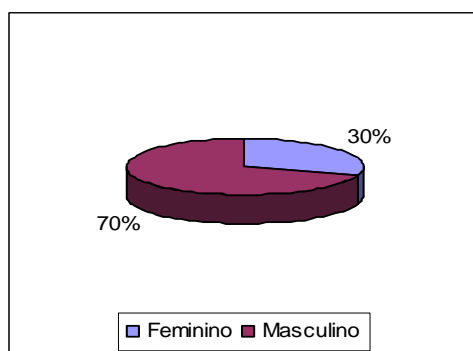


Figura 4 – Distribuição da amostra por sexo

Dos 302 doentes em estudo a maior parte são homens, tal que estes representam cerca de 70% da amostra (esta percentagem está em concordância com a distribuição por sexos verificada no País em geral).

Com vista a verificar se a proporção de doentes do sexo masculino é efectivamente superior à proporção de doentes do sexo feminino, pode fazer-se um teste de hipóteses para a diferença de proporções. Dado o tamanho da amostra, pode utilizar-se a distribuição normal (Teorema do Limite Central). As hipóteses a colocar seriam:

$$H_0 : p_H = p_M ,$$

$$H_1 : p_H > p_M$$

onde  $p_H$  é a proporção de homens e  $p_M$  é a proporção de mulheres. Pela estatística do teste  $Z=6,441$ , e assumindo um nível de significância de 1% verifica-se que a hipótese nula deve ser rejeitada, pelo que se conclui que a proporção de doentes homens é superior à proporção de doentes mulheres.

A Tabela 3 apresenta os resultados de várias medidas amostrais no que respeita às variáveis idade na entrada, idade no diagnóstico e idade actual.

Medidas Amostras	Idade na entrada			Idade no diagnóstico			Idade Actual (Activos)		
	Fem.	Masc.	Total	Fem.	Masc.	Total	Fem.	Masc.	Total
Média	34.01	36.79	<b>35.96</b>	30.99	33.97	<b>33.07</b>	39.06	41.05	<b>40.33</b>
IC a 95% para a Média – Limite Inferior	31.83	35.35	<b>34.75</b>	28.57	32.42	<b>31.76</b>	36.07	39.13	<b>38.71</b>
IC a 95% para a Média – Limite Superior	36.20	38.23	<b>37.16</b>	33.41	35.51	<b>34.37</b>	42.04	42.98	<b>41.95</b>
Mediana	32.00	35.00	<b>34.00</b>	28.00	32.00	<b>31.00</b>	36.50	40.00	<b>39.00</b>
Moda	32.00	29.00	<b>29.00</b>	25.00	37.00	<b>25.00</b>	42.00	35.00	<b>35.00</b>
Variância	107.69	111.17	<b>111.39</b>	128.97	124.07	<b>126.97</b>	114.92	85.30	<b>96.26</b>
Desvio-Padrão	10.38	10.54	<b>10.55</b>	11.36	11.14	<b>11.27</b>	10.72	9.24	<b>9.81</b>
Coefficiente de Variação	0.31	0.29	<b>0.29</b>	0.37	0.33	<b>0.34</b>	0.27	0.22	<b>0.24</b>
Mínimo	17.00	19.00	<b>17.00</b>	16.00	19.00	<b>16.00</b>	25.00	23.00	<b>23.00</b>
Máximo	71.00	86.00	<b>86.00</b>	71.00	86.00	<b>86.00</b>	72.00	90.00	<b>90.00</b>
Amplitude	54.00	67.00	<b>69.00</b>	55.00	67.00	<b>70.00</b>	47.00	67.00	<b>67.00</b>
Percentil 10	23.00	26.00	<b>25.00</b>	20.80	23.00	<b>22.00</b>	27.30	32.20	<b>29.00</b>
Percentil 25	27.00	29.00	<b>29.00</b>	24.00	26.00	<b>25.00</b>	32.00	35.00	<b>34.00</b>
Percentil 75	39.00	41.00	<b>40.00</b>	36.00	38.00	<b>37.50</b>	44.50	46.00	<b>45.00</b>
Percentil 90	44.00	49.00	<b>48.00</b>	44.40	47.10	<b>45.00</b>	52.10	49.00	<b>49.60</b>
Amplitude Interquartil	12.00	12.00	<b>11.00</b>	12.00	12.00	<b>13.00</b>	13.00	11.00	<b>11.00</b>
Assimetria	1.54	1.67	<b>1.59</b>	1.68	1.84	<b>1.76</b>	1.32	1.82	<b>1.56</b>
Erro-Padrão (Assimetria)	.26	.17	<b>.14</b>	.26	.17	<b>.14</b>	.33	.25	<b>.20</b>
Curtose	3.05	4.59	<b>4.01</b>	3.09	5.02	<b>4.26</b>	2.07	7.93	<b>4.82</b>
Erro-Padrão (Curtose)	.51	.34	<b>.28</b>	.51	.34	<b>.29</b>	.65	.50	<b>.40</b>

Tabela 3 – Medidas amostrais quanto á idade na entrada, idade no diagnóstico e idade actual

Como observável na tabela acima verifica-se que, em média, o doente foi diagnosticado com a infecção aproximadamente 3 anos antes de ter entrado para o acompanhamento na instituição. Esse aspecto verifica-se em ambos os sexos, e além disso pode verificar-se que as mulheres foram diagnosticadas e entraram para a consulta mais novas do que o sexo oposto.

O doente foi assim, e em média, diagnosticado com 33 anos (desvio-padrão,  $s=11$  anos), entrou para a consulta com 36 anos (com  $s=11$  anos), e se actualmente o seu estado é activo então terá cerca de 40 anos (com  $s=10$  anos), o que quer dizer que este indivíduo está a viver com a doença há cerca de 7 anos.

O valor do desvio-padrão nas médias referidas é de facto elevado. Como se tratam de distribuições diferentes, com médias diferentes, analisa-se o coeficiente de variação donde se verifica que para todas elas a dispersão é forte, ou seja, os dados não são totalmente

homogéneos, tal que, os desvios em relação à média perfazem cerca de 29%, 34% e 24% dessa medida de localização central nas três distribuições em estudo: idade de entrada, idade diagnóstico e idade actual, respectivamente. A maior dispersão face à média verifica-se na idade de diagnóstico. Se analisarmos por sexo, o coeficiente de variação é sempre superior nos valores obtidos para as mulheres. Ainda assim estes são os dados reais e por isso não devem ser desprezados.

Uma possível solução é considerar também os valores que aparecem mais vezes através da análise da moda. Nesse aspecto, a maior parte dos indivíduos foi diagnosticada com 25 anos, entrou para a consulta com 29 anos, e se actualmente o seu estado é activo, tem 35 anos.

Quanto aos valores máximos observados na amostra, verifica-se que o indivíduo mais velho foi diagnosticado com 86 anos, é do sexo masculino e entrou para a consulta nesse mesmo ano. Se não tiver ainda falecido tem 90 anos.

Por outro lado, existem indivíduos do sexo feminino que foram diagnosticados com 16 anos, tendo entrado para a consulta no ano seguinte. Actualmente, os indivíduos mais novos a serem seguidos no hospital através da consulta têm 21 e 25 anos, para o sexo masculino e feminino, respectivamente.

Analisando mais pormenorizadamente as distribuições das variáveis em estudo, verifica-se que:

1. A média não é de facto representativa. Os dados não seguem distribuição normal, como é visível pela observação dos respectivos histogramas (*vide* Figura 5). De facto, para as variáveis que representam as idades em estudo, os respectivos dados não se distribuem de forma simétrica, revelando em vez disso, enviesamento para a direita ou assimetria positiva (coeficiente de assimetria  $> 0$ ), donde a média  $>$  mediana  $>$  moda em todas as distribuições;
2. Quanto ao grau de achatamento, é identificável pela observação dos histogramas (Figura 5) e corroborado pelos valores de curtose obtidos que todas as distribuições são do tipo leptocúrtico, tal que as frequências revelam maior intensidade na vizinhança dos valores centrais;

3. Para além do referido, da análise aos gráficos caixas de bigode representados na Figura 6, detecta-se que existem algumas observações que estão a influenciar de forma tendencial a média. Em todas as distribuições (no geral e mesmo quando analisadas por sexo), existem observações extremas que parecem estar a inflacionar a média. São *ouliers* moderados e *ouliers* severos. Uma curiosidade acerca destes últimos é que correspondem maioritariamente a indivíduos do sexo masculino.

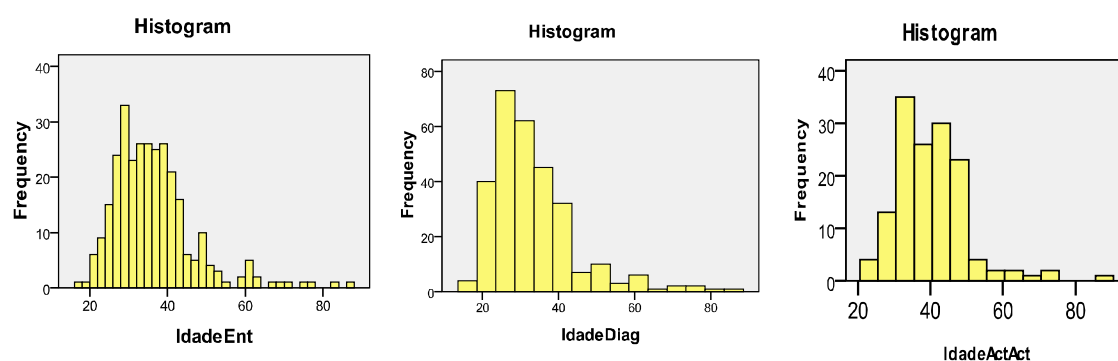


Figura 5 – Histogramas das variáveis: idade na entrada, idade no diagnostico e idade actual

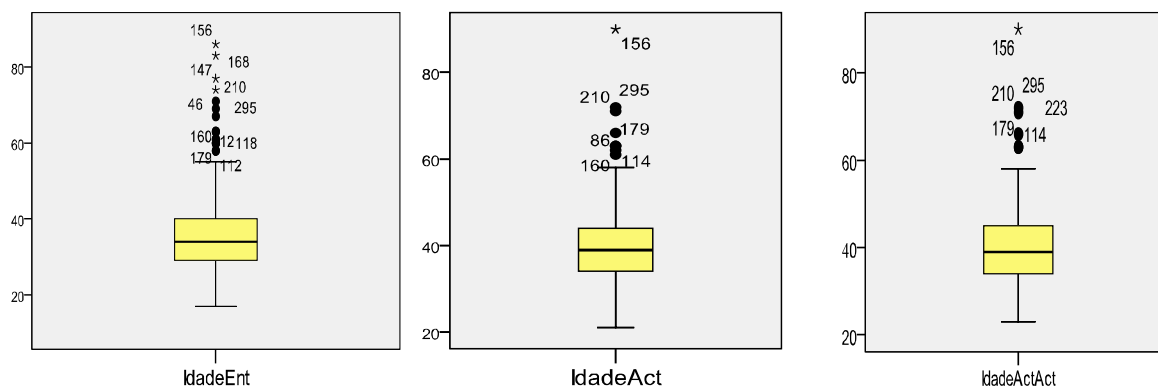
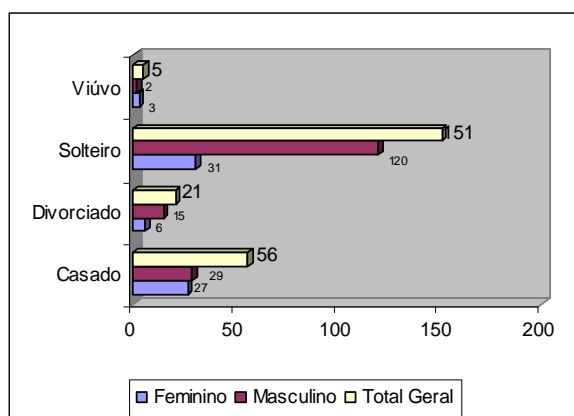


Figura 6 – Caixas de bigodes das variáveis: idade na entrada, idade no diagnostico e idade actual



Quanto ao estado civil, embora não se tenha conseguido saber esta informação acerca de 69 dos indivíduos em estudo (cerca de 23% do total), de acordo com o exposto na Figura 7

pode verificar-se que a maior parte dos indivíduos são solteiros (50%) dos quais cerca de 80% são homens.

Figura 7 - Estado civil dos indivíduos

Seguem-se os casados (19%), com peso semelhante para ambos os sexos, e os divorciados (7%) na sua maioria homens. Uma pequena percentagem pertence aos viúvos (1%), única categoria onde prevalecem as mulheres.

#### 4.2.2 – Sector de Actividade e Profissão

Como observável pelo gráfico na Figura 8 a maior parte dos indivíduos trabalham (ou trabalharam) no sector primário/secundário e apenas 9% no sector terciário.

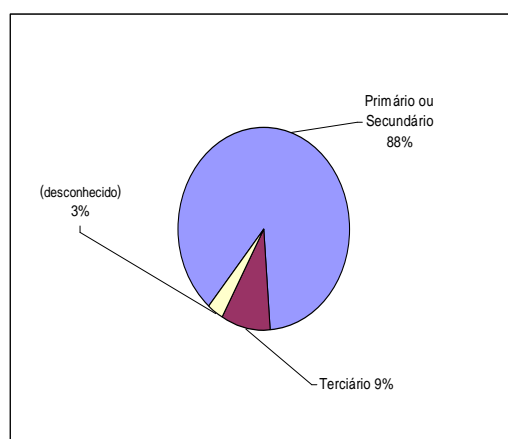


Figura 8 – Distribuição dos indivíduos por sector de actividade

De facto, não é de estranhar estes resultados uma vez que na região as actividades com maior peso têm sido ao longo do tempo a agricultura, pecuária, extracção mineira e, por outro lado, a construção civil.

No que respeita às profissões em concreto, 31 dos indivíduos em estudo já se encontram reformados. Dos restantes, e no que respeita aos dados que se conseguiram extrair podem referir-se com maior peso as profissões de Pedreiro, Servente, Armador de Ferro, Pintor ou Serralheiro da Construção Civil, Carpinteiro, Empregada Doméstica, Trabalhador Agrícola, Pastor, Operário Fabril, Padeiro, Empregado Administrativo/ Escriturário e Empregado de mesa.

### **4.2.3 – Distrito e Concelho de Residência**

De acordo com os dados dispostos na Tabela 4, o distrito onde residem a maior parte dos doentes seguidos é Beja. De facto, estes representam cerca de 89% do total de indivíduos. A maior parte provém dos concelhos de Beja (cerca de 42% face ao total de indivíduos, valor que confirma os dados publicados e já aqui referidos da DDI-URVE), Odemira, Aljustrel, Castro Verde e Serpa. Não obstante será o facto destes estarem entre os concelhos mais populosos e por isso com maior probabilidade de surgir a infecção e referenciação para o Hospital distrital.

Quanto a Odemira há que referir a existência do estabelecimento prisional de mulheres, das quais as que estão infectadas são seguidas em Beja não importando a sua residência originária (Faria, 2010). De resto, o Hospital está também a seguir (ou seguiu durante o período em estudo) utentes provenientes dos distritos de Lisboa, Faro, Setúbal, Évora, Coimbra e Portalegre.

Uma possível explicação dessa ocorrência poderá relacionar-se com a população universitária proveniente de diversos pontos do País na cidade de Beja e, por outro lado, talvez com maior peso, refiram-se todos os utentes toxicodependentes provenientes de outros distritos que estão integrados em clínicas e instituições de desintoxicação/ reabilitação na zona de Beja.

<b>Distrito</b>	<b>Concelho</b>	<b>Total</b>	<b>%</b>
<b>Beja</b>	Beja	127	<b>89.1%</b>
	Odemira	46	
	Aljustrel	19	
	Castro Verde	12	
	Serpa	12	
	Mértola	11	
	Moura	11	
	Ferreira do Alentejo	8	
	Almodôvar	7	
	Alvito	6	
	Vidigueira	5	
	Cuba	2	
	Ourique	2	
	Barrancos	1	
<b>Coimbra</b>	Coimbra	1	<b>0.3%</b>
<b>Évora</b>	Évora	2	<b>0.7%</b>
<b>Faro</b>	Loulé	2	<b>2.6%</b>
	Olhão	2	
	Albufeira	1	
	Monchique	1	
	Portimão	1	
	Silves	1	
<b>Lisboa</b>	Lisboa	6	<b>3.0%</b>
	Cascais	3	
<b>Portalegre</b>	Castelo de Vide	1	<b>0.3%</b>
<b>Setúbal</b>	Santiago do Cacém	2	<b>2.0%</b>
	Alcácer do Sal	1	
	Sesimbra	1	
	Setúbal	1	
	Sines	1	
<b>(desconhecido)</b>		6	<b>2.0%</b>
<b>TOTAL GERAL</b>		<b>302</b>	

Tabela 4 – Distrito e Concelho de residência

Uma análise mais detalhada que compare o peso da infecção nos concelhos do distrito de Beja face aos respectivos totais de população residente permite identificar quais os concelhos onde o risco de infecção parece ser maior, por revelarem maiores taxas de prevalência (considera-se aqui a proporção por 1000 habitantes). De facto, contrapondo os dois tipos de informação obtém-se o exposto na Tabela 5.

Concelhos do distrito de Beja	População Residente (Dados INE)	%	Concelhos do distrito de Beja	População infectada	%	Taxa de Prevalência do Concelho (por 1000 habitantes)
Beja	35762	22.2%	Beja	127	47.2%	3.6
Odemira	26106	16.2%	Odemira	46	17.1%	1.8
Serpa	16723	10.4%	Aljustrel	19	7.1%	1.1
Moura	16590	10.3%	Castro Verde	12	4.5%	0.7
Aljustrel	10567	6.6%	Serpa	12	4.5%	1.1
Ferreira do Alentejo	9010	5.6%	Mértola	11	4.1%	1.2
Mértola	8712	5.4%	Moura	11	4.1%	1.3
Almodôvar	8145	5.1%	Ferreira do Alentejo	8	3.0%	1.0
Castro Verde	7603	4.7%	Almodôvar	7	2.6%	0.9
Ourique	6199	3.8%	Alvito	6	2.2%	1.0
Vidigueira	6188	3.8%	Vidigueira	5	1.9%	0.8
Cuba	4994	3.1%	Cuba	2	0.7%	0.4
Alvito	2688	1.7%	Ourique	2	0.7%	0.7
Barrancos	1924	1.2%	Barrancos	1	0.4%	0.5
<b>Total</b>	<b>161211</b>	<b>100%</b>	<b>Total</b>	<b>269</b>	<b>100%</b>	<b>1.7</b>

Tabela 5 – População residente e população infectada nos concelhos do distrito de Beja

Por um lado, analisando cada um dos concelhos isoladamente, os que apresentam maior taxa de prevalência e serão, portanto, mais preocupantes incluem Beja, Odemira, Moura, Mértola, Serpa e Aljustrel. No entanto, se a análise for efectuada tendo em conta o distrito como um todo, e a contribuição de cada concelho, verifica-se que os concelhos onde a infecção parece ter um peso relativo superior ao esperado tendo em conta o peso de cada concelho no total de habitantes do distrito de Beja, são efectivamente Beja, Aljustrel, Alvito e Castro Verde. A maior concentração da infecção verifica-se no concelho de Beja.

Positivamente há que destacar em particular os concelhos de Ferreira do Alentejo e Ourique uma vez que o peso relativo da infecção face ao total de habitantes do distrito é muito inferior ao que seria de esperar, comparando o peso da população residente nestes concelhos face ao total do distrito com o peso da população infectada proveniente destes concelhos face aos totais da infecção no distrito.

No geral, a taxa de prevalência calculada para o distrito de Beja é de 1,7, donde, aproximadamente 2 em cada 1000 habitantes estarão infectados com o VIH/ SIDA.

Uma outra forma de interpretar estes resultados considera que um distrito ou concelho que apresente maior número de casos face ao esperado não indica necessariamente que a situação



nessa região seja particularmente preocupante. Nesta outra perspectiva de análise, esses concelhos podem simplesmente ter melhores condições para a detecção da doença e daí o número de casos verificados.

#### 4.2.4 – Estado Actual, Tipo de Vírus e Grupos/ Comportamentos de Risco

De acordo com a Figura 9, 47% dos doentes encontram-se efectivamente activos (estão vivos e a ser seguidos no hospital). Os restantes faleceram (16%) ou foram transferidos para outra instituição (28%), ou simplesmente abandonaram o acompanhamento hospitalar (9%).

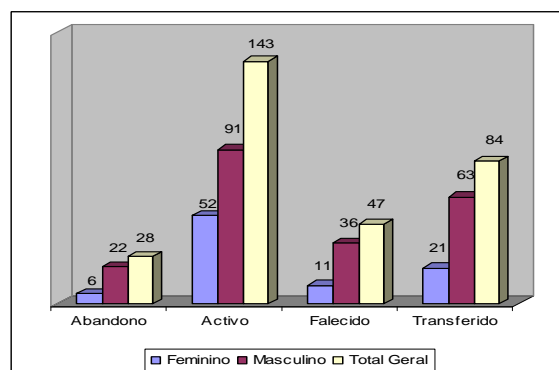


Figura 9 – Estado Actual dos indivíduos em estudo

Como evidenciado na Figura 10, e relativamente ao tipo de vírus, cerca de 97% dos indivíduos estão infectados com VIH 1 (valor semelhante à taxa nacional), e os restantes contraíram o tipo VIH 2 ou ambos os tipos.

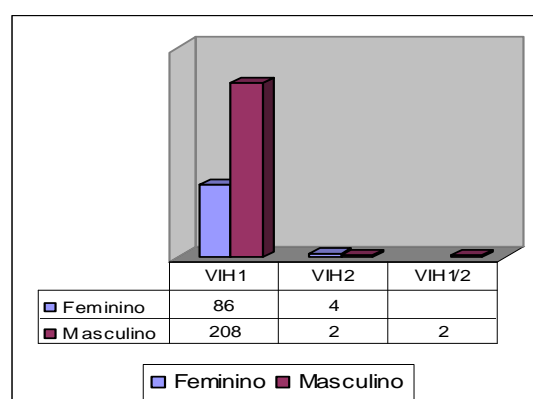


Figura 10 – Total de indivíduos por tipo de vírus

Quanto às categorias de transmissão ou grupos de risco, cuja informação se encontra na Tabela 6, nos indivíduos para os quais se conseguiu recolher informação (98% da amostra), no que respeita à orientação sexual, aqueles distribuem-se em Heterossexuais (cerca de 79%) e Homo/Bissexuais (cerca de 19%). Ao analisar a utilização de drogas injectáveis, 59% dos indivíduos são toxicodependentes.

Apenas um dos indivíduos contraiu a doença via transfusão de sangue.

<b>Grupo de Risco</b>	<b>Feminino</b>	<b>Masculino</b>	<b>Total Geral</b>	<b>%</b>
Heterossexual	49 (54%)	44 (21%)	<b>93</b>	<b>31%</b>
Heterossexual e Toxicodependente	24 (27%)	119 (56%)	<b>143</b>	<b>47%</b>
Heterossexual e Transfusão de Sangue	1 (1%)	0	<b>1</b>	<b>0%</b>
Homo/Bissexual	5 (6%)	18 (9%)	<b>23</b>	<b>8%</b>
Homo/Bissexual e Toxicodependente	9 (10%)	26 (12%)	<b>35</b>	<b>12%</b>
(desconhecido)	2 (2%)	5 (2%)	<b>7</b>	<b>2%</b>
<b>Total Geral</b>	<b>90</b>	<b>212</b>	<b>302</b>	<b>100%</b>

<b>Grupo de Risco na População Reclusa</b>	<b>Feminino</b>	<b>Masculino</b>	<b>Total Geral</b>	<b>%</b>
Heterossexual	2	2	4	9%
Heterossexual e Toxicodependente	9	19	28	64%
Homo/Bissexual	1		1	2%
Homo/Bissexual e Toxicodependente	5	3	8	18%
<b>(desconhecido)</b>	<b>2</b>	<b>1</b>	<b>3</b>	<b>7%</b>
<b>Total Geral</b>	<b>19</b>	<b>25</b>	<b>44</b>	<b>100%</b>

Tabela 6 – Grupos de risco *versus* sexo

Ao analisar as categorias conjuntas verifica-se que a maior parte dos doentes são heterossexuais e/ ou toxicodependentes. De facto, estas parecem ser as categorias de transmissão com maior peso na região, e deverão ser por isso o maior foco de preocupação das autoridades de saúde.

Como já mencionado, a população reclusa constitui um dos grupos de maior risco porque integra características muito propensas à transmissão do vírus. No total de indivíduos em estudo, aproximadamente 15% são (ou foram) reclusos, e destes, mais uma vez, prevalecem as categorias Heterossexual e/ou Toxicodependente.

Com o objectivo de verificar se as categorias (ou variáveis) grupo de risco e sexo são independentes, foi elaborado um teste não-paramétrico – Teste de Fisher (pois têm-se uma ou mais células com valor esperado igual ou inferior a 5, impossibilitando a interpretação do Teste de Independência do Qui-Quadrado), obtendo assim o valor exacto do p-value. As hipóteses em causa são:

$H_0$  : Grupos de risco e Sexo são variáveis independentes

$H_1$  : Grupos de risco e Sexo não são variáveis independentes.

Assumindo um nível de significância  $\alpha = 1\%$ , conclui-se que a hipótese nula de independência é rejeitada (visto ter-se obtido  $p\text{-value} < 0,001 < \alpha$ ). Desta forma, rejeita-se  $H_0$  e assume-se a existência de associação entre as variáveis, tal que o sexo do indivíduo não é independente do grupo de risco a que pertence.

#### 4.2.5 – Situação da infecção aquando da data de entrada no acompanhamento hospitalar

No que respeita à população linfática aquando da entrada, dos valores obtidos para a variável “contagem de células CD4”, pelo seu valor mais baixo até aquela data, têm-se os resultados apresentados na Tabela 7.

Medidas Amostrais	ContCD4
Média	390.61
IC a 95% para a Média - Limite Inferior	342.94
IC a 95% para a Média - Limite Superior	438.29
Mediana	350.00
Moda	6.00
Variância	86721.48
Desvio-Padrão	294.49
Coefficiente de Variação	0.75
Mínimo	4.00
Máximo	1368.00
Amplitude	1364.00
Percentil 10	29.00
Percentil 25	167.50
Percentil 75	565.50
Percentil 90	1059.00
Amplitude Interquartil	398.00
Assimetria	.87
Erro-Padrão (Assimetria)	.20
Curtose	.51
Erro-Padrão (Curtose)	.40

Em média, os indivíduos que começaram a ser seguidos no Hospital de Beja apresentavam na data de entrada cerca de 390 células CD4 por microlitro de sangue (com desvio-padrão de 294). O respectivo coeficiente de variação indica que os desvios em relação à média perfazem 75% desta, o que é bastante elevado. De facto, os dados revelam muita heterogeneidade, o que também é visível nos resultados para o cálculo da amplitude (1364 células de diferença entre o valor mais elevado e o mais baixo). Uma possível causa talvez seja que apenas se conseguiu recolher informação para cerca de 49% da amostra (149 indivíduos).

Tabela 7 – Contagem de células CD 4

Tendo em conta que o normal, numa pessoa saudável, são valores entre os 800 e 1300 células CD4 por microlitro de sangue, pode considerar-se que à data de entrada na consulta de acompanhamento alguns indivíduos já vinham com o sistema imunitário bastante debilitado.

Mais uma vez está-se perante uma variável que segue uma distribuição assimétrica positiva, leptocúrtica, e onde se detectam *outliers* moderados, como se pode confirmar pela observação da Figura 11.

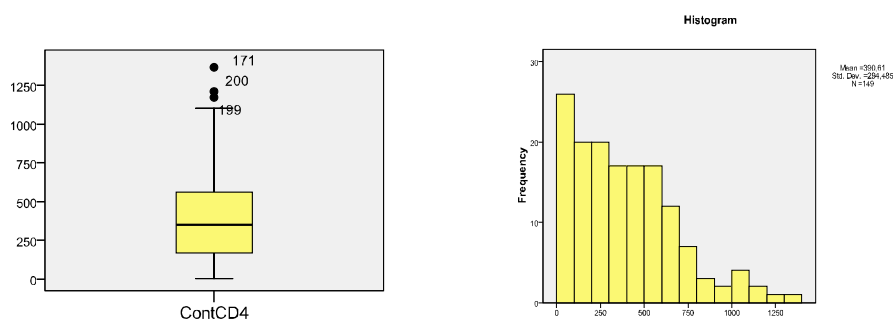


Figura 11 – Distribuição de probabilidade das células CD4

Considerando que os indivíduos foram classificados quanto à sua situação clínica e contagem de células CD4 (pelo seu valor mais baixo) na data de entrada, de acordo com as categorias definidas pelo CDC e já referidas no Capítulo 2 (ponto 2.1.5 - Fases da Doença) obteve-se a Tabela 8 cujos dados se analisam.

Valor mais baixo de células CD4 (à data de entrada)	Categoria Clínica (à data de entrada)			Total Geral
	A/B	C	(desconhecido)	
$\geq 500$	30	4		34
200-499	103	16		119
$< 200$	38	77		115
(desconhecido)	25	1	8	34
<b>Total Geral</b>	<b>196</b>	<b>98</b>	<b>8</b>	<b>302</b>

Tabela 8 – Categoria clínica dos indivíduos

Da observação da Tabela 8 verifica-se que apenas 11% dos doentes registou como valor mais baixo de células CD4 por microlitro de sangue um valor pertencente à categoria  $\geq 500$  cel/ul. Isto evidencia que, embora, em média, e como já referido, o doente tenha sido diagnosticado apenas 3 anos antes de iniciar o acompanhamento nesta instituição, esse período foi o suficiente para que a infecção se propagasse rapidamente e o organismo tenha ficado com o sistema de defesa em situação debilitada. De facto, muitos doentes obtiveram contagens de células CD4 tão baixas ( $< 200$  cel/ul), tais que, de acordo com o sistema de classificação do CDC se pode

considerar que já atingiram o estadió de SIDA (cerca de 38%). Para além disso, considerando também os doentes que, embora se situassem noutras categorias de contagem de células CD4, já apresentavam doença definidora de SIDA (categoria C na situação clínica) pode afirmar-se que afinal cerca de 45% dos indivíduos já estão no estadió mais avançado da infecção por VIH, e portanto com maior possibilidade de contrair doença oportunista, e, eventualmente, falecer a curto prazo. Estes doentes requerem, por isso, uma maior atenção por parte da equipa de acompanhamento.

Considerando as categorias clínicas A e B (Assintomático e Sintomático, respectivamente) com valores de células CD4 conhecidos e acima de 200 células por microlitro de sangue verifica-se uma proporção semelhante à verificada para o estadió SIDA (embora este tenha sido confirmado não apenas pela categoria clínica, mas também pela contagem de células CD4), com cerca de 44% face ao total de indivíduos em estudo.

Mais uma vez, verifica-se que existe interesse em avaliar a independência estatística entre 2 variáveis, mais concretamente a contagem de células CD4 e categoria clínica. Foi realizado o teste não-paramétrico de Fisher, donde, para um nível de significância  $\alpha = 1\%$ , se rejeita  $H_0$  e se assume a existência de associação entre as variáveis ( $p\text{-value} < 0,001 < \alpha$ ). Este resultado coincide assim com o esperado face à classificação do CDC dos doentes com VIH/SIDA nos vários estádios da doença.

A referir são também a quantidade de dados que não se conseguiram extrair no que respeita, especialmente, à contagem de células CD4. De facto, estes representam aproximadamente 11% do total de doentes analisados e o conhecimento do seu valor poderia “desempatar” e clarificar a sua distribuição pelas várias fases da doença. De qualquer forma, a partir do momento em que o indivíduo contrai a infecção, e não apenas quando as suas células CD4 atingem determinado valor, deve começar a ser devidamente acompanhado para controlar a progressão do vírus.

Se a análise for efectuada por sexo, constata-se que, pela análise da Tabela 9, nas mulheres, a proporção de doentes no estadió SIDA é de aproximadamente 46%, e nos homens é de 45%. Por

tais resultados pode afirmar-se que por sexo não existem diferenças relevantes quanto à fase avançada da doença com a qual os doentes foram recebidos no hospital.

Categoria Clínica (à data de entrada) – Sexo Feminino				Categoria Clínica (à data de entrada) – Sexo Masculino					
Valor mais baixo de células CD4 (à data de entrada)	A/B	C	(desc.)	Total Geral	Valor mais baixo de células CD4 (à data de entrada)	A/B	C	(desc.)	Total Geral
>=500	9	1		10	>=500	21	3		24
200-499	33	6		39	200-499	70	10		80
<200	11	22		33	<200	27	55		82
(desconh.)	5	1	2	8	(desconh.)	20		6	26
<b>Total Geral</b>	<b>58</b>	<b>30</b>	<b>2</b>	<b>90</b>	<b>Total Geral</b>	<b>138</b>	<b>68</b>	<b>6</b>	<b>212</b>

Tabela 9 – Categoria clínica dos indivíduos por sexo

Nas categorias A e B, existem ligeiras diferenças na análise por sexo. De facto, a proporção de mulheres nestas categorias (com contagens de células CD4 superiores a 200 *cel/ul*, tal como na análise anterior), é de 47% enquanto que nos homens se verifica um peso de 43%. Portanto, nas mulheres a infecção parece não ter progredido na mesma medida do que nos homens. De facto, estes valores, embora a sua diferença não seja muito significativa, podem indiciar que o vírus é mais agressivo nos homens do que nas mulheres.

Quanto aos dados que não se conseguiram recolher verifica-se que no sexo masculino essa dificuldade foi superior, tal que a categoria “desconhecido” revela um maior peso quando comparada à sua homóloga no sexo feminino.

#### 4.2.6 – Falecidos

Neste ponto analisam-se os dados referentes unicamente aos indivíduos cujo estado actual seja falecido.

Registaram-se 47 falecimentos no total de indivíduos activos em estudo. Na Figura 12. Estão representados as proporções por sexo no total de falecidos.

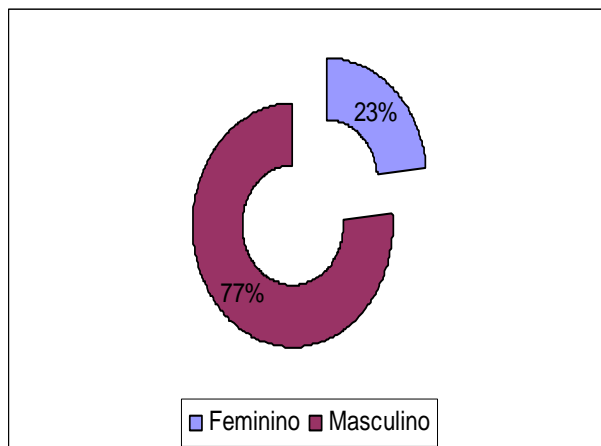


Figura 12 – Proporção de falecimentos por sexo

Embora, como referido anteriormente, na população que começou a ser seguida no hospital os homens representassem cerca de 70%, a sua proporção é superior quando se analisam os indivíduos falecidos. De facto, cerca de 77% destes eram homens.

Mais uma vez, e de acordo com o exposto no ponto anterior, o vírus parece atacar de uma forma um pouco mais agressiva o sexo masculino. Talvez pelo facto da sua resposta imunológica inicial à infecção ser mais forte e daí o vírus, ao infectar ainda mais linfócitos, progredir mais fácil e rapidamente.

Quando analisada a causa da morte (Figura 13), verifica-se que para 31 daqueles indivíduos (cerca de 66% do total de falecidos) o seu falecimento decorreu efectivamente na sequência da doença.

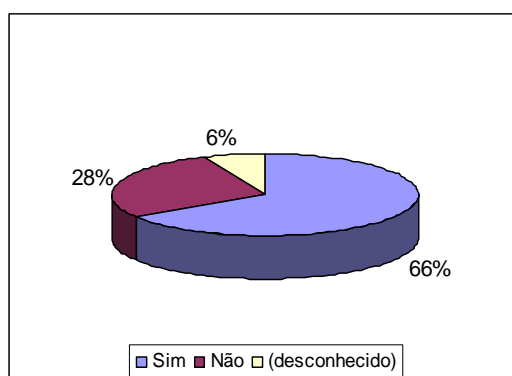


Figura 13 – Causa da morte dos indivíduos

De facto, a maioria apresentava já um conjunto de sintomas ou doença oportunista definidores de SIDA, nomeadamente síndrome de emaciação por VIH ou *wasting* (cerca de 23% dos indivíduos), Pneumonia por *Pneumocystis Carinii* ou Pneumonia por *Pneumocystis Jirovecii* (cerca de 13%), Linfoma não-Hodgkin, Linfoma Primário do Sistema Nervoso Central, Neoplasia do Pulmão e Carcinoma do Colo do Útero (cerca de 12%), Infecção Citomegalovirus Generalizada (cerca de 9%), Tuberculose Pulmonar ou Disseminada (cerca de 6%), Septicemia e Encefalopatia associada ao VIH.

Dos indivíduos que não faleceram directamente do impacto da infecção, registaram-se causas de morte tais como Cirrose Hepática ou Mista, *Overdose* e AVC Hemorrágico.

No que respeita à idade com que os indivíduos faleceram, observe-se a informação constante nas Tabela 10 e Figura 14.

Medidas Amostrais	TempSobrTF
Média	45.66
IC a 95% para a Média - Limite Inferior	40.43
IC a 95% para a Média - Limite Superior	50.88
Mediana	39.00
Moda	39.00
Variância	252.66
Desvio-Padrão	15.90
Coefficiente de Variação	0.35
Mínimo	25.00
Máximo	83.00
Amplitude	58.00
Percentil 10	28.80
Percentil 25	34.00
Percentil 75	57.00
Percentil 90	72.20
Amplitude Interquartil	23.00
Assimetria	.86
Erro-Padrão (Assimetria)	.38
Curtose	-.20
Erro-Padrão (Curtose)	.75

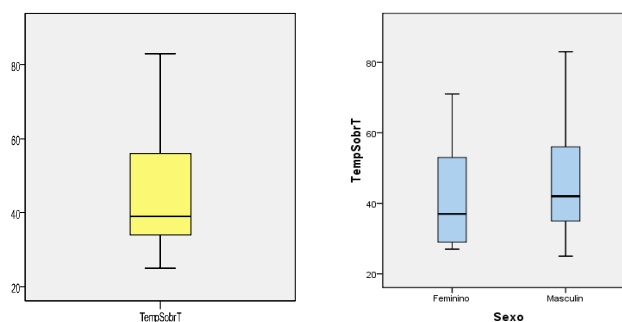


Figura 14 – Idade de falecimento

Em média, o indivíduo faleceu com 46 anos (42 anos para as mulheres e 47 anos para os homens), com um desvio-padrão de 16 anos. O coeficiente de variação indica que o desvio dos dados face à média representa 35% desta.

Tabela 10 – Medidas da estatística descritiva para a idade de falecimento



O indivíduo mais velho cujo falecimento se registou era do sexo masculino, tinha 83 anos, e estava a seguir a terapêutica antiretroviral. Sabe-se também que foi diagnosticado nesse mesmo ano. A sua causa de morte foi Cardiopatia Isquémica, que por si só não consta das doenças definidoras de SIDA do CDC. Ainda assim, e como informação adicional, refira-se que vários estudos recentes [13] apontam para o relacionamento da referida patologia com a infecção por VIH pois acredita-se que a dislipidemia associada à terapêutica antiretroviral da infecção VIH (nomeadamente ao nível dos inibidores da protease e inibidores não nucleósidos da transcriptase reversa) pode contribuir para o aumento da incidência de Cardiopatia Isquémica. Neste caso, a idade avançada do indivíduo falecido também terá contribuído, naturalmente, para o aparecimento ou agravamento de patologias do foro cardíaco.

Outro indício a retirar destes resultados é o de que as mulheres morrem menos do que o sexo oposto, mas mais cedo. Não alheio será o facto de que, em média, as mulheres em estudo foram diagnosticadas com uma idade inferior à dos homens.

No outro extremo da distribuição dos dados tem-se o indivíduo mais novo, que faleceu com 25 anos, sexo masculino e, coincidentemente, havia também sido diagnosticado nesse mesmo ano. Também este doente estava sob terapêutica antiretroviral.

Ainda relativamente aos indivíduos falecidos e quanto ao número de anos que decorreu entre a data do diagnóstico e a data do falecimento, analise-se os dados referentes à variável “TemSobrDF”, cujas medidas amostrais constam da Tabela 11.

Medidas Amostras	TempSobrDF		
	Feminino	Masculino	Total
Média	4.89	5.45	5.32
IC a 95% para a Média - Limite Inferior	1.80	3.53	3.75
IC a 95% para a Média - Limite Superior	7.97	7.37	6.88
Mediana	4.00	4.00	4.00
Moda	2.00	2.00	2.00
Variância	16.11	25.40	22.76
Desvio-Padrão	4.01	5.04	4.77
Coeficiente de Variação	0.82	0.93	0.90
Mínimo	1.00	0.00	0.00
Máximo	14.00	17.00	17.00
Amplitude	13.00	17.00	17.00
Percentil 10	1.00	0.00	0.00
Percentil 25	2.00	1.00	1.75
Percentil 75	6.50	8.80	8.00
Percentil 90	-	14.00	14.00
Amplitude Interquartil	5.00	8.0	6.00
Assimetria	1.67	.80	.91
Erro-Padrão (Assimetria)	.72	.43	.38
Curtose	3.03	-.37	-.12
Erro-Padrão (Curtose)	1.40	.86	.75

Tabela 11 – Medidas da estatística descritiva da variável “TemSobrDT”

Em média, os indivíduos cujo estado actual é falecido viveram cerca de 5 anos com a doença, desde o ano em que foram diagnosticados até à data do óbito.

Por sexo, verifica-se que o período que decorre entre a data do diagnóstico e a data do óbito para as mulheres é, em média, inferior em 0,56 anos (cerca de 7 meses) face aos homens.

Dados os valores obtidos, poderá inferenciar-se se a média de tempo de vida com a doença, desde a data do respectivo diagnóstico, será menor que 7 anos, por exemplo. As hipóteses a colocar são:

$$H_0 : \mu = 7$$

$$H_1 : \mu < 7$$

Analise-se o resultado da estatística t. Para um nível de significância ou probabilidade de cometer o erro tipo I (rejeitar  $H_0$  quando ela é verdadeira) de 5%, o valor da estatística de teste T = -2,176 e p-value = 0,036, ou seja, rejeita-se a hipótese nula. Desta forma pode afirmar-se com

um nível de confiança de 95% que estes indivíduos vivem em média menos de 7 anos após a data do diagnóstico.

De facto, a amostra revela que o sexo feminino apresenta uma média de 4,89 anos (com desvio-padrão de 4 anos) e o sexo masculino tem média de 5,45 anos (com desvio-padrão de 5 anos).

O coeficiente de variação calculado para os dois sexos indica que os desvios em relação à média face ao valor que esta apresenta são maiores no sexo masculino (o desvio dos dados face à média é cerca de 92% desta), comparativamente ao sexo feminino, onde os dados parecem ser um pouco mais homogéneos do que aqueles, embora com um coeficiente de variação também muito elevado, de 82%.

#### 4.2.7 – Terapêutica Antiretroviral (TAR): Início, Adesão versus Interrupção e Antiretrovirais Prescritos

A terapêutica antiretroviral veio revolucionar a forma de lidar com a infecção por VIH e, na maior parte dos casos, veio proporcionar um aumento de longevidade a esses doentes, não obstante os seus efeitos secundários. Sabe-se que não foi apenas aquando do acompanhamento no Hospital de Beja que os indivíduos tomaram contacto com a TAR. De facto, tenha-se em conta que alguns deles vieram de outras instituições onde já haviam iniciado (com interrupção ou não) este tipo de terapêutica. Ainda assim, a maior parte dos doentes que entraram para o acompanhamento em Beja não tinha iniciado TAR antes dessa data.

As Tabelas 12 e 13 apresentam o total de indivíduos por sexo e estado actual que iniciaram ou não TAR.

Terapêutica Antiretroviral (TAR)	Sexo	Estado Actual				Total Geral
		Abandono	Activo	Falecido	Transferido	
Não	Feminino	2	3	1	13	19
	Masculino	11	17	7	29	64
<b>Não Total</b>		<b>13</b>	<b>20</b>	<b>8</b>	<b>42</b>	<b>83</b>
Sim	Feminino	4	49	10	8	71
	Masculino	11	74	29	34	148
<b>Sim Total</b>		<b>15</b>	<b>123</b>	<b>39</b>	<b>42</b>	<b>219</b>
<b>Total Geral</b>		<b>28</b>	<b>143</b>	<b>47</b>	<b>84</b>	<b>302</b>

Tabela 12 – Indivíduos que iniciaram ou não TAR, por sexo

Estado Actual	N.º de Indivíduos em estudo	Proporção de Indivíduos Total	N.º de Indivíduos em TAR	Proporção de Indivíduos em TAR no total de TAR	Proporção de Indivíduos em TAR face ao Total de Indivíduos por Estado Actual
Abandono	28	9.3%	15	6.8%	53.6%
Activo	143	47.4%	123	56.2%	86.0%
Falecido	47	15.6%	39	17.8%	83.0%
Transferido	84	27.8%	42	19.2%	50.0%
<b>Total</b>	<b>302</b>	<b>100.0%</b>	<b>219</b>	<b>100.0%</b>	<b>72.5%</b>

Tabela 13 – Indivíduos que iniciaram ou não TAR, por tipo de estado

Dos 302 indivíduos em estudo cerca de 73% iniciou a TAR, dos quais actualmente 123 estão em estado activo. Estes representam cerca de 56% no total de indivíduos que iniciaram TAR.

Dos indivíduos que iniciaram TAR cerca de 68% são do sexo masculino.

Ao analisar a Tabela 12 por estado actual, constata-se que tanto para os indivíduos que acabaram por ser transferidos como para os que abandonaram o acompanhamento hospitalar, o facto de terem iniciado a TAR parece não ter sido muito relevante para essa situação actual. De facto, tanto para a categoria sim (iniciou TAR) como para a categoria não (não iniciou TAR), nestes dois tipos de estado, o número de doentes naquelas condições é igual ou muito semelhante. No entanto, ao comparar o total de doentes que iniciaram TAR, por tipo de estado, face ao total geral de doentes em TAR, os estados “Transferido” e “Abandono” são dos que revelam menor peso (cerca de 19% e 7%, respectivamente). Poder-se-ia dizer que esse facto foi influenciado por estes tipos de estado terem efectivamente menos doentes e portanto o seu peso no total (quer tenham iniciado TAR ou não) é sempre baixo. Mas essa dedução pode ser refutada. De facto, se se analisar com mais atenção, verifica-se que a proporção de doentes cujo estado actual é Activo que iniciaram TAR, é superior à esperada face ao peso do estado “Activo” no total de estados. De facto, esperar-se-ia um peso de aproximadamente 47%, quando afinal estes indivíduos (que iniciaram a TAR) representam cerca de 56%. A mesma conclusão se retira para o estado Falecido (de 16% para 18%) mas não é tão relevante. Esta tendência não se verifica nos restantes estados. Assim, parece existir indício que nos estados Abandono e Transferido (especialmente neste último), existe decepção quanto ao número de indivíduos que iniciaram TAR. Isso poderá não ter a ver apenas com a falta de adesão destes doentes, mas também opção da equipa médica de acordo com as características que aqueles apresentavam.

Para os indivíduos que faleceram, a maioria tinha iniciado a TAR (cerca de 83%), mas possivelmente já não foi a tempo de travar a reprodução do vírus ou limitar a resistência à própria medicação.

Quanto aos indivíduos activos é grande a proporção dos que, entre estes, iniciaram o tratamento (cerca de 86% do total de indivíduos activos), sendo este o estado actual que revela o maior número de doentes medicados pela TAR.

Por sexo, verifica-se que os homens que iniciaram TAR representam cerca de 49% do total de indivíduos em estudo (e as mulheres apenas 24%).

Tendo em conta que a amostra é essencialmente constituída por homens (cerca de 70%), então pode afirmar-se que neste sexo o início de TAR foi afinal relativamente menor. Se estão 212 homens em estudo e 90 mulheres, então a proporção de homens que iniciaram TAR face ao total de homens é baixa, de apenas 69%, enquanto que nas mulheres é de 79%. Isto indica que, comparativamente, e de um ponto de vista relativo, as mulheres foram o sexo mais proposto para iniciar TAR.

Com vista a melhor avaliar esta situação do ponto de vista estatístico, pode realizar-se um teste de hipóteses para a diferença de proporções entre os dois sexos. As hipóteses a ter em conta são:

$$H_0 : p_H = p_M$$

$$H_1 : p_H > p_M$$

onde  $p_H$  é a proporção de homens e  $p_M$  é a proporção de mulheres. Considerando um nível de significância de 1% e utilizando a distribuição normal (pela aplicação do Teorema do Limite Central dado o tamanho da amostra), o valor obtido para a estatística de teste  $Z = -0,00565$ . Este valor conduz à não rejeição da hipótese nula. Conclui-se que, com um nível de significância de 1%, não existem diferenças entre a proporção de homens e mulheres que iniciaram TAR.

Por sexo e estado actual, verifica-se que a maior parte das mulheres que não iniciou TAR foi transferida, e apenas uma faleceu. Das que iniciaram TAR, cerca de 69% estão em estado activo.

Nos homens, é de realçar também a diferença que existe para os que iniciaram o tratamento e se encontram em estado activo. De facto, dos 91 homens activos cerca de 81% iniciaram TAR. No entanto, esta mesma percentagem verifica-se nos homens que faleceram.

A tabela seguinte evidencia as idades médias dos indivíduos, por sexo e estado actual, aquando do início da Terapêutica Antiretroviral.

De acordo com os dados na Tabela 14, em média, os indivíduos que iniciaram a TAR tinham 36 anos (com desvio-padrão de 11 anos).

<b>Estado Actual</b>	<b>Feminino</b>	<b>Masculino</b>	<b>Total Geral</b>
Abandono	28	33	31
Activo	35	36	36
Falecido	39	44	42
Transferido	31	34	34
<b>Total Geral</b>	<b>35</b>	<b>37</b>	<b>36</b>

Tabela 14 – Idade média no início da TAR

Considerando que, também em média, e como referido anteriormente, os doentes entraram para o acompanhamento hospitalar com 36 anos, então o início de TAR está a acontecer no ano de entrada, para um indivíduo com as características médias ou usuais. De facto, para 101 dos 219 indivíduos que iniciaram TAR, esse ano coincidiu com o ano de entrada no acompanhamento hospitalar.

Analisando por sexo, as mulheres (que entram para a consulta com 34 anos, em média) iniciam a TAR com 35 anos, e os homens (que entram para a consulta com 37) iniciam a TAR no ano de entrada. Poder-se-ia à partida afirmar que embora a TAR esteja a ser prescrita mais às mulheres (inter-sexo) isso está a fazer-se no ano seguinte à sua entrada, mas tal não corresponde à verdade. Ainda que a média (em anos) da diferença entre o início da TAR e a data de entrada no Hospital de Beja seja igual a 1 ano, se se analisarem mais pormenorizadamente os dados verifica-se que existem muitas mulheres que já tinham iniciado a TAR há mais tempo (por exemplo, uma mulher que iniciou a TAR 10 anos antes de começar a ser acompanhada no Hospital de Beja), influenciando a média, que neste caso é pouco representativa. Na análise dos dados, detecta-se que, pela moda, a maior percentagem de mulheres também iniciou TAR no ano de entrada (34 das 71 mulheres em TAR), como os homens.

Por estado actual, os falecidos foram o grupo que iniciou mais tarde (em média, com 42 anos) a TAR, e os que abandonaram foram o grupo que iniciou mais cedo a terapêutica. Possivelmente essa poderá ter sido uma das razões do abandono. Quanto mais jovem o indivíduo que inicia a medicação antiretroviral, mais fácil e rapidamente a infecção é controlada, o que induz o indivíduo a achar que já está melhor, a sentir-se mais saudável e a desprezar a toma contínua dos medicamentos. Por outro lado, os efeitos adversos da medicação podem mais facilmente desencantar um indivíduo jovem do que um indivíduo com mais maturidade, levando muitas vezes à situação de abandono e descrédito no acompanhamento hospitalar.

Tem-se informação que o indivíduo mais velho a iniciar TAR tinha 83 anos, sexo masculino e actualmente o seu estado é falecido. A mulher mais velha a iniciar TAR tinha na altura 71 anos e actualmente o seu estado é activo.

O indivíduo mais novo a iniciar TAR tinha 20 anos (existiam 3 indivíduos nessas condições) nessa data e actualmente o seu estado pode ser abandono, activo ou transferido. A mulher mais nova tinha também 20 anos e actualmente foi transferida para acompanhamento noutra instituição.

	Feminino	Masculino
Média	2.03	2.41
Máximo	12.00	15.00
Mínimo	0.00	0.00

Quanto ao número de anos após o diagnóstico com que os indivíduos iniciaram TAR (Tabela 15), verifica-se que em média essa etapa ocorreu após 2 anos do diagnóstico da infecção. O maior número de anos sem medicação registado é de um indivíduo do sexo masculino, que apenas iniciou TAR 15 anos depois de ter sido diagnosticado.

Tabela 15 – Número de anos no início da TAR após diagnóstico da infecção

Categoria para células CD4 (valor mais baixo)	Feminino	Masculino	Total Geral
>=500	6	6	12
200-499	33	57	90
<200	32	77	109
(desconhecido)		8	8
<b>Total Geral</b>	<b>71</b>	<b>148</b>	<b>219</b>

Categoria Clínica	Feminino	Masculino	Total Geral
A/B	42	86	128
C	29	61	90
(desconhecido)		1	1
<b>Total Geral</b>	<b>71</b>	<b>148</b>	<b>219</b>

No que respeita à categoria clínica e categoria de contagem de células CD4 pelo valor mais baixo à data de entrada, verifica-se que a maior parte dos indivíduos que iniciaram TAR já tinham apresentado uma contagem inferior a 200 células CD4 por microlitro de sangue (evidenciando estadio SIDA) embora não tivessem ainda apresentado doença definidora de SIDA (inscritos nas situações clínicas A/B) (*vide* Tabela 16).

Tabela 16 – Categoria clínica e categoria de contagem das células CD4

Quanto à adesão, uma coisa é iniciar a TAR, outra bem diferente é continuar o tratamento sem interrupções devidamente justificadas. Os efeitos adversos da medicação são muitas vezes dissuasores da efectiva adesão do doente à terapêutica prescrita pelo médico.

De acordo com a informação na Figura 15 quando analisados os indivíduos em estudo, dos 219 que iniciaram TAR, cerca de 16% (34 doentes) interromperam o tratamento, por diversas razões. Em média, estes indivíduos tinham iniciado a TAR com 33 anos (para ambos os sexos). Dos 34 doentes que interromperam a TAR cerca de 56% foram homens.

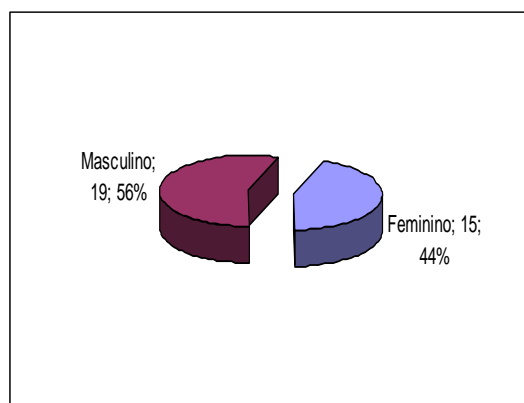


Figura 15 – Indivíduos que interromperam a TAR, por sexo



Se se considerar como indicador uma taxa de adesão efectiva calculada com os doentes que cumprem a prescrição face ao total de doentes que iniciaram o tratamento, então 84% dos indivíduos que iniciaram TAR são devidamente responsáveis e encontram-se motivados tal que respeitam o esquema terapêutico.

Quanto ao seu estado actual, a maioria dos indivíduos que interromperam a TAR encontram-se activos.

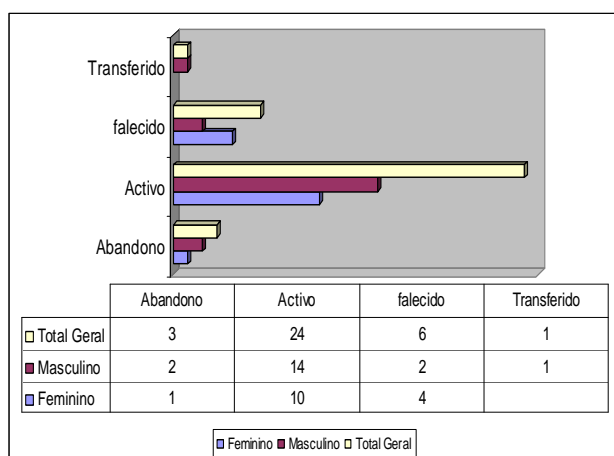


Figura 16 – Estado actual dos indivíduos que interromperam a TAR

Da análise à Figura 16 verifica-se que nos indivíduos que faleceram (47 doentes no total), 6 tinham interrompido a terapêutica. Tal pode ter influenciado a progressão da infecção para níveis mortais. De facto, em apenas um desses casos a morte não se deveu directamente à infecção VIH/SIDA.

Da análise realizada aos indivíduos em estudo verifica-se que para cerca de 43% dos doentes que iniciaram TAR não se tem indicação de quais foram os antiretrovirais (AR) e respectivos grupos de medicamentos prescritos. Ainda assim, dos restantes indivíduos, sabe-se que a maioria foi prescrito com antiretrovirais dos grupos NITR, NNITR e NITR, IP. A Tabela 17 e o gráfico na Figura 17 apresentam, por grupo das substâncias antiretrovirais, a prescrição aos indivíduos em estudo.

Grupo dos Antiretrovirais	Total de Indivíduos (início TAR)	%
NITR, NNITR	61	27.90%
NITR, IP	58	26.50%
NITR, IP, II	3	1.40%
NITR	2	0.90%
(desconhecido)	95	43.40%
<b>Total Geral</b>	<b>219</b>	<b>100.00%</b>

Figura 17 – Grupos de antiretrovirais

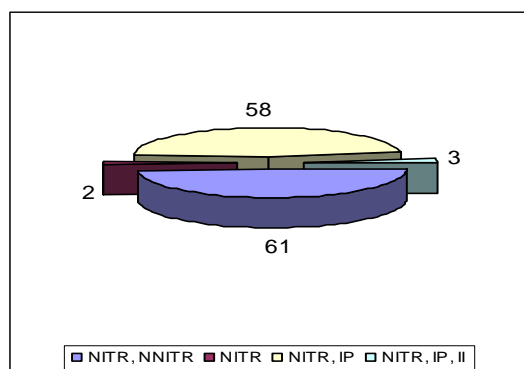


Tabela 17 – Grupos de antiretrovirais

Os antiretrovirais mais prescritos foram os inibidores da enzima transcriptase reversa (nucleósidos – NITR – e não nucleósidos - NNITR) e inibidores da protease (IP). O efeito desejado parece ser evitar a propagação do vírus inibindo a enzima através da qual o vírus se reproduz e/ou inibir a enzima que cliva as proteínas do vírus após aquela transcrição. Tenta-se portanto agir logo no início do processo como o vírus se propaga no organismo, ou seja, a estratégia principal para estes doentes (cujas características e condicionantes hão-de ter influenciado a escolha do antiretroviral prescrito) é limitar a capacidade do vírus transformar o seu código genético de ARN em ADN nas células e a sua reprodução.

Como observável na Figura 18 no caso dos indivíduos que, por qualquer razão, interromperam o tratamento, mais uma vez não se conseguiu extrair informação para todos eles, mas dos dados disponíveis sabe-se que a maioria havia sido prescrita com antiretrovirais dos tipos NITR, IP e NITR, NNITR.

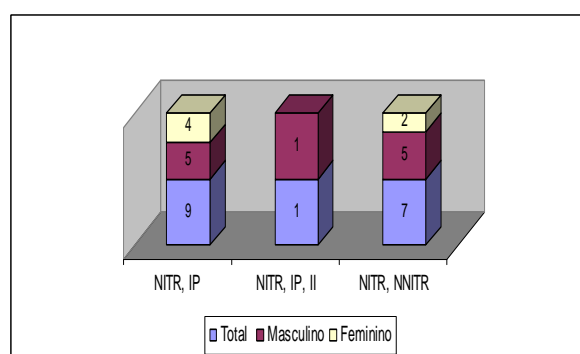


Figura 18 – Grupos de antiretrovirais para indivíduos que interromperam o tratamento

Neste caso, o maior peso recai sobre os grupos NITR, IP, o que pode ser um indício, ainda que ténue, que o fármaco inibidor da protease pode ter alguma influência na desistência ou abandono da terapêutica.

#### 4.2.8 – Carga Viral após TAR

Após o início da terapêutica antiretroviral é comum avaliar a eficácia desta através da contabilização da carga viral. De facto, o objectivo primordial do tratamento é a supressão da carga viral para níveis indetectáveis.

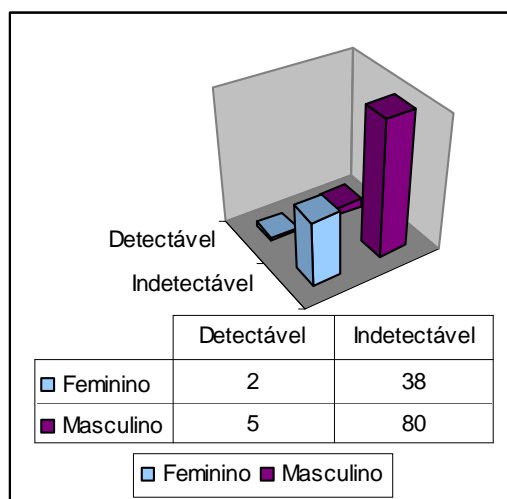


Figura 19 – Carga viral após TAR

Nos indivíduos em estudo verificou-se que (*vide* Figura 19), para aqueles em que se conseguiu extrair essa informação (cerca de 57% dos que iniciaram TAR), a carga viral apresentou-se em praticamente todos indetectável (em 94% dos 125 indivíduos).

#### 4.2.9 – Patologias Verificadas e Comportamentos de Adicção

De entre toda a informação de que se dispunha relativamente às patologias apresentadas pelos indivíduos em estudo, e considerando a importância de cada uma no objectivo principal do próprio, optou-se por analisar as seguintes: Tuberculose, Hepatite B, Hepatite C, Depressão Nervosa, Alcoolismo e Tabagismo. Os resultados obtidos constam na Tabela 18 que a seguir se apresenta.

<b>Patologias</b>	<b>Feminino</b>	<b>Masculino</b>	<b>Total Geral</b>	<b>%</b>
Hepatite C	13	73	<b>86</b>	39.30%
Depressão Nervosa	14	10	<b>24</b>	11.00%
Hepatite C e Depressão Nervosa	9	13	<b>22</b>	10.00%
Hepatite C e Alcoolismo	3	12	<b>15</b>	6.80%
Hepatite C e Tuberculose	4	8	<b>12</b>	5.50%
Tuberculose	6	3	<b>9</b>	4.10%
Tabagismo	2	6	<b>8</b>	3.70%
Hepatite B e Hepatite C	2	5	<b>7</b>	3.20%
Alcoolismo		6	<b>6</b>	2.70%
Hepatite C e Tabagismo		5	<b>5</b>	2.30%
Hepatite C, Alcoolismo e Tabagismo		5	<b>5</b>	2.30%
Depressão Nervosa e Tuberculose	1	3	<b>4</b>	1.80%
Alcoolismo e Tuberculose		3	<b>3</b>	1.40%
Hepatite B, Hepatite C e Tuberculose		2	<b>2</b>	0.90%
Hepatite C, Depressão Nervosa e Tabagismo	1	1	<b>2</b>	0.90%
Hepatite C, Depressão Nervosa e Tuberculose	1	1	<b>2</b>	0.90%
Alcoolismo e Tabagismo		1	<b>1</b>	0.50%
Depressão Nervosa e Alcoolismo		1	<b>1</b>	0.50%
Hepatite B, Hepatite C, Alcoolismo e Tuberculose		1	<b>1</b>	0.50%
Hepatite B, Hepatite C, Depressão Nervosa e Tabagismo		1	<b>1</b>	0.50%
Hepatite C, Alcoolismo e Tuberculose		1	<b>1</b>	0.50%
Hepatite C, Alcoolismo, Tabagismo e Tuberculose		1	<b>1</b>	0.50%
Hepatite C, Tabagismo e Tuberculose		1	<b>1</b>	0.50%
<b>Total Geral</b>	<b>56</b>	<b>163</b>	<b>219</b>	<b>100.00%</b>

Tabela 18 – Patologias e comportamentos de adicção

Da análise à Tabela 18 é notória a prevalência de Hepatite C entre os indivíduos em estudo. De facto, dos que se conseguiu apurar essa informação, 163 indivíduos (cerca de 74%) sofrem de infecção no fígado causada pelo vírus da hepatite C (com maior proporção no sexo masculino). Como já foi referido, este aspecto pode ser explicado pelo facto de ambos os vírus da Hepatite e Imunodeficiência Humana partilharem as mesmas vias de transmissão. Se analisarmos o grupo de risco que mais contribui para este resultado verifica-se que 94% dos que apresentam Hepatite C são toxicodependentes. Na população reclusa verifica-se que cerca de 75% apresentam também Hepatite C, o que não é de estranhar pois a maior parte dos reclusos é Toxicodependente (hetero e homo/bissexual).

Por outro lado, o que se sabe também é que a Hepatite C normalmente afecta a evolução da infecção por VIH, no sentido em que pode aumentar o risco de progressão para SIDA e morte

associada, pode diminuir a recuperação dos linfócitos CD4 em resposta à terapêutica antiretroviral e aumenta o risco de toxicidade hepática e metabólica [13]. Esta situação é preocupante e coloca estes indivíduos num estado mais fragilizado, necessitando por isso de maior atenção e cuidado. Sabe-se ainda que dos 163 indivíduos infectados por Hepatite C, 115 iniciaram terapêutica antiretroviral.

Quanto à hepatite B não tem um impacto tão forte. De facto, apenas 4% dos indivíduos em estudo apresentam essa patologia, e mais uma vez com maior prevalência no sexo masculino.

Os dados da OMS/ONUSIDA revelam que Portugal é o terceiro país da União Europeia com mais casos de co-infecção VIH/ SIDA e Tuberculose. Nos indivíduos em estudo, esta doença infecciosa representa cerca de 13%, e neste caso a sua distribuição por sexo é mais equilibrada dos que nas restantes patologias em análise. Curiosamente, apenas 2 dos indivíduos “Reclusos” apresentam a doença e esses são toxicodependentes heterossexuais.

Na análise da patologia do foro psicológico, Depressão Nervosa (diagnosticada e referida pela equipa que segue os doentes), verifica-se que 56 dos indivíduos em estudo, e para os quais se conseguiu extrair esta informação, sofrem de transtorno depressivo. Neste caso, o sexo feminino ganha expressão face ao sexo oposto comparativamente às restantes patologias. Se se avaliar por sexo e grupo de risco, a maior parte dos indivíduos com depressão nervosa são no sexo masculino os toxicodependentes heterossexuais e no sexo feminino a maior parte são heterossexuais. Quanto ao seu estado actual, a maior parte encontra-se activo.

À partida existirá alguma relação entre o facto do indivíduo se encontrar deprimido moralmente e, no caso de ter iniciado a terapêutica, ter também interrompido a medicação até à data. Por um lado porque suporta com mais dificuldade os efeitos adversos dos antiretrovirais, por outro porque essa medicação pode agravar o seu estado moral. Cerca de 82% dos indivíduos que sofrem de depressão nervosa iniciaram a TAR. E, surpreendentemente apenas 13% desses interromperam o tratamento (com peso semelhante entre os sexos).

Da tabela em análise pode verificar-se também que cerca de 40% dos indivíduos que apresentam depressão nervosa sofrem de Hepatite C.

Quanto aos comportamentos de adicção, que nos seus efeitos também podem ser considerados como patologias, verifica-se que nos indivíduos para os quais se conseguiu extrair essa informação, cerca de 13% sofrem de alcoolismo, sendo a maior parte (cerca de 91%) homens. Muitos destes apresentam também Hepatite C.

Quanto ao tabaco e respectivo abuso, cerca de 9% dos indivíduos referem tabagismo, entre os quais a maior parte (cerca de 88%) são do sexo masculino. Aproximadamente 25% destes indivíduos sofre também de alcoolismo.

#### 4.2.10 – Frequência das consultas de acompanhamento – Enfermagem e Psicologia

Na Tabela 19 com os dados extraídos sobre os indivíduos em estudo, verifica-se que ao longo do período em análise, cerca de 42% dos indivíduos frequentaram a consulta de Enfermagem e 24% frequentaram a consulta de Psicologia. Do total de indivíduos, cerca de 16% frequentaram ambas as especialidades de consulta de acompanhamento.

	Frequência das consultas de acompanhamento		
	Enfermagem	Psicologia	Ambas as consultas
Utentes atendidos	126 (42%)	72 (24%)	47 (16%)
Utentes faltosos (todos os motivos)	14 (11%)	47 (65%)	2 (4%)

Tabela 19 – Frequência de consultas de acompanhamento

Os motivos para o indivíduo ter faltado (uma ou mais vezes) a uma consulta programada podem ser vários, alguns justificáveis e com marcação de consulta posterior, outros não. Sabe-se que existem utentes que descuram a ida a estas consultas e quando comparadas as duas especialidades, essa situação é muito mais frequente na consulta de Psicologia. De facto, 65% dos indivíduos supostos de frequentarem a consulta não cumpriram a assiduidade prevista. Essa taxa é muito mais baixa na consulta de enfermagem, que ainda assim revela que 11% dos indivíduos que frequentam (ou frequentaram) a consulta faltaram uma ou mais vezes.

A aplicação de um teste de independência qui-quadrado entre as variáveis consulta de enfermagem e consulta de psicologia resulta num p-value inferior ao nível de significância estabelecido (1%) tal que se pode afirmar que as variáveis não são independentes.

Dos 126 indivíduos atendidos na consulta de enfermagem, cerca de 68% são do sexo masculino e solteiros. Apenas 10% desses indivíduos são reclusos.

A maior parte dos doentes atendidos nesta consulta são toxicodependentes heterossexuais ou somente heterossexuais no que respeita ao seu grupo de risco. Cerca de 74% encontram-se em estado activo (e cerca de 10% já faleceram). Quanto à TAR, 105 dos indivíduos iniciaram o tratamento, entre os quais 22 (cerca de 21%) já interromperam o tratamento (de forma definitiva ou não).

Quanto às patologias identificadas, a maior parte sofrem de Depressão Nervosa (principalmente o sexo feminino) e/ou Hepatite C (principalmente o sexo masculino).

O p-value obtido no teste de independência para a hipótese nula de inexistência de associação entre a variável sexo e a variável consulta de enfermagem permite confirmar que o sexo do indivíduo não influencia o facto de frequentar ou não a consulta de enfermagem (p-value = 0,523 no teste Qui-Quadrado para um nível de significância de 5%).

Considera-se também pertinente verificar se o estado actual do doente (abandono, activo, transferido ou falecido) e frequência da consulta de enfermagem são variáveis independentes. Neste contexto é elaborado o respectivo teste de independência, donde, para um nível de significância  $\alpha = 1\%$ , se rejeita  $H_0$  e se assume a existência de associação entre as variáveis ( $p\text{-value} < \alpha$ ). Conclui-se que existe evidência estatística de associação entre o facto de frequentar a consulta e o estado actual do indivíduo no acompanhamento efectuado no hospital. Por outro lado, a maior parte dos indivíduos activos frequenta esta consulta, tal como os dados revelam (*vide* Tabela 20).

Estado Actual	Frequência da Consulta de Enfermagem			Total
	Sim	Não	(desconh.)	
Abandono	5	23	0	28
Activo	93	50	0	143
Falecido	12	34	1	47
Transferido	16	66	2	84
<b>Total</b>	<b>126</b>	<b>173</b>	<b>3</b>	<b>302</b>

Tabela 20 – Frequência de consultas de acompanhamento por estado actual do indivíduo

Na consulta de Psicologia, do total de 72 utentes atendidos, cerca de 58% são homens, 21% são reclusos, a maior parte são Heterossexuais (cerca de 51%), 82% encontram-se em estado activo (e cerca de 6% já faleceram). De resto, 59 dos indivíduos iniciaram TAR, tendo 14 desses vindo a interromper o tratamento (cerca de 23%) e mais uma vez, no que respeita às patologias verificadas, a maior parte apresenta Depressão Nervosa e/ou Hepatite C (com maior impacto nos homens).

O p-value obtido no teste de independência para a hipótese nula de inexistência de associação entre a variável sexo e a variável consulta de psicologia permite confirmar que, neste caso, o sexo do indivíduo influencia o facto de frequentar ou não esta consulta (p-value = 0,012 no teste Qui-Quadrado para um nível de significância de 5%).

### 4.3 – Análise Descritiva por Anos

Como já foi referido, o presente estudo incide sobre os anos 2000 a 2009. Na análise anterior os dados foram abordados de forma essencialmente descritiva e como um todo, isto é, não se consideraram possíveis diferenças nos resultados face ao período concreto a que dizem respeito. Nesta secção descrevem-se os dados por período anual.

#### 4.3.1 – Ano de Nascimento

Como observável no gráfico da Figura 20 a maior parte dos indivíduos nasceu nas décadas de 60 e 70. De facto, 22 dos indivíduos em estudo nasceram em 1974. O indivíduo mais velho nasceu em 1919 e o mais novo em 1988.

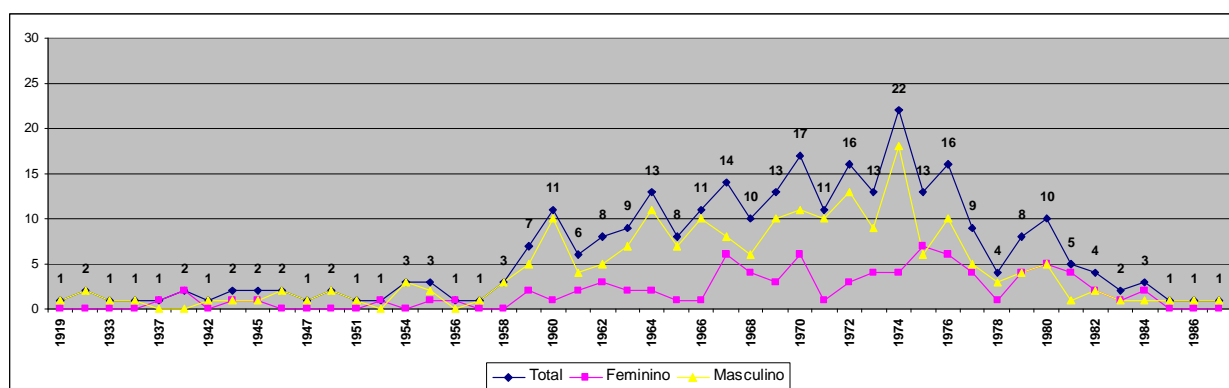


Figura 20 – Ano de nascimento dos indivíduos



Por sexo, em apenas 7 dos anos o número de mulheres foi superior ao dos homens, tendo-se registado em 1981 a maior diferença desse tipo. De resto, foi mesmo em 1974 que o número de homens foi aproximadamente 5 vezes superior ao número de mulheres nascidas.

#### 4.3.2 – Ano de Diagnóstico

De acordo com a Figura 21, o ano em que foram diagnosticados o maior número dos indivíduos em estudo com a infecção por VIH/SIDA foi em 1999.

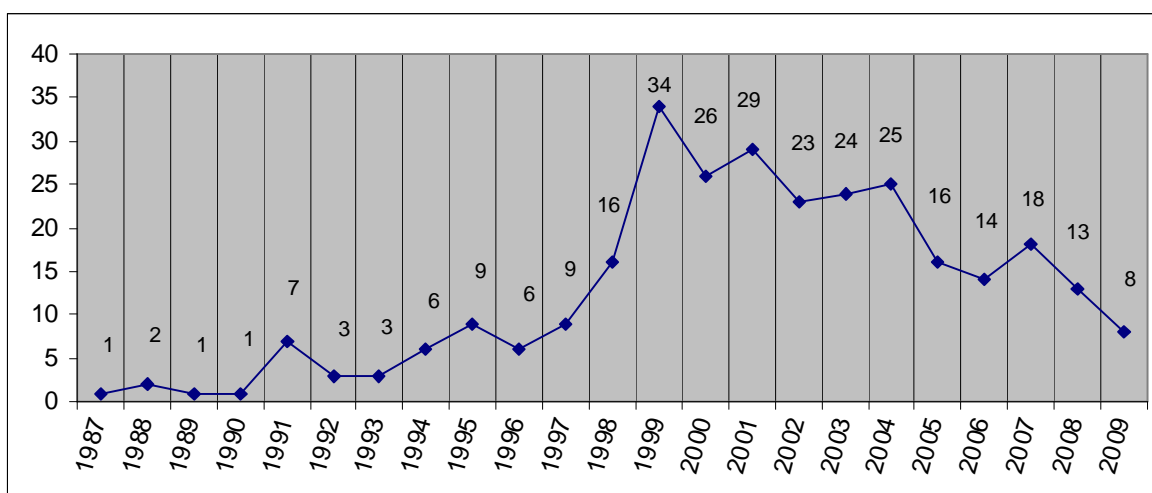


Figura 21 – Ano de diagnóstico dos indivíduos

Repare-se também a tendência decrescente que os dados assumem a partir dessa data, ainda que entre os anos posteriores existam pequenas flutuações. No outro extremo verifica-se a existência de indivíduos que quando começaram a ser seguidos no Hospital de Beja já conviviam com a doença há vários anos (o mais antigo há 13 anos).

Na análise por sexo (Figura 22) verifica-se que, dos indivíduos em estudo, apenas nos anos 1987 e 1992 o número de mulheres diagnosticadas foi superior ao do sexo oposto.

Nos restantes anos o número de homens diagnosticados foi sempre superior e a maior diferença entre os sexos ocorreu no ano 2004 no qual foram diagnosticados 22 dos homens em estudo e apenas 3 das mulheres.

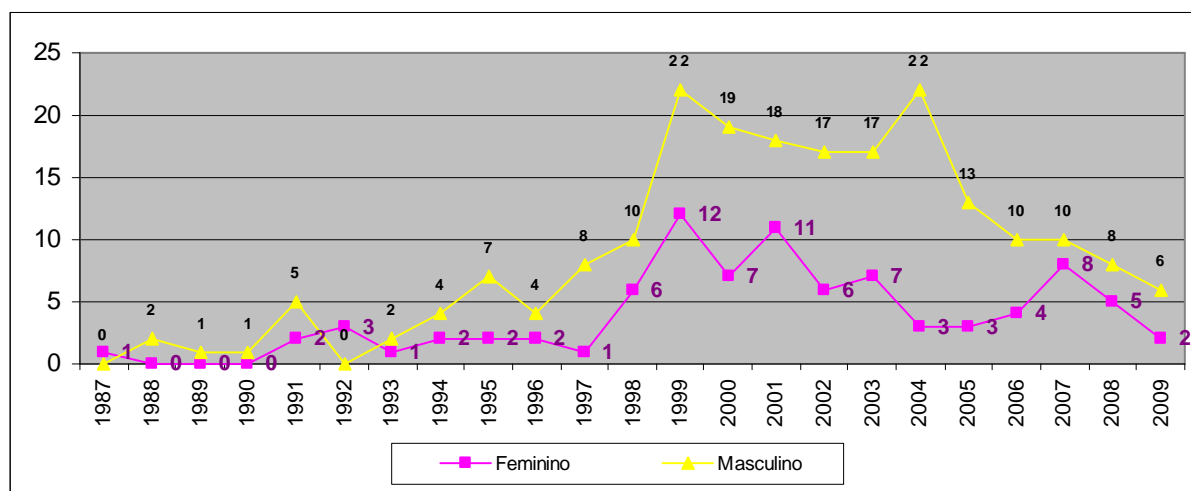


Figura 22 – Sexo dos indivíduos

#### 4.3.3 – Ano de entrada no acompanhamento hospitalar – Hospital de Beja

Como se pode observar na Figura 23, dos 10 anos em estudo, o ano 2001 registou o maior número de novas entradas no acompanhamento efectuado no Hospital de Beja. Por outro lado, em 2005 registou-se o menor número de novos indivíduos a serem seguidos no hospital.

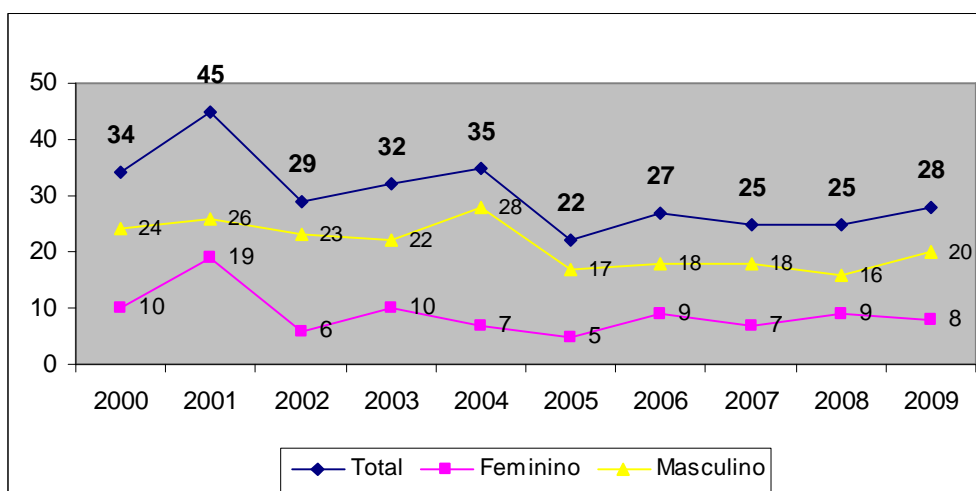


Figura 23 – Evolução do número de indivíduos por ano de entrada

De facto, a tendência desde o ano 2001 tem sido maioritariamente decrescente, embora com flutuações. Talvez aqui exista indicação de que a incidência (surgimento de novos casos) está efectivamente a diminuir ou a permanecer relativamente estável na região.

Em todos os anos o número de homens admitidos foi sempre superior ao das mulheres, embora nos anos 2001 e 2008 os seus valores estivessem próximos. Em 2004 foi quando se registou a maior diferença entre o número de indivíduos entrados no acompanhamento do Hospital de Beja, por sexo.

#### 4.3.4 – Ano de Falecimento

Por ano de falecimento (*vide* Figura 24) verifica-se que o maior número de mortes entre os indivíduos em estudo ocorreu em 2008 (não esquecendo que existem 9 dos falecimentos para os quais não se sabe qual o ano exacto dessa ocorrência, o que poderia alterar a análise). Em 2009, 2007 e 2002 foram os anos em que se registaram menor número de falecimentos.

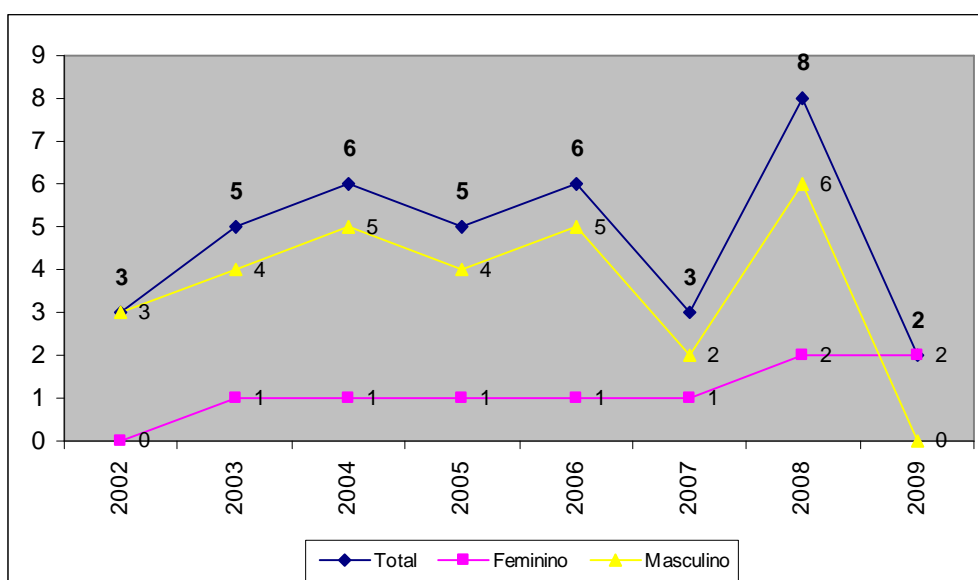


Figura 24 – Ano de falecimento dos indivíduos

Por sexo, verifica-se que no ano 2002 não faleceu nenhuma mulher, e no ano 2009 não faleceu nenhum homem. Nos restantes anos, o número de homens falecidos é superior ao das mulheres, sendo a diferença menor no ano 2007.

#### 4.3.5 – Ano de início da TAR

Da análise da Figura 25 verifica-se que o número de indivíduos a iniciar a TAR tem sofrido grandes alterações e oscilações ao longo do tempo. De facto, a partir de 1999 ocorreu uma subida acentuada dos indivíduos em TAR (o que não será alheio ao facto de esse ter sido o ano em que

se registaram o maior número de diagnósticos, o que mais uma vez leva a concluir que a maior parte dos indivíduos em estudo está a iniciar a TAR no ano em que são diagnosticados).

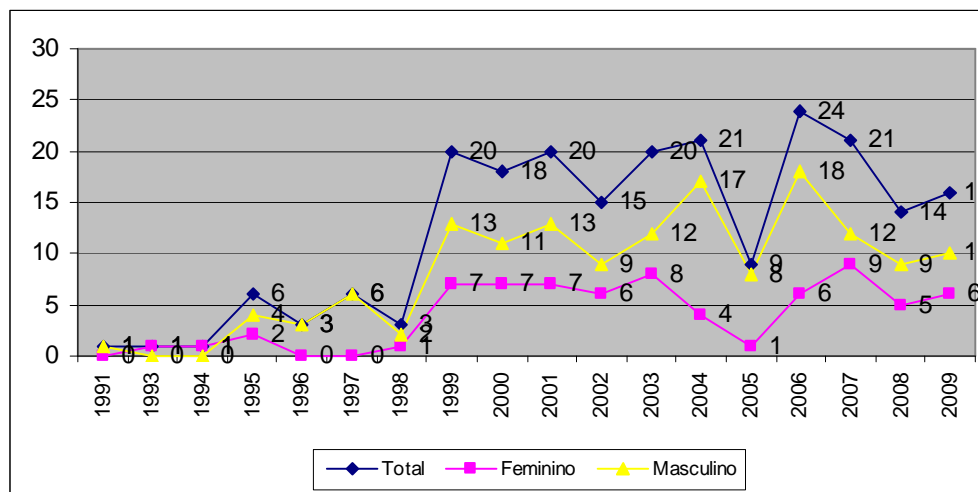


Figura 25 – Ano de início da TAR

Apesar de, a partir dessa data esse número se situar nos 18 indivíduos, em média, por ano, há que registar uma descida também acentuada no ano 2005, onde apenas 9 dos indivíduos iniciaram o tratamento. Ainda assim, no ano seguinte, em 2006, verificou-se o maior número de doentes a iniciar os antiretrovirais.

Por sexo, apenas em 1991, 1993 e 1994 o número de mulheres a iniciar TAR foi superior ao dos homens, mas sem grande expressão pois trata-se apenas de uma mulher em cada um desses anos. De resto, em todos os anos o número de homens é superior ao das mulheres, sendo essa diferença menor em 1998 e maior em 2004 e 2006.

#### Síntese – Capítulo 4

Neste capítulo foram abordados os dados em estudo de forma essencialmente descritiva. Verificou-se que ao nível das suas características gerais os indivíduos infectados com o VIH/SIDA que foram acompanhados no Hospital de Beja nos anos 2000 a 2009 apresentam alguns resultados semelhantes àqueles que caracterizam a infecção a nível regional, nomeadamente no que respeita à proporção de infectados por sexo, patologias apresentadas, comportamentos de risco, tipo de vírus e idades mais afectadas. Este aspecto pode revelar-se importante caso exista interesse em inferir os resultados obtidos para a população infectada no país e não apenas da região do Baixo Alentejo.

## 5. Análise de Sobrevida

Um dos objectivos do presente estudo é averiguar quais as variáveis estatisticamente significativas na explicação do tempo de sobrevivência (em dias) que os indivíduos apresentam desde a data de diagnóstico da infecção até à ocorrência do óbito. Embora na análise anterior se tenham já detectado alguns indícios sobre que factores podem estar a influenciar o tempo até à ocorrência de morte relacionada com a infecção VIH/ SIDA, importa agora aplicar métodos estatísticos que permitam inferir e modelar esse tempo.

Através de análise univariada testar-se-á a significância de cada uma das variáveis face à variável resposta. Depois dessa avaliação construir-se-á um modelo multivariado que explique o máximo possível do tempo de sobrevivência em estudo.

Para as variáveis que se revelem não significativas *a priori*, será igualmente ponderada a sua importância no modelo, para o que se recorrerá ao conhecimento empírico e científico, bem como a possíveis transformações e/ou interacções que venham a justificar a sua inclusão efectiva no modelo explicativo.

### 5.1 – Dados Censurados

É considerado como dado censurado o tempo associado a um indivíduo que não experimenta o evento de interesse (morte com associação directa ao VIH/SIDA) ou quando o seu seguimento é interrompido por alguma razão e por esse facto não se têm os dados completos.

Quanto ao tipo de censura geral os dados dizem respeito a indivíduos que entraram no estudo de forma simultânea ou progressiva e foram censurados (ao mesmo tempo no final do estudo) caso não tenha ocorrido o seu falecimento por SIDA. Está-se portanto perante o Tipo I de censura geral tal que o número de eventos no final do estudo é uma variável aleatória.

A Tabela 21 contém o total de indivíduos em estudo consoante o seu status: censurado ou não censurado.

	Feminino	Masculino	Total	%
Censurado	81	190	271	89.7%
Não Censurado	9	22	31	10.3%
Total	90	212	302	100.0%

Tabela 21 – Total de Indivíduos por status

Como observável na tabela acima, para os dados disponíveis à data, apenas 31 do total de indivíduos tinha falecido por causa directa de VIH/ SIDA aquando do final do período em análise. Espera-se que esse facto não retire potência e validade às conclusões encontradas.

Para efeitos de modelação e respectivo encaixe aos dados em estudo, consideram-se aptos para análise os indivíduos para os quais se tenha informação completa face às variáveis em estudo. Assim, eliminando da base de dados todas as linhas ou indivíduos com um ou mais dados em falta fica-se com um total de 244 indivíduos sobre os quais se aplica o modelo de sobrevivência. Daqueles, 27 correspondem a óbitos decorridos na sequência directa da infecção.

Uma análise preliminar à variável Tempo de Sobrevivência (“TempSobrTD”) permite verificar que o tempo mediano de sobrevivência é de 6 anos, aproximadamente) e o tempo médio de sobrevivência dos indivíduos em estudo é de aproximadamente 7 anos (desvio-padrão,  $s = 3$  meses). O tempo mínimo observado no período em análise é de 50 dias (de um indivíduo já diagnosticado com 83 anos) e o valor máximo é de 23 anos.

## 5.2 – Análise Univariada

### 5.2.1 – Análise univariada a variáveis qualitativas – Estimativas de Kaplan-Meier

Para as variáveis do tipo qualitativo começam-se por considerar as estimativas de Kaplan-Meier (KM), curvas de sobrevivência e respectivos testes de comparação ( $H_0$ : As curvas de sobrevivência são estatisticamente idênticas) no sentido de avaliar a sua significância no tempo de sobrevivência dos indivíduos após o diagnóstico.

Considere-se, nesta fase, um nível de significância ( $\alpha$ ) de 25%. Este valor é estabelecido para a entrada das variáveis no modelo, e tenta respeitar o aconselhado por autores como Hosmer e Lemeshow.

### 5.2.1.1- Sexo

De acordo com as estimativas obtidas (Tabela 22) as mulheres vivem, em média, cerca de 19 anos após o diagnóstico da infecção. Existe uma diferença de 3 anos entre os sexos, mas não parece ser muito significativa.

Means and Medians for Survival Time								
Sexo	Mean <sup>a</sup>				Median			
	Estimate	Std. Error	95% Confidence Interval		Estimate	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound			Lower Bound	Upper Bound
Feminino	6951,463	432,880	6103,018	7799,909	.	.	.	.
Masculino	6012,698	323,488	5378,663	6646,734	.	.	.	.
Overall	6653,438	306,853	6052,006	7254,871	.	.	.	.

a. Estimation is limited to the largest survival time if it is censored.

Tabela 22 – Estimativas de KM para a variável Sexo

Como se repara não se obtiveram quaisquer resultados para a mediana, muito possivelmente devido ao reduzido número de observações não censuradas que permitissem o seu cálculo.

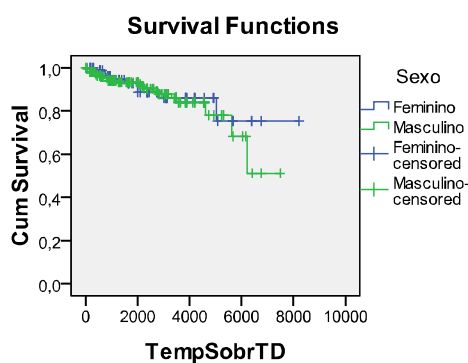


Figura 26 – Curvas de Sobrevivência de KM – Sexo

A representação gráfica das curvas de sobrevivência para ambos os sexos (Figura 26) evidencia (ou aparenta evidenciar) que o risco de morte entre homens e mulheres é muito semelhante, principalmente até aos 6000 dias (ou 16 anos) após o diagnóstico, a partir dos quais o risco de morte para os homens é superior.

	Chi-Square	df	Sig.
Log Rank (Mantel-Cox)	,201	1	,654
Breslow (Generalized Wilcoxon)	,064	1	,801
Tarone-Ware	,061	1	,805

Test of equality of survival distributions for the different levels of Sexo.

Como as curvas se cruzam analisa-se o Teste de Breslow (vide Tabela 23), cujo  $p\text{-value} > \alpha$  não permite rejeitar a  $H_0$  de que as curvas de sobrevivência são estatisticamente idênticas. Por conseguinte, considera-se que não existem diferenças significativas entre os sexos e como tal esta variável não parece ser importante na explicação do tempo de sobrevivência.

Tabela 23 – Testes de comparação de curvas de sobrevivência – Sexo

### 5.2.1.2- Estado Civil

De acordo com os dados expostos na Tabela 24, e consoante o seu estado civil, os indivíduos viúvos são os que, em média, sobrevivem menos tempo após serem diagnosticados com VIH/SIDA (em média, cerca de 9 anos). Seguem-se os indivíduos divorciados, com aproximadamente 13 anos, menos 3 anos do que os solteiros. Os casados são os que, também em média, sobrevivem mais tempo: 19 anos (com desvio-padrão de aproximadamente 2 anos).

EstCivil	Mean <sup>a</sup>				Median			
	Estimate	Std. Error	95% Confidence Interval		Estimate	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound			Lower Bound	Upper Bound
Casado	6873,564	688,947	5523,228	8223,901	.	.	.	.
Solteiro	5960,773	336,294	5301,637	6619,909	.	.	.	.
Divorciado	4632,211	746,886	3168,315	6096,107	4604,000	2215,753	261,124	8946,876
Viúvo	3126,000	243,952	2647,854	3604,146	2781,000	.	.	.
Overall	6390,172	348,061	5707,972	7072,371	.	.	.	.

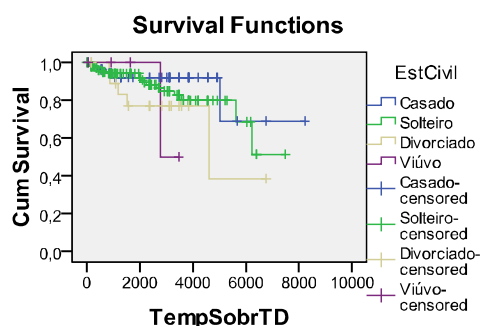
a. Estimation is limited to the largest survival time if it is censored.

Tabela 24 – Estimativas de KM para a variável – Estado Civil

A mediana foi calculada apenas para os viúvos e divorciados. Metade dos indivíduos os referidos estados civis sobrevivem 2781 (cerca de 8 anos) e 4604 dias (cerca de 13 anos), respectivamente, após o diagnóstico.



A representação gráfica das curvas de sobrevivência (Figura 27) revela que, para além do que já foi referido, e embora não se trate de uma diferença significativa, os casados apresentam um risco de morte quase sempre inferior aos solteiros, excepto num pequeno intervalo de tempo após o diagnóstico: dos 5000 aos 6000 dias (14 a 16 anos), no qual apresentam um risco superior.



	Chi-Square	df	Sig.
Log Rank (Mantel-Cox)	3,407	3	,333
Breslow (Generalized Wilcoxon)	1,785	3	,618
Tarone-Ware	2,454	3	,484

Test of equality of survival distributions for the different levels of EstCivil.

Figura 27 – Curvas de Sobrevivência de KM – Estado Civil

Tabela 25 - Testes de comparação de curvas de sobrevivência – Estado Civil

Como as curvas se cruzam analisa-se novamente o Teste de Breslow (Tabela 25), cujo p-value não permite rejeitar  $H_0$ , donde a variável estado civil não é estatisticamente importante na explicação do tempo de sobrevivência, tal que as respectivas curvas são idênticas.

### 5.2.1.3- Recluso

As estimativas obtidas (*vide* Tabela 26) indicam que, em média, o indivíduo recluso sobrevive cerca de 16 anos e o indivíduo não recluso cerca de 18 anos. A diferença não é de facto muito relevante. Existem por isso fortes indícios na não explicabilidade desta variável.

	Mean <sup>a</sup>				Median			
	Estimate	Std. Error	95% Confidence Interval		Estimate	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound			Lower Bound	Upper Bound
Recluso								
Sim	5814,335	337,843	5152,162	6476,508	.	.	.	.
Não	6521,972	341,981	5851,690	7192,254	.	.	.	.
Overall	6653,438	306,853	6052,006	7254,871	.	.	.	.

a. Estimation is limited to the largest survival time if it is censored.

Tabela 26 – Estimativas de KM para a variável – Recluso

As curvas de sobrevivência das duas categorias cruzam-se (Figura 28), pelo que se avalia o Teste de Breslow (Tabela 27). O p-value não permite rejeitar a  $H_0$ , logo pode afirmar-se que as curvas são estatisticamente idênticas (não há diferença entre as categorias).

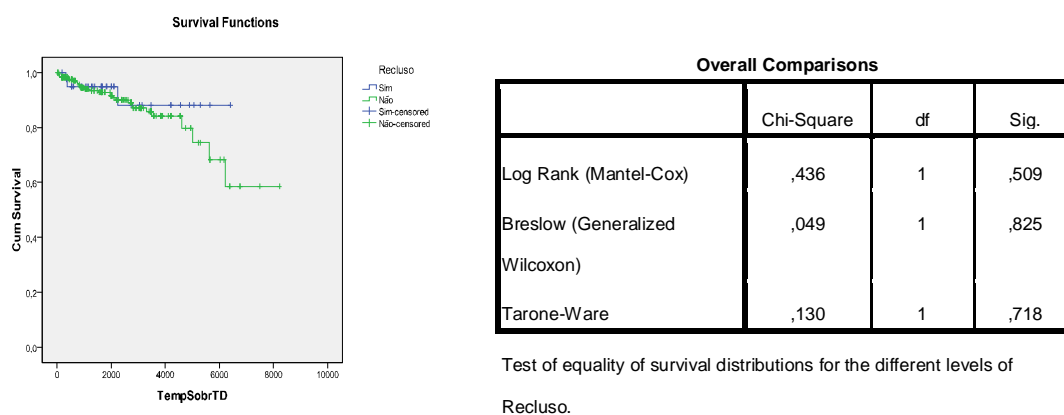


Figura 28 – Curvas de Sobrevivência de KM – Recluso

Tabela 27 - Testes de comparação de curvas de sobrevivência – Recluso

### 5.2.1.4– Grupo/ Comportamento de Risco: Heterossexual

A maior parte dos doentes em estudo são heterossexuais, e de facto este pode ser considerado um efectivo grupo de risco, embora a maioria da população se inclua nesta categoria no que respeita à sua orientação sexual. Em média, um indivíduo heterossexual sobrevive em média cerca de 18 anos após ser diagnosticado com o VIH/ SIDA (*vide* Tabela 28).

	Mean <sup>a</sup>				Median			
	Estimate	Std. Error	95% Confidence Interval		Estimate	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound			Lower Bound	Upper Bound
Heteross								
Sim	6606,372	328,673	5962,173	7250,572	.	.	.	.
Não	5637,795	418,912	4816,728	6458,862	.	.	.	.
Overall	6653,438	306,853	6052,006	7254,871	.	.	.	.

a. Estimation is limited to the largest survival time if it is censored.

Tabela 28 – Estimativas de KM para a variável – Heterossexual

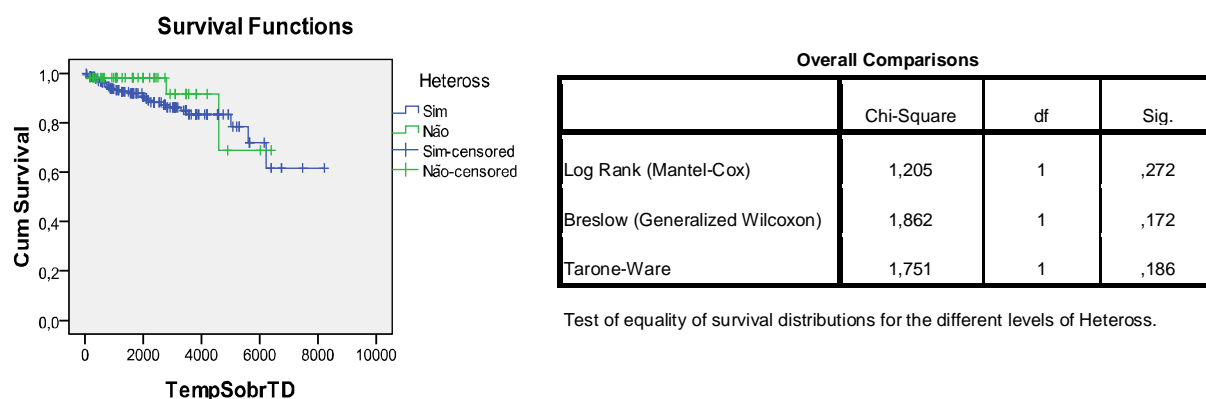


Figura 29 – Curvas de Sobrevivência de KM – Heterossexual

Tabela 29 - Testes de comparação de curvas de sobrevivência – Heterossexual

Da análise às curvas de sobrevivência para as categorias ser ou não heterossexual (Figura 29), verifica-se que até aproximadamente aos 4500 dias (cerca de 12 anos) após o diagnóstico, os indivíduos heterossexuais têm um maior risco de morte face aos restantes indivíduos. A partir desse período, os heterossexuais apresentam um risco de morte inferior e parecem sobreviver mais tempo.

A análise ao teste de comparação entre as curvas (Teste de Breslow – *vide* Tabela 29) leva à rejeição da  $H_0$ , para  $\alpha = 25\%$ , logo pode afirmar-se que existem diferenças quanto ao facto de ser ou não heterossexual no tempo de sobrevivência após o diagnóstico. A variável “Heterossexual” considera-se por isso significativa para entrar no modelo de sobrevivência a apurar.

### 5.2.1.5 – Grupo/ Comportamento de Risco: Toxicodependente

Os valores obtidos nas estimativas de K-M (*vide* Tabela 30) indicam que, em média, e surpreendentemente, os indivíduos toxicodependentes apresentam um tempo de sobrevivência superior aos restantes indivíduos. Os primeiros sobrevivem, em média, cerca de 18 anos e os não toxicodependentes cerca de 16 anos (com mediana igual a 15 anos). Mas os valores dos respectivos IC intersectam-se (um inclui o outro, neste caso), e portanto aquela diferença poderá não ser significativa.

Means and Medians for Survival Time

Toxicodep	Mean <sup>a</sup>				Median			
	Estimate	Std. Error	95% Confidence Interval		Estimate	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound			Lower Bound	Upper Bound
Sim	6479,072	237,698	6013,184	6944,960	.	.	.	.
Não	5865,571	596,597	4696,240	7034,902	5618,000	740,299	4167,014	7068,986
Overall	6653,438	306,853	6052,006	7254,871	.	.	.	.

a. Estimation is limited to the largest survival time if it is censored.

Tabela 30 – Estimativas de KM para a variável – Toxicodependente

A análise gráfica às curvas de sobrevivência (Figura 30) e respectivo teste de hipóteses (Teste de Breslow – Tabela 31) indicam que aquelas são estatisticamente idênticas (o p-value obtido não permite rejeitar essa  $H_0$ ), logo, como se previa, não existem diferenças significativas entre as categorias e por conseguinte não se considera explicativa nem útil a variável “Toxicodependente” no modelo de sobrevivência que se pretende construir.

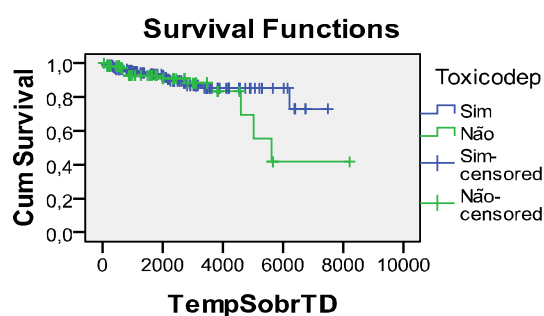


Figura 30 – Curvas de Sobrevivência de KM –Toxicodependente

Overall Comparisons

	Chi-Square	df	Sig.
Log Rank (Mantel-Cox)	,766	1	,381
Breslow (Generalized Wilcoxon)	,054	1	,816
Tarone-Ware	,139	1	,709

Test of equality of survival distributions for the different levels of Toxicodep.

Tabela 31 - Testes de comparação de curvas de sobrevivência- Toxicod.

### 5.2.1.6– Grupo/Comportamento de Risco: Transfusão de Sangue

Como apenas um dos indivíduos em estudo pertence a este grupo (e terá sido infectado por essa via), constituindo este um dado censurado (não tendo experienciado o evento de interesse), não há dados suficientes para calcular as estimativas de K-M, e obviamente esta variável não se considera significativa no tempo de sobrevivência em análise.

### 5.2.1.7 - Categoria Clínica

Os resultados do cálculo para as estimativas de KM (Tabela 32) resultou em aproximadamente 20 anos (em média) de tempo de sobrevivência após o diagnóstico para os indivíduos que, à data de entrada no acompanhamento do Hospital de Beja, apresentavam sinais clínicos que se enquadram nas categorias A ou B (de acordo com a classificação do CDC). Por outro lado, os indivíduos que se incluíam na categoria C (só por si definindo o estadio de SIDA), sobrevivem menos tempo, cerca de 14 anos (com desvio-padrão de 1 ano e 2 meses).

CatClinic	Mean <sup>a</sup>				Median			
	Estimate	Std. Error	95% Confidence Interval		Estimate	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound			Lower Bound	Upper Bound
A/B	7254,039	364,334	6539,945	7968,133	.	.	.	.
C	5241,971	429,901	4399,365	6084,576	.	.	.	.
Overall	6653,358	306,856	6051,920	7254,795	.	.	.	.

a. Estimation is limited to the largest survival time if it is censored.

Tabela 31 – Estimativas de KM para a variável – Categoria Clínica

Da observação do gráfico da Figura 31 que representa as curvas de sobrevivência confirma-se que os indivíduos na categoria clínica A/B (assintomáticos e sintomáticos) têm um risco de morte inferior, isto é, sobrevivem mais tempo, quando comparados aos indivíduos da categoria C (estadio SIDA). Para estes últimos, o período de maior risco vai até aos 3500 dias (cerca de 10 anos) após o diagnóstico do vírus.

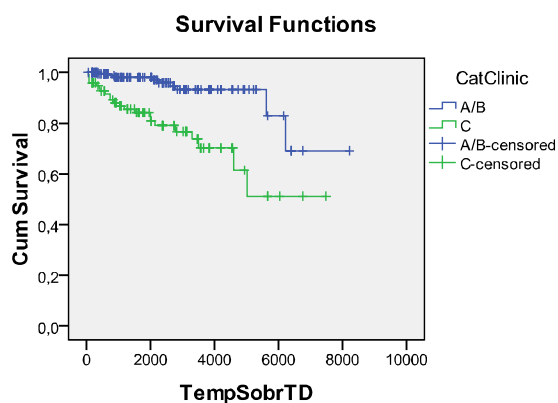


Figura 31 – Curvas de Sobrevivência de KM – Cat. Clínica

	Chi-Square	df	Sig.
Log Rank (Mantel-Cox)	19,680	1	,000
Breslow (Generalized Wilcoxon)	20,763	1	,000
Tarone-Ware	21,352	1	,000

Test of equality of survival distributions for the different levels of CatClinic.

Tabela 32 - Testes de comparação de curvas de sobrevivência- cat. Clínica

Como as curvas de sobrevivência não se cruzam realiza-se o Teste de Log Rank (*vide* Tabela 33), cujo p-value = 0,000 permite rejeitar a  $H_0$  no teste de comparação das curvas. Assim, admite-se que existem diferenças estatisticamente significativas entre a categoria clínica em que o indivíduo se encontra quando começa a ser seguido no que respeita ao tempo que sobrevive após o diagnóstico da doença.

### 5.2.1.8 - Categoria de Contagem de Células CD4

Da observação à Tabela 33 por categoria de contagem de células CD4 (pelo seu valor mais baixo), os indivíduos que já haviam registado menos do que 200 células CD4, apresentam um tempo de sobrevivência médio, após o diagnóstico, de 14 anos.

Nas restantes categorias, sabe-se que o número de dados não censurados é bastante reduzido, o que pode ter influenciado os resultados. De facto, os indivíduos que à partida se encontravam em melhor situação de defesa do seu sistema imunitário ( $\geq 500$  células CD4) apresentam um tempo médio de sobrevivência de 11 anos.

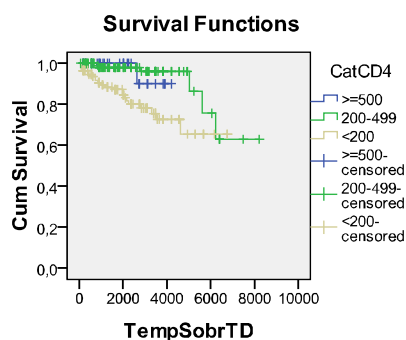
A categoria que, em média, parece sobreviver mais tempo após o diagnóstico é a dos indivíduos que, na data de entrada no acompanhamento no hospital, obtiveram entre 200 a 499 células CD4. Sobrevivem, em média, 20 anos (com o maior desvio-padrão apresentado nas três categorias, de aproximadamente 1 ano).

CatCD4	Mean <sup>a</sup>				Median			
	Estimate	Std. Error	95% Confidence Interval		Estimate	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound			Lower Bound	Upper Bound
>=500	4044,400	149,512	3751,356	4337,444	.	.	.	.
200-499	7122,602	420,481	6298,459	7946,745	.	.	.	.
<200	5227,146	295,076	4648,797	5805,494	.	.	.	.
Overall	6639,105	312,447	6026,708	7251,502	.	.	.	.

a. Estimation is limited to the largest survival time if it is censored.

Tabela 33 – Estimativas de KM para a variável – Categoria Células CD4

As curvas de sobrevivência cruzam-se (Figura 32) pelo que se analisa o Teste de Breslow. O p-value obtido neste teste (*vide* Tabela 32), por ser inferior ao nível de significância considerado, permite que se rejeite a  $H_0$ , donde se assume que existem diferenças significativas entre as categorias da variável “ContCD4”.



Overall Comparisons			
	Chi-Square	df	Sig.
Log Rank (Mantel-Cox)	13,217	2	,001
Breslow (Generalized Wilcoxon)	16,579	2	,000
Tarone-Ware	16,630	2	,000

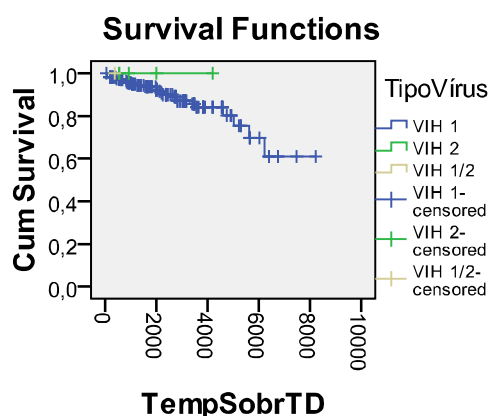
Test of equality of survival distributions for the different levels of CatCD4.

Figura 32 – Curvas de Sobrevivência de KM – Cat. Células CD4 e Tabela 34 – Testes de comparação de curvas de sobrevivência- Cat. Cél. CD4

### 5.2.1.9 - Tipo de Vírus

Todos os indivíduos que vieram a falecer por causa directamente relacionada com a infecção VIH/ SIDA estavam infectados com o tipo VIH 1.

A variável Tipo de Vírus não se considera explicativa no tempo de sobrevivência em estudo (confirme-se com o p-value obtido nos testes comparação – Tabela 35).



Overall Comparisons			
	Chi-Square	df	Sig.
Log Rank (Mantel-Cox)	,369	2	,832
Breslow (Generalized Wilcoxon)	,330	2	,848
Tarone-Ware	,356	2	,837

Test of equality of survival distributions for the different levels of TipoVirus.

Figura 33 – Curvas de Sobrevivência de KM – Tipo Vírus

Tabela 35 – Testes de comparação de curvas de sobrevivência- Tipo Vírus

### 5.2.1.10 - Início de TAR (Terapêutica Antiretroviral)

A Tabela 36 evidencia que, em média, os indivíduos que iniciaram o tratamento antiretroviral apresentam um tempo de sobrevivência superior. Estes sobrevivem, em média, cerca de 18 a 19 anos enquanto os restantes indivíduos que experienciaram o evento de interesse apresentam um tempo de sobrevivência após diagnóstico de aproximadamente 14 anos.

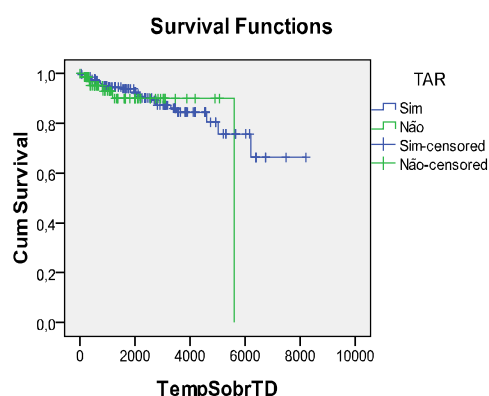
TAR	Mean <sup>a</sup>				Median			
	Estimate	Std. Error	95% Confidence Interval		Estimate	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound			Lower Bound	Upper Bound
Sim	6750,845	313,010	6137,345	7364,344	.	.	.	.
Não	5121,479	235,736	4659,436	5583,522	5618,000	,000	.	.
Overall	6653,438	306,853	6052,006	7254,871	.	.	.	.

a. Estimation is limited to the largest survival time if it is censored.

Tabela 36 – Estimativas de KM para a variável – TAR

O p-value calculado no Teste de Breslow (Tabela 37) não permite rejeitar a hipótese das curvas de sobrevivência para os indivíduos que iniciaram TAR e que não iniciaram serem idênticas. Logo, não se assumem diferenças entre estas categorias pelo que a sua inclusão num modelo explicativo do tempo de sobrevivência não é estatisticamente significativa.

O período de maior risco de morte nos indivíduos que iniciaram TAR é até aos 4000 dias (cerca de 11 anos) após a data de diagnóstico. Para os restantes, o período de maior risco é até aos 1500 dias (cerca de 4 anos) após o diagnóstico da doença.



	Chi-Square	df	Sig.
Log Rank (Mantel-Cox)	,271	1	,603
Breslow (Generalized Wilcoxon)	,280	1	,597
Tarone-Ware	,209	1	,648

Test of equality of survival distributions for the different levels of TAR.

Figura 34 – Curvas de Sobrevivência de KM – TAR

Tabela 37 – Testes de comparação de curvas de sobrevivência- TAR



### 5.2.1.11 - Interrupção da TAR

Na sequência da não significância estatística da variável anterior, é expectável que o facto do indivíduo que iniciou TAR ter interrompido ou não o tratamento não seja também uma variável estatisticamente significativa (por existir na sequência da anterior) na explicação do tempo de sobrevivência em estudo.

A análise aos valores obtidos nas estimativas de KM (Tabela 38) indicam que, em média, os indivíduos que interrompem o tratamento sobrevivem menos 2 anos do que os que não o fazem.

Interrupt AR	Mean <sup>a</sup>				Median			
	Estimate	Std. Error	95% Confidence Interval		Estimate	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound			Lower Bound	Upper Bound
	Sim	5867,527	365,839	5150,484	6584,571	.	.	.
Não	6639,715	374,254	5906,177	7373,252	.	.	.	.
Overall	6750,845	313,010	6137,345	7364,344	.	.	.	.

a. Estimation is limited to the largest survival time if it is censored.

Tabela 38 – Estimativas de KM para a variável – Interrupção TAR

Da análise à representação gráfica das curvas de sobrevivência (Figura 35) verifica-se um aspecto curioso. A partir dos 4500 dias ou 12 anos (aproximadamente), os indivíduos que, tendo iniciado TAR não interromperam o tratamento têm um risco de morte superior.

O p-value no Teste de Breslow (*vide* tabela 39) não permite rejeitar a hipótese na qual as curvas são estatisticamente idênticas. Assim, o facto do indivíduo ter interrompido (de forma contínua ou não) o tratamento não influencia significativamente o seu tempo de sobrevivência após o diagnóstico.

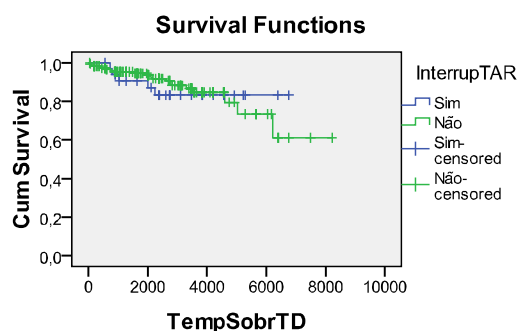


Figura 35 – Curvas de Sobrevivência de KM – Interrup. TAR

Overall Comparisons			
	Chi-Square	df	Sig.
Log Rank (Mantel-Cox)	,044	1	,835
Breslow (Generalized Wilcoxon)	,524	1	,469
Tarone-Ware	,368	1	,544

Test of equality of survival distributions for the different levels of InterrupTAR.

Tabela 39 – Testes de comparação de curvas de sobrevivência – Interrup. TAR

### 5.2.1.12 - Grupo Antiretrovirais

Da análise à Tabela 40, especificamente ao p-value obtido no Teste de Breslow (porque as curvas de sobrevivência das quatro categorias de antiretrovirais se cruzam), não rejeitamos a hipótese nula, donde a variável em análise não tem, por si só, influência estatisticamente significativa no tempo de sobrevivência após o diagnóstico.

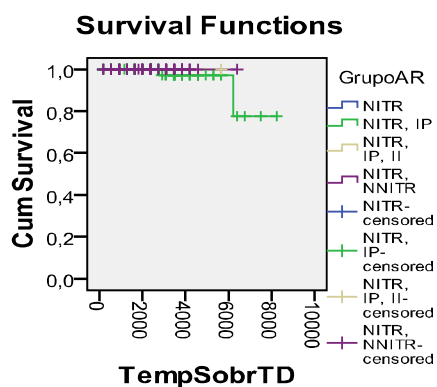


Figura 36 – Curvas de Sobrevivência de KM – Antiretrovirais

Overall Comparisons			
	Chi-Square	df	Sig.
Log Rank (Mantel-Cox)	,990	3	,804
Breslow (Generalized Wilcoxon)	,879	3	,831
Tarone-Ware	,971	3	,808

Test of equality of survival distributions for the different levels of GrupoAR.

Tabela 40 – Testes de comparação de curvas de sobrevivência – Antiretrovirais

### 5.2.1.13 - Carga Viral

No que respeita à carga viral dos indivíduos em estudo, em média, os que apresentam carga viral indetectável sobrevivem cerca de 21 anos após o diagnóstico da infecção, uma diferença de cerca de 6 anos face aos indivíduos com carga viral detectável (*vide* Tabela 41).

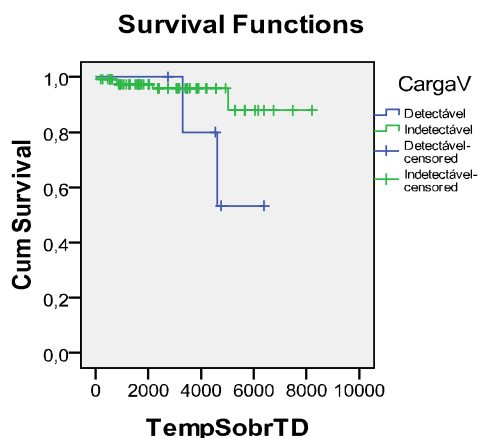
Means and Medians for Survival Time								
CargaV	Mean <sup>a</sup>				Median			
	Estimate	Std. Error	95% Confidence Interval		Estimate	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound			Lower Bound	Upper Bound
Detectável	5297,133	592,752	4135,339	6458,928	.	.	.	.
Indetectável	7675,803	280,399	7126,221	8225,385	.	.	.	.
Overall	7425,025	320,667	6796,517	8053,533	.	.	.	.

a. Estimation is limited to the largest survival time if it is censored.

Tabela 41 – Estimativas de KM para a variável – Carga Viral

A análise às curvas de sobrevivência (Figura 37) revela que o risco de morte é maior naqueles cuja carga viral é detectável, principalmente passados 10 anos (cerca de 3500 dias) após o diagnóstico.

O p-value no Teste de Breslow (Tabela 42) leva à não rejeição da respectiva hipótese nula, pelo que se admite que não existem diferenças significativas entre as categorias, não influenciando assim esta variável o tempo de sobrevivência após o diagnóstico de VIH/ SIDA.



Overall Comparisons			
	Chi-Square	df	Sig.
Log Rank (Mantel-Cox)	3,722	1	,054
Breslow (Generalized Wilcoxon)	,427	1	,513
Tarone-Ware	1,637	1	,201

Test of equality of survival distributions for the different levels of CargaV.

Figura 37 – Curvas de Sobrevivência de KM –Carga Viral

Tabela 42 – Testes de comparação de curvas de sobrevivência – Carga Viral

### 5.2.1.14 - Patologia: Tuberculose

Tal como evidenciam as estimativas de K-M (Tabela 43), os indivíduos que não foram diagnosticados com tuberculose sobrevivem, em média, cerca de 19 anos após o diagnóstico da infecção, enquanto que os doentes tuberculosos sobrevivem, também é média, cerca de 14 anos.

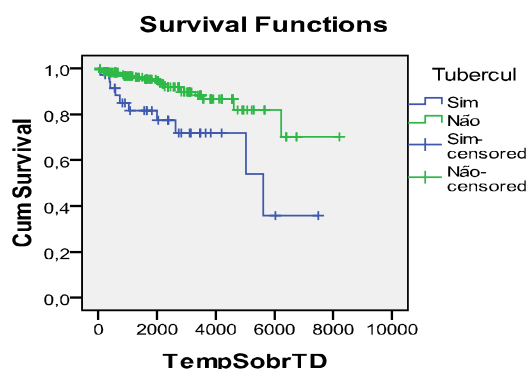
Means and Medians for Survival Time								
Tubercu l	Mean <sup>a</sup>				Median			
	Estimate	Std. Error	95% Confidence Interval		Estimate	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound			Lower Bound	Upper Bound
Sim	4933,148	600,283	3756,593	6109,704	5618,000	1541,354	2596,947	8639,053
Não	6993,424	318,864	6368,451	7618,397	.	.	.	.
Overall	6623,556	315,728	6004,730	7242,383	.	.	.	.

a. Estimation is limited to the largest survival time if it is censored.

Tabela 44 – Estimativas de KM para a variável – Tuberculose

Da análise às curvas de sobrevivência representadas na Figura 38, verifica-se que o risco de morte é menor nos indivíduos não tuberculosos desde que são diagnosticados. O período de maior risco para os indivíduos que apresentam tuberculose é até aos 2500 dias (cerca de 7 anos) e para os restantes esse período crítico vai até aos 3500 dias (cerca de 10 anos) após a data do diagnóstico de VIH/SIDA.

Na Tabela 44, no Teste de Breslow, cujo p-value leva à rejeição da  $H_0$ , pode afirmar-se que existem diferenças estatisticamente significativas entre o facto de ter ou não tuberculose. Desta forma esta é uma das variáveis a considerar no modelo de sobrevivência pois influencia a respectiva variável resposta.



Overall Comparisons			
	Chi-Square	df	Sig.
Log Rank (Mantel-Cox)	10,301	1	,001
Breslow (Generalized Wilcoxon)	10,508	1	,001
Tarone-Ware	10,606	1	,001

Test of equality of survival distributions for the different levels of Tubercul.

Figura 38 – Curvas de Sobrevivência de KM –Tuberculose

Tabela 44 – Testes de comparação de curvas de sobrevivência – Tuberculose

### 5.2.1.15 - Patologia: Hepatite B

Os indivíduos que não estão infectados com o vírus da Hepatite B parecem sobreviver menos tempo do que aqueles que o estão (*vide* Tabela 45). De facto, os primeiros sobrevivem cerca de 17 anos e os segundos cerca de 20 anos. Para este resultado inesperado poderá ter contribuído o facto de apenas 1 dos indivíduos que vieram a falecer de VIH/ SIDA ter apresentado Hepatite B.

Means and Medians for Survival Time								
HepB	Mean <sup>a</sup>				Median			
	Estimate	Std. Error	95% Confidence Interval		Estimate	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound			Lower Bound	Upper Bound
Sim	7478,091	707,383	6091,621	8864,561	.	.	.	.
Não	6063,258	274,205	5525,817	6600,698	.	.	.	.
Overall	6623,556	315,728	6004,730	7242,383	.	.	.	.

a. Estimation is limited to the largest survival time if it is censored.

Tabela 45 – Estimativas de KM para a variável – Hepatite B

O gráfico onde estão representadas as curvas de sobrevivência (Figura 39) revela que o risco de morte é superior para os indivíduos que não contraíram Hepatite B a partir dos 5 anos após o diagnóstico.

No entanto, o p-value no Teste de Breslow (Tabela 46) leva a afirmar que, estatisticamente, a variável "HepB" não é significativa para explicar o tempo de sobrevivência em estudo.

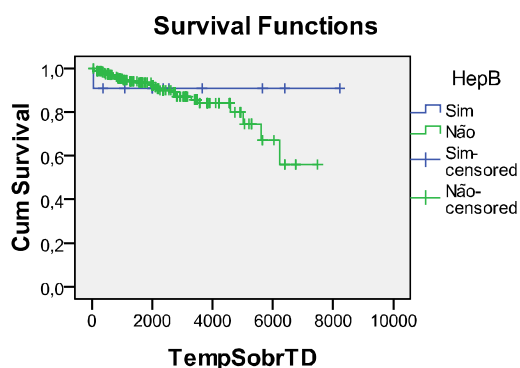


Figura 39 – Curvas de Sobrevivência de KM – Hepatite B

Overall Comparisons			
	Chi-Square	df	Sig.
Log Rank (Mantel-Cox)	,481	1	,488
Breslow (Generalized Wilcoxon)	,102	1	,750
Tarone-Ware	,005	1	,942

Test of equality of survival distributions for the different levels of HepB.

Tabela 46 – Testes de comparação de curvas de sobrevivência – Hepatite B

### 5.2.1.16 - Patologia: Hepatite C

Podendo estar outros factores a influenciar esta conclusão, na verdade, e de acordo com os valores obtidos nas estimativas de KM (*vide* Tabela 47) os indivíduos que apresentam o vírus da hepatite C sobrevivem mais tempo do que os restantes. Sobrevivem, em média, 18 anos, enquanto que os restantes sobrevivem menos 3 anos. Repare-se que os respectivos IC se intersectam.

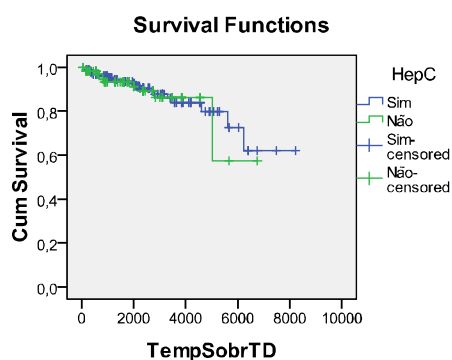
Means and Medians for Survival Time								
HepC	Mean <sup>a</sup>				Median			
	Estimate	Std. Error	95% Confidence Interval		Estimate	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound			Lower Bound	Upper Bound
Sim	6666,732	352,415	5975,999	7357,465	.	.	.	.
Não	5533,571	456,861	4638,123	6429,018	.	.	.	.
Overall	6623,556	315,728	6004,730	7242,383	.	.	.	.

a. Estimation is limited to the largest survival time if it is censored.

Tabela 47 – Estimativas de KM para a variável – Hepatite C

De facto, e de acordo com a representação gráfica das curvas de sobrevivência (Figura 40), o risco de morte parece ser idêntico para as duas categorias (ter ou não contraído VHC) até, aproximadamente, 4500 dias (cerca de 12 anos) após o diagnóstico de VIH/ SIDA. A partir daí, o risco é superior nos indivíduos não infectados por VHC.

O p-value = 0,793 calculado no Teste de Breslow (Tabela 48) leva a não rejeitar  $H_0$ , donde se admite que o facto do indivíduo estar ou não co-infectado por VHC e VIH/ SIDA não influencia o seu tempo de sobrevivência após o diagnóstico.



Overall Comparisons			
	Chi-Square	df	Sig.
Log Rank (Mantel-Cox)	,054	1	,816
Breslow (Generalized Wilcoxon)	,069	1	,793
Tarone-Ware	,057	1	,812

Test of equality of survival distributions for the different levels of HepC.

Figura 40 – Curvas de Sobrevivência de KM – Hepatite C

Tabela 48 – Testes de comparação de curvas de sobrevivência – Hepatite C

### 5.2.1.17 - Patologia: Depressão Nervosa

O IC calculado nas estimativas de KM (Tabela 49) indica que, em 95% das vezes, um indivíduo que apresente ou tenha apresentado depressão nervosa sobrevive, em média, 17 a 19 anos após ter sido diagnosticado com a doença. Enquanto que um indivíduo que nunca tenha apresentado aquela patologia sobreviva, em média, 16 a 19 anos.

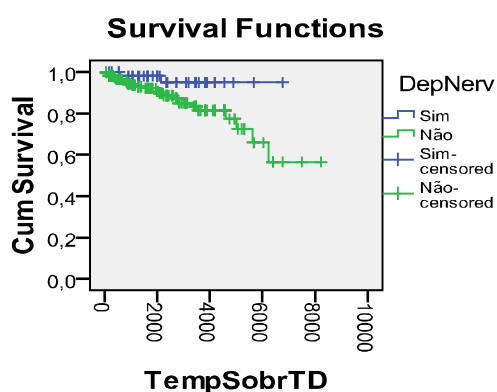
DepNerv	Means and Medians for Survival Time							
	Mean <sup>a</sup>				Median			
	Estimate	Std. Error	95% Confidence Interval		Estimate	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound			Lower Bound	Upper Bound
Sim	6494,379	183,175	6135,356	6853,401	.	.	.	.
Não	6378,299	350,287	5691,736	7064,861	.	.	.	.
Overall	6623,556	315,728	6004,730	7242,383	.	.	.	.

a. Estimation is limited to the largest survival time if it is censored.

Tabela 49 – Estimativas de KM para a variável – Depressão Nervosa

O gráfico da Figura 41 demonstra que o risco de morte é superior nos indivíduos que não apresentam sintomatologia de depressão nervosa.

Como as curvas de sobrevivência não se cruzam, interpreta-se o Teste Log Rank (Tabela 50), donde o resultado obtido para o p-value respectivo permite rejeitar a  $H_0$  de semelhança estatística entre as curvas. Logo, admite-se que existem diferenças significativas entre o facto de ter ou não depressão nervosa e o tempo que o indivíduo sobrevive após o diagnóstico de VIH/ SIDA.



	Chi-Square	df	Sig.
Log Rank (Mantel-Cox)	3,842	1	,050
Breslow (Generalized Wilcoxon)	2,944	1	,086
Tarone-Ware	3,340	1	,068

Test of equality of survival distributions for the different levels of DepNerv.

Figura 41 – Curvas de Sobrevivência de KM – Dep. Nervosa

Tabela 50 – Testes de comparação de curvas de sobrevivência – Dep. Nerv.

### 5.2.1.18 - Alcoolismo

De acordo com os valores da Tabela 51 os doentes que revelam hábitos alcoólicos sobrevivem, em média, 14 anos após o diagnóstico da infecção por VIH/ SIDA. De facto, 50% destes sobrevivem 15 anos. Os indivíduos não alcoólicos sobrevivem, em média, 19 anos.

Alcool	Mean <sup>a</sup>				Median			
	Estimate	Std. Error	95% Confidence Interval		Estimate	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound			Lower Bound	Upper Bound
Sim	5250,746	395,750	4475,076	6026,417	5618,000	2130,465	1442,289	9793,711
Não	6898,069	298,194	6313,609	7482,528	.	.	.	.
Overall	6629,787	315,366	6011,670	7247,905	.	.	.	.

a. Estimation is limited to the largest survival time if it is censored.

Tabela 51 – Estimativas de KM para a variável – Alcoolismo



O risco de morte parece ser idêntico tanto para os indivíduos alcoólicos como para os restantes até, aproximadamente, aos 12 anos após o diagnóstico de VIH/SIDA (*vide* Figura 42).

Como as curvas de sobrevivência se cruzam, analisa-se o Teste de Breslow (Tabela 52) cujo  $p\text{-value} > \alpha$  não permite rejeitar a  $H_0$ , logo assume-se que não existem diferenças significativas entre as categorias da variável em análise no que respeita ao tempo de sobrevivência dos indivíduos.

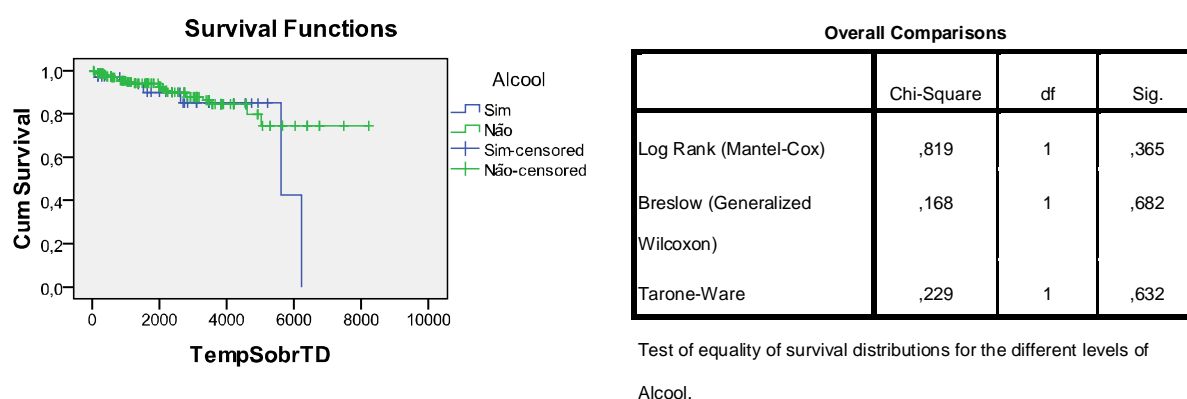


Figura 42 – Curvas de Sobrevivência de KM – Alcoolismo

Tabela 52 – Testes de comparação de curvas de sobrevivência – Álcool.

### 5.2.1.19 - Tabagismo

De acordo com a estimativa de Kaplan Meier (Tabela 53), em média, os indivíduos que revelam toxicomania associada ao consumo de tabaco, sobrevivem cerca de 13 anos, enquanto que os indivíduos que não sofrem daquela dependência sobrevivem cerca de 18 anos.

	Mean <sup>a</sup>				Median			
	Estimate	Std. Error	95% Confidence Interval		Estimate	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound			Lower Bound	Upper Bound
Tabag								
Sim	4771,859	296,799	4190,133	5353,585	.	.	.	.
Não	6595,808	323,897	5960,969	7230,646	.	.	.	.
Overall	6623,556	315,728	6004,730	7242,383	.	.	.	.

a. Estimation is limited to the largest survival time if it is censored.

Tabela 53 – Estimativas de KM para a variável – Tabagismo

Como as curvas de sobrevivência se cruzam (*vide* Figura 43) analisa-se o Teste de Breslow (Tabela 54) na comparação das curvas e o p-value obtido neste teste não permite rejeitar a  $H_0$  para  $\alpha = 25\%$ , donde não se considera que existam diferenças estatisticamente significativas no facto do indivíduo apresentar dependência física e psicológica do tabaco no tempo que sobrevive após ser diagnosticado com VIH/ SIDA.

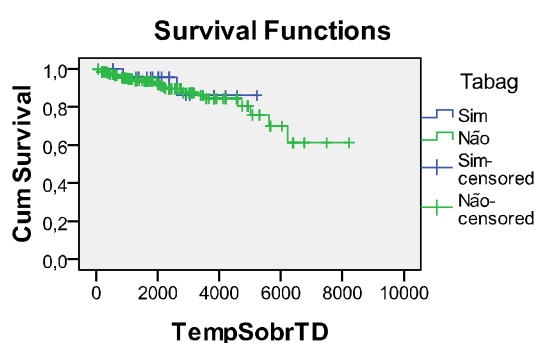


Figura 43 – Curvas de Sobrevivência de KM – Tabagismo

Overall Comparisons			
	Chi-Square	df	Sig.
Log Rank (Mantel-Cox)	,153	1	,695
Breslow (Generalized Wilcoxon)	,225	1	,635
Tarone-Ware	,183	1	,669

Test of equality of survival distributions for the different levels of Tabag.

Tabela 54 – Testes de comparação de curvas de sobrevivência – Tabag.

### 5.2.1.20 - Sector de Actividade

No que respeita ao sector de actividade em que o indivíduo desenvolve a sua profissão, os que se incluem no Sector 1 e/ou 2 (agricultura, pecuária e/ou indústria) sobrevivem em média, e de acordo com a estimativa de KM, mais 2 anos do que os indivíduos que trabalham no sector terciário (serviços), o que é um pouco contraditório pois à partida, os indivíduos incluídos no sector 3 têm maior acesso à informação e conhecimento do que os restantes.

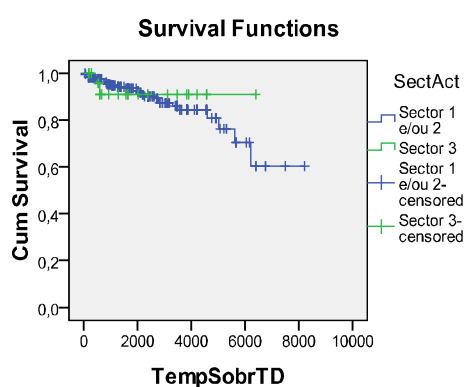
De facto, e segundo as estimativas de KM (Tabela 55), um indivíduo cuja profissão pertença ao Sector 3 sobrevive, em média, 16 anos.

SectAct	Mean <sup>a</sup>				Median			
	Estimate	Std. Error	95% Confidence Interval		Estimate	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound			Lower Bound	Upper Bound
Sector 1 e/ou 2	6607,083	326,324	5967,488	7246,678	.	.	.	.
Sector 3	5865,629	356,915	5166,075	6565,183	.	.	.	.
Overall	6669,048	307,486	6066,375	7271,722	.	.	.	.

a. Estimation is limited to the largest survival time if it is censored.

Tabela 55 – Estimativas de KM para a variável – Sector de Actividade

A análise à representação gráfica das curvas de sobrevivência (Figura 44) permite verificar que, aproximadamente, até aos 2000 dias (cerca de 5,5 anos) o risco de morte é inferior para os indivíduos pertencentes ao sector 1 e/ou 2. A partir dessa altura a situação parece inverter-se.



	Chi-Square	df	Sig.
Log Rank (Mantel-Cox)	,078	1	,780
Breslow (Generalized Wilcoxon)	,059	1	,808
Tarone-Ware	,003	1	,955

Test of equality of survival distributions for the different levels of SectAct.

Figura 44 – Curvas de Sobrevivência de KM – Sector Act.

Tabela 56 – Testes de comparação de curvas de sobrevivência – Sector Act.

### 5.2.1.21 - Concelho de Residência

De acordo com a representação gráfica das curvas de sobrevivência (Figura 45), existem alguns concelhos cujos dados são integralmente do tipo censurado (é o caso dos concelhos de Almodôvar, Alvito, Barrancos, Castro Verde, Vidigueira e Outros não pertencentes ao distrito de

Beja). Dos restantes, o risco de morte parece ser superior para os indivíduos com residência no concelho de Serpa e Moura, e menor para os residentes nos concelhos de Odemira e Mértola.

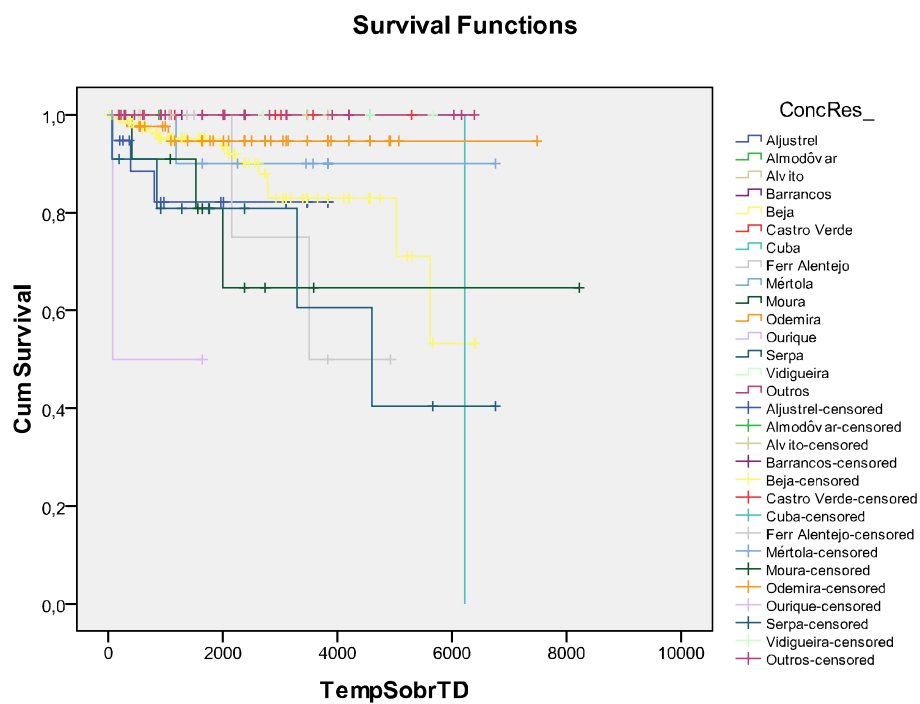


Figura 45 – Curvas de Sobrevivência de KM – Concelho de Residência

No que respeita à importância que a variável Concelho de Residência assume no tempo de sobrevivência do indivíduo após o diagnóstico, o p-value no Teste de Breslow (Tabela 57) permite rejeitar a  $H_0$ . Assim, a variável é significativa na modelação do tempo de sobrevivência em análise.

**Overall Comparisons**

	Chi-Square	df	Sig.
Log Rank (Mantel-Cox)	28,066	14	,014
Breslow (Generalized Wilcoxon)	29,944	14	,008
Tarone-Ware	29,495	14	,009

Test of equality of survival distributions for the different levels of ConcRes\_.

Tabela 57 – Testes de comparação de curvas de sobrevivência – Conc. Resid.

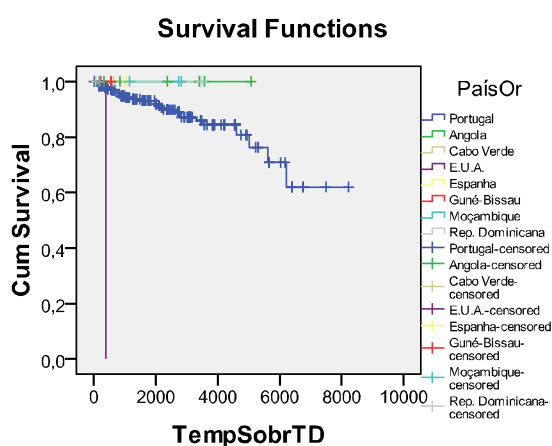
### 5.2.1.22 - Distrito de Residência: Beja

No que respeita ao distrito onde o indivíduo reside (Distrito de Beja ou Outros), verifica-se que todos os dados não censurados correspondem a doentes residentes no distrito de Beja. Por essa razão não se considera que esta variável seja devidamente explicativa no modelo que se visa encontrar.

### 5.2.1.23 - País de Origem

Quando observados os dados verifica-se que dos não censurados, 30 pertencem a Portugal e apenas 1 pertence a outro país, neste caso E.U.A. Por tal razão as estimativas KM não foram calculadas.

O p-value no Teste de Breslow (Tabela 58) é tal que se rejeita a hipótese das curvas serem estatisticamente idênticas. Por esse facto a variável País de Origem (dos 9 países em estudo) considera-se significativa.



Overall Comparisons			
	Chi-Square	df	Sig.
Log Rank (Mantel-Cox)	38,513	7	,000
Breslow (Generalized Wilcoxon)	33,623	7	,000
Tarone-Ware	36,049	7	,000

Test of equality of survival distributions for the different levels of PaísOr.

Figura 46 – Curvas de Sobrevivência de KM –País Origem

Tabela 58 – Testes de comparação de curvas de sobrevivência – País Origem

Veja-se se criando outra variável reduzindo o número de categorias da variável “PaísOr” se identifica outra possível forma de considerar o efeito do país de onde o indivíduo provém no seu tempo de sobrevivência após o diagnóstico de VIH/SIDA.

### 5.2.1.24- País de Origem: Portugal

De acordo com o exposto na Tabela 59 a estimativa de KM para a média indica que o indivíduo cujo país de origem não é Portugal apresenta cerca de 5 anos a menos (18 anos para categoria Portugal e 13 anos para categoria Outros) no tempo de sobrevivência após o diagnóstico de VIH/SIDA. De realçar a expectável não significância desta variável na explicação do tempo de sobrevivência em estudo uma vez que ocorreu apenas um evento num indivíduo fora de Portugal.

PaísOrPort	Mean <sup>a</sup>				Median			
	Estimate	Std. Error	95% Confidence Interval		Estimate	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound			Lower Bound	Upper Bound
Portugal	6626,523	311,728	6015,535	7237,511	.	.	.	.
Outros - Não Portugal	4736,786	322,057	4105,554	5368,017	.	.	.	.
Overall	6653,438	306,853	6052,006	7254,871	.	.	.	.

a. Estimation is limited to the largest survival time if it is censored.

Tabela 59 – Estimativas de KM para a variável – País de Origem Portugal

O p-value no Teste de Breslow (Tabela 60) não permite rejeitar a respectiva  $H_0$ , donde se pode afirmar estatisticamente que não existe diferença significativa entre o facto do indivíduo ser proveniente de Portugal ou de outro país (não discriminado).

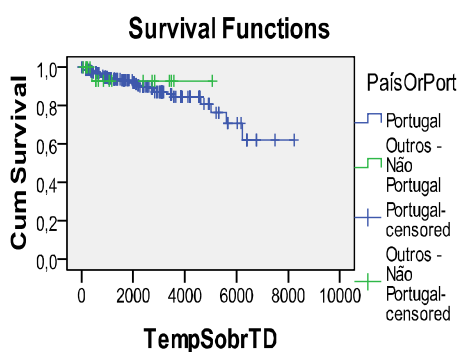


Figura 47 – Curvas de Sobrevivência de KM –País Portugal

	Chi-Square	df	Sig.
Log Rank (Mantel-Cox)	,133	1	,715
Breslow (Generalized Wilcoxon)	,003	1	,953
Tarone-Ware	,035	1	,852

Test of equality of survival distributions for the different levels of PaísOrPort.

Tabela 60 – Testes de comparação de curvas de sobrevivência – País Portugal

### 5.2.1.25 - Consulta de Enfermagem

De acordo com as estimativas de KM calculadas (*vide* Tabela 61), os indivíduos que frequentam a consulta de enfermagem de apoio no acompanhamento hospitalar sobrevivem, em média, 20 anos, e em contrapartida, os indivíduos que não frequentam a consulta apresentam um tempo de sobrevivência de, em média, 14 anos, o que perfaz uma diferença de 6 anos face àqueles, e os respectivos IC a 95% não se intersectam.

A frequência da consulta de enfermagem parece ser um factor importante na definição do tempo de sobrevivência do indivíduo.

ConsEnf	Mean <sup>a</sup>				Median			
	Estimate	Std. Error	95% Confidence Interval		Estimate	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound			Lower Bound	Upper Bound
Sim	7134,363	364,922	6419,115	7849,610	.	.	.	.
Não	5140,374	258,152	4634,397	5646,351	.	.	.	.
Overall	6650,466	306,961	6048,822	7252,111	.	.	.	.

a. Estimation is limited to the largest survival time if it is censored.

Tabela 61 – Estimativas de KM para a variável – Consulta de Enfermagem

Do gráfico que representa as curvas de sobrevivência (Figura 48) ressalta o facto do risco de morte ser visivelmente superior para aqueles que não frequentam a consulta de enfermagem.

De facto, o p-value e respectiva rejeição da  $H_0$  no Teste de Breslow (Tabela 62) permite afirmar que existem diferenças estatisticamente significativas no facto do doente frequentar ou não a consulta de enfermagem no seu tempo de sobrevivência após o diagnóstico de VIH/SIDA.

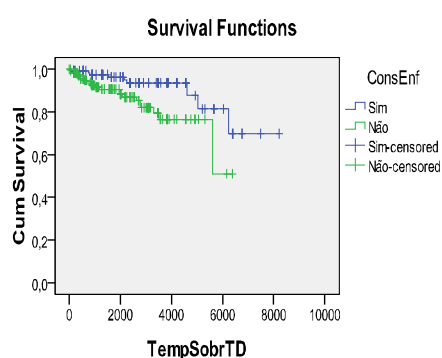


Figura 48 – Curvas de Sobrevivência de KM – Cons. Enf.

	Chi-Square	df	Sig.
Log Rank (Mantel-Cox)	6,732	1	,009
Breslow (Generalized Wilcoxon)	5,726	1	,017
Tarone-Ware	6,406	1	,011

Test of equality of survival distributions for the different levels of ConsEnf.

Tabela 62 – Testes de comparação de curvas de sobrevivência – Cons. Enf.

### 5.2.1.26 - Consulta de Psicologia

Em média, o indivíduo que frequenta a consulta de enfermagem sobrevive cerca de 17 anos e o indivíduo que não frequenta essa consulta sobrevive cerca de 18 anos. Os respectivos IC intersectam-se, e por conseguinte, não é totalmente eficiente analisar o impacto desta variável através da média calculada e apresentada na Tabela 63.

ConsPsic	Mean <sup>a</sup>				Median			
	Estimate	Std. Error	95% Confidence Interval		Estimate	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound			Lower Bound	Upper Bound
Sim	6083,215	313,986	5467,802	6698,627	.	.	.	.
Não	6621,397	339,935	5955,125	7287,670	.	.	.	.
Overall	6653,438	306,853	6052,006	7254,871	.	.	.	.

a. Estimation is limited to the largest survival time if it is censored.

Tabela 63 – Estimativas de KM para a variável – Consulta de Psicologia

O risco de morte é superior para os indivíduos que não frequentam a consulta de Psicologia até aos 18 anos após o diagnóstico da infecção por VIH/SIDA. A partir dessa altura esse risco passa a ser inferior face aos indivíduos que recorrem ao psicólogo.



O teste de Breslow (Tabela 64) calculado para a comparação das curvas de sobrevivência (que se cruzam) e respectivo valor obtido para o p-value leva à rejeição da hipótese nula, pelo que se assume que o facto de frequentar a consulta de Psicologia tem influência no tempo de sobrevivência do indivíduo infectado por VIH/SIDA após o respectivo diagnóstico.

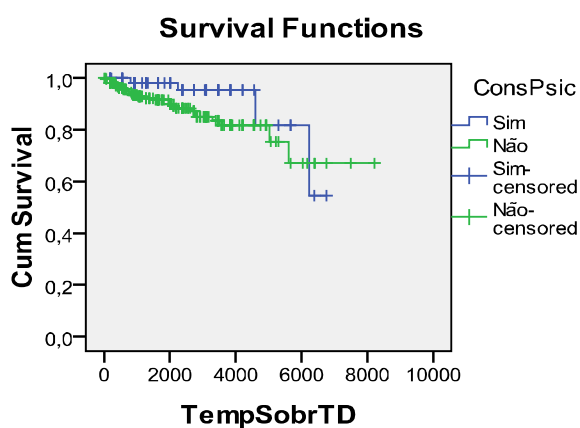


Figura 49 – Curvas de Sobrevivência de KM – Cons. Psic.

Overall Comparisons			
	Chi-Square	df	Sig.
Log Rank (Mantel-Cox)	2,552	1	,110
Breslow (Generalized Wilcoxon)	3,832	1	,050
Tarone-Ware	3,713	1	,054

Test of equality of survival distributions for the different levels of ConsPsic.

Tabela 64 – Testes de comparação de curvas de sobrevivência – Cons. Psic.

### 5.2.2 – Análise Univariada – Variáveis categóricas – Modelo de regressão de Cox

Para avaliar a sua significância no tempo de sobrevivência dos indivíduos após o diagnóstico, considera-se um modelo Cox simples e respectivo Teste Wald ( $H_0: \beta_i = 0$ ) (e/ou Teste da razão de verossimilhanças, se necessário) para as variáveis categóricas que se consideram importantes de avaliar o respectivo risco relativo no tempo de sobrevivência.

Considere-se um nível de significância ( $\alpha$ ) de 25%.

A interpretação de cada um dos resultados do risco de morte nas categorias consideradas significativas é válida mantendo tudo o resto constante, *ceteris paribus*.

#### 5.2.2.1- Categoria de Contagem de Células CD4

De acordo com os valores obtidos (*vide* Tabela 65) a variável aqui analisada considera-se significativa (p-value no Teste Wald permite rejeitar a respectiva hipótese nula).

Variables in the Equation								
	B	SE	Wald	df	Sig.	Exp(B)	95.0% CI for Exp(B)	
							Lower	Upper
CatCD4			11,142	2	,004			
CatCD4(1)	-1,551	1,026	2,286	1	,131	,212	,028	1,583
CatCD4(2)	-1,436	,463	9,618	1	,002	,238	,096	,590

Tabela 65 – Estimativas Modelo de Cox para a covariável – CatCD4

Devem permanecer em estudo as 3 categorias que constituem esta variável. Assim, da análise à razão de riscos verifica-se que num indivíduo que no início do acompanhamento apresentasse um número de células CD4 entre 200 a 499 células por microlitro de sangue ( $\exp(B) = 0,238$ ), o risco de morte diminui em aproximadamente 76% face a um indivíduo cuja contagem tenha resultado em valores “ $<200$ ” (considerando esta como categoria de referência). O respectivo IC a 95% indica que essa diminuição no risco de morte pode variar entre os 41% e 90%. Interpretando esse resultado de outra forma, pode igualmente afirmar-se que o risco de morte aumenta cerca de 4 vezes se o indivíduo obtiver uma contagem de células CD4 inferior a 200 células face a outro indivíduo cuja contagem se inclua na categoria 200-499 células por microlitro de sangue.

Por outro lado, para resultados de contagem de células CD4 acima das 500 células por microlitro de sangue o risco de morte desse indivíduo diminui em cerca de 79% (pois  $\exp(B) = 0,212$ ) face a um resultado de contagem que se inclua na categoria  $<200$  células. Também aqui se pode afirmar que o risco de morte aumenta aproximadamente 5 vezes quando comparados os indivíduos que se incluíam na categoria  $<200$  face à categoria  $\geq 500$  células por microlitro de sangue.

### 5.2.2.2 - Concelho de Residência

Esta variável é particularmente grande no sentido em que é constituída por 15 categorias. As primeiras 14 dizem respeito aos Concelhos que pertencem ao Distrito de Beja e a última categoria inclui todos os outros. Apure-se então a sua significância na explicação do tempo de sobrevivência do indivíduo.

De facto, quando analisado o p-value associado à variável “Concelho de Residência” e respectivas categorias no Teste Wald, considerando a categoria “Outros” como referência, verifica-se que nenhuma é significativa (respectivos p-value > 25% e os 95% IC da razão de riscos Exp (B) incluem a unidade). Esta variável deixa portanto de considerar-se importante no modelo.

Neste caso não se considerou útil colapsar categorias e avaliar o impacto dessa operação na significância da variável concelho de residência. O objectivo em contrapor concelhos era justamente analisar a importância de dado concelho na explicação do tempo de sobrevivência dos indivíduos (caso contrário considerar-se-ia apenas o distrito de residência), e por essa razão não faz sentido agrupar concelhos, até porque não haveria uma regra de agrupamento justificável.

### 5.2.2.3 - País de Origem

Embora a variável em si seja significativa, contém categorias que não o são. De facto, apenas a categoria “E.U.A.” parece ser explicativa. Vejam-se os resultados na Tabela 66, apenas naquela categoria o p-value no Teste Wald permite rejeitar a respectiva hipótese nula e o IC da razão de risco não inclui a unidade.

Mas, como já referido anteriormente, tal deve-se ao facto de se ter registado apenas um evento num indivíduo cuja residência é fora de Portugal, e como tal a respectiva interpretação do risco relativo está condicionada.

Variables in the Equation								
	B	SE	Wald	df	Sig.	Exp(B)	95.0% CI for Exp(B)	
							Lower	Upper
PaísOr			12,092	7	,098			
PaísOr(1)	-12,238	517,367	,001	1	,981	,000	,000	.
PaísOr(2)	-12,382	2605,308	,000	1	,996	,000	,000	.
PaísOr(3)	3,756	1,080	12,091	1	,001	42,790	5,150	355,509
PaísOr(4)	-12,283	1901,667	,000	1	,995	,000	,000	.
PaísOr(5)	-12,417	1905,738	,000	1	,995	,000	,000	.
PaísOr(6)	-12,253	799,555	,000	1	,988	,000	,000	.
PaísOr(7)	-12,232	1163,855	,000	1	,992	,000	,000	.

Tabela 66 – Estimativas Modelo de Cox para a covariável – País de Origem

Vejamos se considerando apenas 3 categorias (Portugal, E.U.A., e Outros) o que acontece (Tabela 67).

Variables in the Equation								
	B	SE	Wald	df	Sig.	Exp(B)	95.0% CI for Exp(B)	
							Lower	Upper
PaisOr_nova			12,092	2	,002			
PaisOr_nova(1)	3,756	1,080	12,091	1	,001	42,790	5,150	355,509
PaisOr_nova(2)	-12,253	385,142	,001	1	,975	,000	,000	.

Tabela 67 – Estimativas Modelo de Cox para a covariável – País de Origem modificada

Mais uma vez apenas a categoria “E.U.A.” é significativa. Por essa razão aquele país está aqui a assumir tanta importância. Ou seja, não é o país em questão que está a influenciar o resultado mas sim a forma como os dados estão distribuídos (e o facto de serem efectivamente reais).

Se se considerar uma variável apenas com as categorias “1 = E.U.A.” e “2 = Outros”, até resulta significativa (*vide* Tabela 68) mas não faz muito sentido considerar esta característica.

O risco de morte de um indivíduo originário dos E.U.A. parece ser 6,7 vezes superior quando comparado com um indivíduo de qualquer outro país. O respectivo IC a 95% indica que aquele valor pode variar dos 2,3 às 19,4 vezes.

Variables in the Equation								
	B	SE	Wald	df	Sig.	Exp(B)	95.0% CI for Exp(B)	
							Lower	Upper
PaisEUA	1,907	,540	12,466	1	,000	6,733	2,336	19,409

Tabela 68 – Estimativas Modelo de Cox para a covariável – País de Origem EUA

No entanto, não faz sentido considerar distinção entre um doente seguido no Hospital de Beja cujo país de origem seja ou não os E.U.A., e como tal a variável não “PaísEUA” não entrará no modelo. Nem mesmo a variável inicial “PaísOr” uma vez que, analisados os respectivos dados verifica-se que a maioria os indivíduos provêm mesmo de Portugal.

### 5.2.3 – Análise Univariada – Variáveis Contínuas – Modelo de regressão de Cox

#### 5.2.3.1 - Idade de entrada

Na Tabela 69 verifica-se que a variável “IdadeEnt” não é significativa no modelo. De facto, o p-value no Teste Wald não permite rejeitar a respectiva hipótese nula, sendo a razão de risco dada por Exp (B) aproximadamente igual a 1. Por conseguinte, não há utilidade estatística em considerar, por enquanto, esta variável na explicação do tempo de sobrevivência.

Variables in the Equation								
	B	SE	Wald	df	Sig.	Exp(B)	95.0% CI for Exp(B)	
							Lower	Upper
IdadeEnt	,019	,019	1,021	1	,312	1,019	,983	1,057

Tabela 69 – Estimativas Modelo de Cox para a covariável – País de Origem EUA

Veja-se se categorizando a variável “IdadeEnt” esta se torna significativa no modelo. Após algumas tentativas, verifica-se que a idade de entrada distribuída pelos grupos etários: 0-55 e 56-100 anos é significativa.

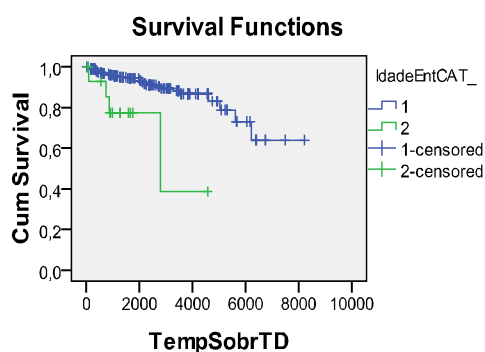
Alguns estudos defendem que a infecção VIH/SIDA por ser agravada pelas patologias que normalmente surgem no indivíduo especialmente após os 50 anos. De facto, entre as Recomendações Portuguesas para o Tratamento da Infecção VIH/Sida (CNLCS, 2009), com base em estudos que foram elaborados por coorte, encontra-se a idade do doente, neste caso “> 55 anos”, como indicação de início da terapêutica, porque a partir desta idade podem surgir mais complicações clínicas. O documento da Coordenação Nacional para a Infecção VIH/ SIDA recomenda que em indivíduos com contagem de células CD4 entre 350 e 500/mm<sup>3</sup>, idade superior a 55 anos, com carga vírica elevada ou declínio rápido das células CD4, e ainda co-infectado por VHC ou VHB, deve iniciar TAR.

A nova variável “IdadeEntCAT\_” é então significativa no modelo (ver p-value <  $\alpha$  na Tabela 71). De acordo com o p-value obtido (no Teste Log Rank) permite rejeitar a  $H_0$ , pelo que se pode afirmar que existem diferenças significativas entre as categorias.

IdadeEnt CAT_	Mean <sup>a</sup>				Median			
	Estimate	Std. Error	95% Confidence Interval		Estimate	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound			Lower Bound	Upper Bound
<=55	6771,007	309,027	6165,314	7376,700	.	.	.	.
>55	2972,887	605,859	1785,403	4160,371	2781,000	1384,515	67,350	5494,650
Overall	6675,444	307,230	6073,273	7277,615	.	.	.	.

a. Estimation is limited to the largest survival time if it is censored.

Tabela 70 – Estimativas de KM para a variável – Idade de Entrada categorizada



	Chi-Square	df	Sig.
Log Rank (Mantel-Cox)	11,168	1	,001
Breslow (Generalized Wilcoxon)	8,771	1	,003
Tarone-Ware	9,879	1	,002

Test of equality of survival distributions for the different levels of IdadeEntCAT\_.

Figura 50 – Curvas de Sobrevivência de KM – Idade Ent. Cat.

Tabela 71 – Testes de comparação de curvas de sobrevivência – Idade Ent.Cat.

Das estimativas de KM (Tabela 71) retira-se que, em média, um indivíduo com mais do que 55 anos sobrevive menos 10 anos do que um indivíduo mais novo. De facto, com idade de entrada no acompanhamento até aos 55 anos o tempo médio de sobrevivência é de 18 anos [95% IC (17; 20)], e se tiver mais do que 55 anos o seu tempo médio de sobrevivência é de aproximadamente 8 anos [95% IC (5; 11)].

De facto, de acordo os valores obtidos no cálculo da razão de riscos (*vide* Tabela 72), o risco de morte num indivíduo com mais do que 55 anos aquando da entrada no acompanhamento

hospitalar é 5,2 vezes superior face a um indivíduo com menos do que 55 anos. O respectivo IC a 95% indica que aquele risco pode variar entre 1,7 e 15,5 vezes.

Variables in the Equation								
	B	SE	Wald	df	Sig.	Exp(B)	95.0% CI for Exp(B)	
							Lower	Upper
IdadeEntCAT_	1,656	,554	8,955	1	,003	5,240	1,771	15,506

Tabela 72 – Estimativas Modelo de Cox para a covariável – Idade de Entrada categorizada

### 5.2.3.2 - Idade de diagnóstico

De acordo com o p-value obtido no Teste Wald (*vide* Tabela 73), a variável é significativa na explicação do tempo de sobrevivência.

Por cada ano a mais de vida na idade de diagnóstico o risco de morte aumenta 3,8% (com IC entre os 0,9% e os 6,8%), independentemente da idade que o doente tenha, uma vez confirmado o pressuposto da proporcionalidade de riscos. Outra forma de analisar os valores obtidos permite afirmar que por cada 10 anos a mais na idade aquando do diagnóstico do VIH/SIDA o risco de morte aumenta cerca de 45% (pois  $\exp 10 (0,037) = 1,447$ ).

Variables in the Equation								
	B	SE	Wald	df	Sig.	Exp(B)	95.0% CI for Exp(B)	
							Lower	Upper
IdadeDiag	,037	,015	6,607	1	,010	1,038	1,009	1,068

Tabela 73 – Estimativas Modelo de Cox para a covariável – Idade de Diagnóstico

### 5.2.3.3- Idade no início da TAR

De acordo com o resultado no Teste Wald (*vide* Tabela 74), a variável é significativa.

Variables in the Equation								
	B	SE	Wald	df	Sig.	Exp(B)	95.0% CI for Exp(B)	
							Lower	Upper
IdadeTAR	,030	,019	2,450	1	,118	1,031	,992	1,071

Tabela 74 – Estimativas Modelo de Cox para a covariável – Idade de Início TAR

No entanto, a esta variável está inerente uma restrição na base de dados, isto é, só é válida e útil no caso dos indivíduos que efectivamente iniciaram a TAR durante o período em análise. Desta forma, não entrará no modelo como uma variável explicativa, tal que se estuda apenas a significância da variável que lhe deu origem: Início da TAR (Sim; Não).

#### 5.2.4 – Resumo Análise Univariada

##### ➤ Variáveis categóricas via Kaplan-Meier

Tendo considerado um nível de significância de 25%, da análise efectuada via KM resultaram significativas as variáveis que constam da Tabela 75.

	Variável								
	Heterossexual	Categoria Clínica	Cat. Células CD4	Tuberculose	Depressão Nervosa	Concelho Residência	País de origem	Consulta Enfermagem	Consulta Psicologia
n.º de categorias	2	2	3	2	2	15	9	2	2
curvas de sobrevivência cruzam-se	sim	não	sim	sim	não	sim	sim	sim	sim
p-value no Teste Breslow	0,172	-	0,000	0,001	-	0,008	0,000	0,017	0,050
p-value no Teste Log-Rank	-	0,000	-	-	0,050	-	-	-	-

Tabela 75 – Resumo das variáveis significativas – Análise Univariada via KM

##### ➤ Variáveis categóricas via Modelo Cox simples

Ainda estabelecendo o nível de significância de 25%, e considerando como categoria de referência a última, das variáveis categóricas analisadas via Cox resultaram significativas a variável Categoria de Contagem células CD4 em todas as suas categorias (*vide* Tabela 76).

	CatCD4	CatCD4(1)	CatCD4(2)
p-value no Teste Wald	0,004	0,131	0,002

Tabela 76 – Resultados Teste Wald para variável CatCD4

A variável Concelho de Residência veio a ser eliminada, e a variável País de Origem, embora significativa *a priori*, apenas uma das suas categorias (E.U.A.) revelou significância estatística, donde se considerou, por ter sido influenciada por uma observação isolada e por não fazer



sentido considerar como factor de distinção a proveniência dos E.U.A, que não entra também no modelo.

### ➤ Variáveis contínuas via Modelo Cox simples

A variável Idade de Entrada começou por resultar não significativa. Tentou-se por isso a sua categorização, tal que a nova variável Idade de Entrada Categorizada (com duas categorias:  $\leq 55$  anos e  $> 55$  anos) é significativa (Tabela 77).

A variável Idade no Diagnóstico foi também considerada significativa.

	Idade Diag.	Idade Ent. Cat.
p-value no Teste Wald	0,010	0,003

Tabela 77 – Resultados Teste Wald para variáveis Idade de Diagnóstico e Idade de Entrada categorizada

### ➤ Geral

No geral, foram consideradas significativas um total de 9 variáveis, das quais 8 categóricas e 1 contínua. Para esta última verificar-se-á oportunamente se o pressuposto da linearidade é válido.

## 5.3 – Análise Multivariada

### 5.3.1– Inclusão variáveis significativas

Tendo sido apuradas as variáveis que, individualmente, influenciam de forma significativa a variável resposta em estudo (Tempo de sobrevivência após o diagnóstico da infecção VIH/SIDA), considera-se agora a sua inclusão conjunta num modelo de sobrevivência.

De acordo com o aconselhado por alguns autores como Hosmer e Lemeshow considera-se nesta fase um nível de significância  $\alpha$  de 10%.

Os valores obtidos para o modelo inicial constam das Tabelas 78 e 79.

Omnibus Tests of Model Coefficients<sup>a,b</sup>

-2 Log Likelihood	Overall (score)			Change From Previous Step			Change From Previous Block		
	Chi-square	df	Sig.	Chi-square	df	Sig.	Chi-square	df	Sig.
203,294	50,810	10	,000	48,106	10	,000	48,106	10	,000

a. Beginning Block Number 0, initial Log Likelihood function: -2 Log likelihood: 251.399

b. Beginning Block Number 1. Method = Enter

Tabela 78 – Resultados do modelo inicial estimado

Relativamente à significância estatística das variáveis incluídas, verifica-se que existem 6 variáveis que, na presença das restantes, não parecem ser explanatórias, uma vez que o respectivo p-value no Teste Wald resultou superior ao  $\alpha$  considerado (*vide* Tabela 79). São as variáveis Heterossexual (p-value = 0,686), Categoria Clínica (p-value = 0,457), Consulta Psicologia (p-value = 0,362), Idade de Entrada categorizada (p-value = 0,291), Idade no Diagnóstico (p-value = 0,278) e Depressão Nervosa (p-value = 0,162). Verifique-se também que embora a variável Categoria de Contagem Células CD4 (CatCD4) seja significativa nesta fase, uma das suas categorias não o é. No entanto, pelas razões entretanto já apresentadas, a variável permanecerá desta forma no modelo.

Variables in the Equation

	B	SE	Wald	df	Sig.	Exp(B)	95.0% CI for Exp(B)	
							Lower	Upper
Heteross	,261	,644	,164	1	,686	1,298	,367	4,590
CatClinic	-,411	,552	,554	1	,457	,663	,224	1,957
<b>CatCD4</b>			<b>10,526</b>	<b>2</b>	<b>,005</b>			
CatCD4(1)	-1,599	1,075	2,214	1	,137	,202	,025	1,660
CatCD4(2)	-1,845	,594	9,659	1	,002	,158	,049	,506
Tubercul	1,116	,458	5,932	1	,015	3,053	1,244	7,496
DepNerv	-1,110	,794	1,953	1	,162	,330	,069	1,564
ConsEnf	-1,323	,483	7,493	1	,006	,266	,103	,687
ConsPsic	-,623	,683	,832	1	,362	,536	,141	2,046
IdadeEntCAT_	-1,162	1,101	1,115	1	,291	,313	,036	2,705
IdadeDiag	,032	,029	1,176	1	,278	1,032	,975	1,093

Tabela 79 – Estimativas Modelo de Cox – inicial

Comece-se por retirar a variável com maior p-value (“Heteross”) e calcule-se o respectivo efeito no modelo através do Teste da Razão de Verosimilhanças.

A diferença entre as Deviances (-2LL) de ambos os modelos (ver Deviance do modelo reduzido na Tabela 80) leva a  $G = 0,173$ , donde o valor p numa  $\chi_1^2$  é igual a 0,6775. Desta forma, não se rejeita a hipótese nula do teste, tal que se o modelo sem a variável “Heteross” é tão explicativo quanto o modelo completo. Elimina-se portanto esta variável.

Omnibus Tests of Model Coefficients<sup>a,b</sup>

-2 Log Likelihood	Overall (score)			Change From Previous Step			Change From Previous Block		
	Chi-square	df	Sig.	Chi-square	df	Sig.	Chi-square	df	Sig.
203,467	50,378	9	,000	47,933	9	,000	47,933	9	,000

a. Beginning Block Number 0, initial Log Likelihood function: -2 Log likelihood: 251.399

b. Beginning Block Number 1. Method = Enter

Tabela 80 – Valor Deviance -2LL do modelo estimado

Depois de retirada a variável “Heteross”, continuam não significativas as variáveis acima referidas. Confirme os valores obtidos através da Tabela 81.

Variables in the Equation

	B	SE	Wald	df	Sig.	Exp(B)	95.0% CI for Exp(B)	
							Lower	Upper
CatClinic	-,456	,540	,712	1	,399	,634	,220	1,826
CatCD4			10,532	2	,005			
CatCD4(1)	-1,605	1,072	2,241	1	,134	,201	,025	1,643
CatCD4(2)	-1,834	,592	9,607	1	,002	,160	,050	,510
Tubercul	1,142	,456	6,267	1	,012	3,133	1,281	7,660
DepNerv	-1,099	,794	1,918	1	,166	,333	,070	1,578
ConsEnf	-1,329	,482	7,599	1	,006	,265	,103	,681
ConsPsic	-,582	,675	,743	1	,389	,559	,149	2,098
IdadeEntCAT_	-1,181	1,101	1,152	1	,283	,307	,035	2,654
IdadeDiag	,031	,029	1,128	1	,288	1,031	,974	1,092

Tabela 81 – Estimativas Modelo de Cox sem variável Heterossexual

Retire-se então a variável “CatClic” e apure-se o respectivo efeito no modelo.

A diferença na Deviance para com o modelo anterior resulta em  $G = 204,203 - 203,467 = 0,736$ , donde o p-value para  $\chi^2_1$  não permite mais uma vez rejeitar a hipótese nula (p-value = 0,3909). Considera-se por isso que a variável “CatClic” pode ser retirada do modelo pois não afecta o ajustamento deste.

Verifique-se que o p-value nas restantes variáveis é cada vez menor (Tabela 82).

Variables in the Equation								
	B	SE	Wald	df	Sig.	Exp(B)	95.0% CI for Exp(B)	
							Lower	Upper
CatCD4			16,052	2	,000			
CatCD4(1)	-1,815	1,042	3,037	1	,081	,163	,021	1,254
CatCD4(2)	-2,036	,543	14,068	1	,000	,131	,045	,378
Tubercul	1,292	,425	9,239	1	,002	3,640	1,582	8,375
DepNerv	-1,152	,792	2,117	1	,146	,316	,067	1,492
ConsEnf	-1,287	,472	7,439	1	,006	,276	,110	,696
ConsPsic	-,703	,660	1,132	1	,287	,495	,136	1,807
IdadeEntCAT_	-1,039	1,074	,935	1	,334	,354	,043	2,906
IdadeDiag	,034	,029	1,343	1	,247	1,034	,977	1,094

Tabela 82 – Estimativas Modelo de Cox sem variáveis Heterossexual e Categoria Clínica

Retire-se a variável “IdadeEntCAT\_”, que apresenta neste momento o maior nível de não significância (p-value = 0,334). Calculando  $G = 205,143 - 204,203 = 0,94$ , e o p-value para o respectivo  $\chi^2_1$  é igual a 0,3322. Não se rejeita a hipótese nula, logo a variável pode ser retirada definitivamente do modelo sem que se afecte o grau de explicabilidade deste.

Nesta fase registaram-se alterações significativas no modelo (*vide* Tabela 83). De facto, as variáveis Idade no Diagnóstico e Depressão Nervosa passaram a ser significativas. Apenas a variável Consulta de Psicologia continua a não o ser.

Variables in the Equation								
	B	SE	Wald	df	Sig.	Exp(B)	95.0% CI for Exp(B)	
							Lower	Upper
CatCD4			16,581	2	,000			
CatCD4(1)	-1,907	1,038	3,378	1	,066	,149	,019	1,135
CatCD4(2)	-2,050	,541	14,342	1	,000	,129	,045	,372
Tubercul	1,242	,423	8,610	1	,003	3,463	1,511	7,941
DepNerv	-1,374	,764	3,235	1	,072	,253	,057	1,131
ConsEnf	-1,310	,474	7,625	1	,006	,270	,107	,684
ConsPsic	-,650	,656	,982	1	,322	,522	,144	1,888
IdadeDiag	,056	,017	10,310	1	,001	1,057	1,022	1,094

Tabela 83 – Estimativas Modelo de Cox sem variáveis Heterossexual, Categoria Clínica e Idade de Entrada categorizada

Retire-se a variável “ConsPsic” e avalie-se o seu impacto no modelo.

Esta variável não parece ser, de facto, significativa no modelo. O p-value obtido no Teste de Razão de Verossimilhanças não permite rejeitar a respectiva hipótese nula ( $G = 206,252 - 205,143 = 1,109$  tal que p-value para  $\chi_1^2$  é igual a 0,2923).

Deixando ficar no modelo apenas as variáveis efectivamente significativas (5 variáveis), obtém-se o exposto na Tabela 84.

Variables in the Equation								
	B	SE	Wald	df	Sig.	Exp(B)	95.0% CI for Exp(B)	
							Lower	Upper
CatCD4			16,083	2	,000			
CatCD4(1)	-1,890	1,038	3,315	1	,069	,151	,020	1,156
CatCD4(2)	-1,995	,535	13,901	1	,000	,136	,048	,388
Tubercul	1,334	,417	10,246	1	,001	3,795	1,677	8,587
DepNerv	-1,419	,759	3,496	1	,062	,242	,055	1,071
ConsEnf	-1,411	,465	9,195	1	,002	,244	,098	,607
IdadeDiag	,055	,017	10,251	1	,001	1,056	1,022	1,093

Tabela 84 – Estimativas Modelo de Cox sem variáveis Heterossexual, Categoria Clínica, Idade de Entrada categorizada e Consulta de Psicologia

### 5.3.2 – Inclusão variáveis não significativas na análise univariada

Na análise univariada foram excluídas algumas variáveis que quando colocadas individualmente a explicar a variável resposta Tempo de Sobrevivência se revelaram estatisticamente não significativas, de acordo com o nível de significância então considerado.

Dessas variáveis consideram-se agora as mais relevantes em termos do objectivo do estudo e cujos dados estejam o mais completo possível.

A Tabela 85 contém os resultados obtidos para o Teste Wald (ou Teste da Razão de Verossimilhanças se necessário) e respectivas considerações aquando da inclusão de cada uma das referidas variáveis no modelo apurado no ponto anterior, por etapa de inclusão. Cada etapa é nomeada assim que seja encontrada uma variável que melhore o modelo já existente, tal que a partir desse momento essa variável é incluída e cada uma de todas as restantes variáveis é comparada face a esse novo modelo.

Etapa de Inclusão	Variável	p-value (Teste Wald)	Valor Wald	Deviance	Observações
1; 2; 3	Sexo	0.669; 0.568; 0.723	0.183; 0.327; 0.125	206.064; 195.749; 188.246	
<b>1</b>	<b>IdadeEnt</b>	<b>0.004</b>	<b>8.282</b>	<b>194.881</b>	<b>G= 11,371, donde p-value para <math>\chi_1^2</math>.no Teste Razão Verossimilhanças igual a 0,00746. A variável é importante no modelo.</b>
1; 2; 3	Recluso	0.252; 0.35; 0.247	1.313; 0.875; 1.339	193.096; 190.846; 186.579	
1; 2; 3	Toxicodep	0.574; 0.443; 0.352	0.316; 0.590; 0.868	194.559; 191.350; 187.513	
1; 2; 3	HomoBissex	0.749; 0.941; 0.944	0.102; 0.006; 0.005	194.774; 191.952; 188.369	
1; 2; 3	EstadoAct	0.992; 0.957; 0.928	0.096; 0.318; 0.043	97.018; 96.036; 95.999	Nenhuma das suas categorias significativa.
1; 2; 3	TipoVirus	1.000; 0.999; 0.999	0.001; 0.000; 0.000	193.924; 190.636; 186.600	
1; 2; 3	TAR	<b>0.087</b> ; 0.16; 0.255	2.938; 1.973; 1.295	192.372; 190.223; 187.197	Etapa 1: G= 2,509, donde p-value para $\chi_1^2$ no Teste Razão Verossimilhança igual a 0,1132. A variável afinal não é importante no modelo.

1; 2; 3	Hep B	0.270; 0.314; 0.178	1.217; 1.014; 1.813	190.704; 190.603; 185.748	
<b>3</b>	<b>Hep C</b>	<b>0.173; 0.069</b>	<b>1.854; 3.310</b>	<b>190.404; 188.378</b>	<b>Etapa 2: G= 17,874, donde p-value para <math>\chi_1^2</math>.no Teste Razão Verosimilhanças igual a 0,00002. A variável é importante no modelo.</b>
<b>2</b>	<b>Álcool</b>	<b>0.068</b>	<b>3.330</b>	<b>191.958</b>	<b>G= 2,923, donde p-value para <math>\chi_1^2</math> no Teste Razão Verosimilhança igual a 0,0873. A variável é importante no modelo.</b>
2; 3	Tabag	0.550; 0.661	0.357; 0.192	191.636; 188.195	
2; 3	SectAct	0.724; 0.963	0.125; 0.002	191.804; 188.345	
2; 3	ConcRes_	0.297; 0.347	16.264; 15.467	159.557; 160.243	Nenhuma das suas categorias significativa.
2; 3	PaísOr	0.764; 0.803	3.349; 3.044	187.083; 184.086	Nenhuma das suas categorias significativa.
2; 3	PaísOrPort	0.778; 0.893	0.079; 0.018	191.873; 188.355	

Tabela 85 – Resultados nas estimativas Modelo de Cox com inclusão de variáveis por etapas

Da análise à Tabela 85 constata-se que na Etapa 1 detectou-se a variável “IdadeEnt” como significativa, tal que passou, a partir desse momento a fazer parte do modelo. Ainda nesta etapa, e após a inclusão da idade de entrada, a variável “TAR” parecia ser também significativa a 10%. No entanto o Teste da Razão de Verosimilhanças veio confirmar que a variável não é importante no modelo, tal que não foi incluída.

Na etapa seguinte, acrescentou-se a variável “Alcool”, e posteriormente, na Etapa 3 considerou-se a inclusão efectiva de “Hep C”.

Assim, o modelo resultante e em vigor até este ponto é o que se apresenta na Tabela 86. O seu valor Deviance é igual a 188,374.

Variables in the Equation								
	B	SE	Wald	df	Sig.	Exp(B)	90.0% CI for Exp(B)	
							Lower	Upper
CatCD4			17,671	2	,000			
CatCD4(1)	-2,364	1,072	4,865	1	,027	,094	,016	,548
CatCD4(2)	-2,199	,568	14,977	1	,000	,111	,044	,282
Tubercul	1,338	,435	9,459	1	,002	3,812	1,864	7,799
DepNerv	-1,323	,786	2,836	1	,092	,266	,073	,970
ConsEnf	-1,354	,471	8,259	1	,004	,258	,119	,560
IdadeDiag	,381	,098	15,079	1	,000	1,464	1,246	1,720
IdadeEnt	-,317	,096	10,924	1	,001	,728	,622	,853
Alcool	1,094	,519	4,450	1	,035	2,987	1,273	7,012
HepC	1,034	,568	3,310	1	,069	2,813	1,104	7,164

Tabela 86 – Estimativas Modelo de Cox após inclusão de variáveis por etapas

O novo modelo significativo passou a incluir 8 variáveis (6 categóricas e 2 contínuas), das quais 5 são significativas a 5% e as restantes a 10%.

## 5.4 – Análise da forma funcional das variáveis contínuas

Verifique-se agora se as variáveis contínuas entretanto apuradas para o modelo se encontram na escala correcta, isto é, lineares, ou se, pelo contrário, necessitam ser alteradas.

### 5.4.1 - Idade de Entrada – IdadeEnt

Depois de categorizada a variável Idade de Entrada em 4 categorias: [min;Q1 [, [Q1;Q2 [, [Q2;Q3 [, [Q3;Max [e incluída no modelo são registados os valores dos respectivos coeficientes, considerando a primeira categoria como referência. Estes são representados graficamente, cujas abcissas constituem os pontos médios de cada categoria.

Da observação ao gráfico obtido (Figura 51) detecta-se uma ligeira subida seguida de tendência decrescente. Tal pode invalidar a linearidade da variável.



Ainda assim, mesmo quando avaliada a mesma variável na escala quadrática (para a qual foi criada a variável  $\text{IdadeEnt}^2$ ), em nada se altera a conclusão anterior, e por outro lado, a inclusão da variável categorizada no modelo também não parece ser solução pois não se revela significativa.

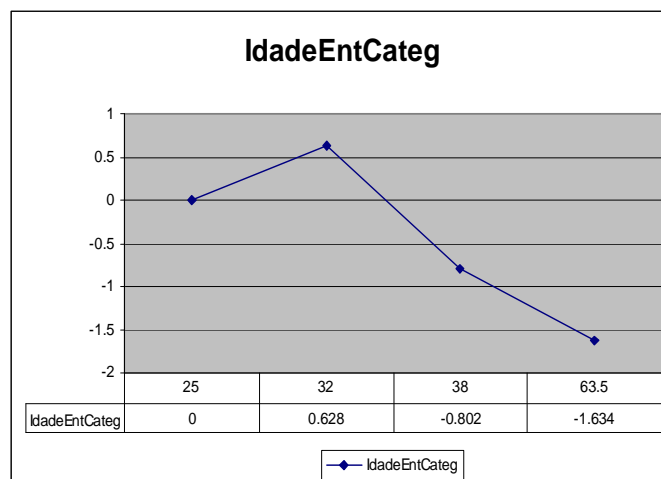


Figura 51 – Representação gráfica da variável Idade de Entrada por categorias

Para os devidos efeitos do estudo considera-se que a variável é aproximadamente linear.

#### 5.4.2 - Idade no Diagnóstico – IdadeDiag

O mesmo procedimento aplicado à variável anterior foi também aplicado à variável Idade no Diagnóstico. Depois de introduzida a respectiva nova variável categorizada apuraram-se os seus coeficientes, donde, representados com os pontos médios de cada uma das suas categorias (sendo a primeira a categoria de referência), resultou o gráfico da Figura 52, no qual se confirma a linearidade da variável. Por essa razão não se efectuarão quaisquer alterações na variável “IdadeDiag”.

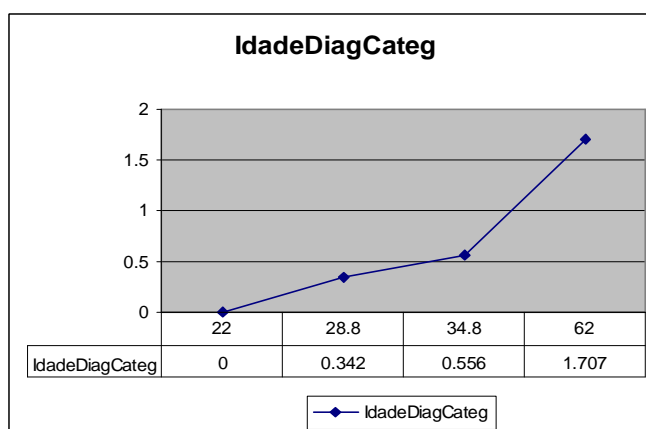


Figura 52 – Representação gráfica da variável Idade de Diagnóstico por categorias

Das análises anteriores conclui-se que ambas as variáveis contínuas do modelo significativo até aqui apurado não necessitam ser alteradas, pelo que o modelo permanece inalterado.

## 5.5 – Avaliação de interacções

Para averiguar possíveis melhoramentos no modelo actual, considera-se a introdução de algumas interacções entre as variáveis entretanto apuradas.

A Tabela 87 resume os resultados em termos de p-value obtido no Teste Wald e valor da Deviance (-2LL) para cada uma das interacções introduzidas.

Interacções	Efeitos no modelo		Interacções (continuação)	Efeitos no modelo	
	p-value Teste Wald	-2LL		p-value Teste Wald	-2LL
CatCD4*Tubercul	0.701	186.178	Hep C*Alcool	0.823	188.325
CatCD4*Hep C	0.298	185.74	Hep C*ConsEnf	0.484	187.88
CatCD4*DepNerv	1	187.979	Hep C*IdadeEnt	0.661	188.183
CatCD4*Alcool	0.463	185.423	Hep C*IdadeDiag	0.59	188.085
CatCD4*ConsEnf	0.745	187.711	DepNerv*Alcool	0.953	188.357
CatCD4*IdadeEnt	0.933	192.098	DepNerv*ConsEnf	0.716	188.242
CatCD4*IdadeDiag	0.903	188.184	DepNerv*IdadeEnt	0.239	187.043
<b>Tubercul*Hep C</b>	<b>0.127</b>	<b>186.069</b>	DepNerv*IdadeDiag	0.296	187.325
Tubercul*DepNerv	0.842	188.334	Alcool*ConsEnf	0.782	188.298
Tubercul*Alcool	0.259	187.121	Alcool*IdadeEnt	0.397	187.704
Tubercul*ConsEnf	0.91	188.362	Alcool*IdadeDiag	0.633	188.154
Tubercul*IdadeEnt	0.425	187.771	ConsEnf*IdadeEnt	0.882	188.352
Tubercul*IdadeDiag	0.412	187.741	ConsEnf*IdadeDiag	0.633	188.154
Hep C*DepNerv	0.98	186.939	IdadeEnt*IdadeDiag	0.777	188.297

Tabela 87 – Resultados no Teste Wald e Razão de Verossimilhanças no estudo de interacções

Como se observa nenhuma das interacções consideradas se revelou significativa. Acerca da interacção Tuberculose\*Hepatite C, cujo p-value no Teste Wald resultou um pouco acima do nível de significância considerado, calculou-se a diferença entre as deviances do modelo reduzido e do modelo com a interacção ( $G= 2,305$ ), a partir da qual se calculou o p-value para

$\chi_1^2$  e constatou-se que de facto esta nova variável explanatória não era importante no modelo (p-value = 0,11289). Assim, o modelo em vigor é o apurado no ponto anterior, ou seja, sem quaisquer interacções.

## **5.6 – Validação de pressupostos e verificação da adequabilidade do modelo**

### **5.6.1 – Proporcionalidade de riscos**

O pressuposto inerente ao modelo de Regressão de Cox consiste em assegurar que a taxa de risco ou função hazard não depende de t (tempo), tal que a influência que as covariáveis exercem sobre o risco não sofre qualquer alteração ao longo do tempo. Se isso ocorresse, invalidaria toda análise e interpretação do risco inerente a cada uma das covariáveis, mesmo mantendo tudo o resto constante.

Existem várias formas de confirmar este pressuposto da proporcionalidade na taxa de risco em função do tempo de vida. Para as variáveis categóricas pode fazer-se via observação gráfica, e no caso das variáveis contínuas recorre-se ao Teste de Rho, cuja hipótese nula defende que o pressuposto é válido.

Como observável nos gráficos da Figura 53 as representações indicam proporcionalidade de riscos para todas as variáveis categóricas consideradas no modelo (as curvas log-log são aproximadamente paralelas para as categorias). Por conseguinte considera-se validado o pressuposto em análise para as variáveis do tipo categórico.

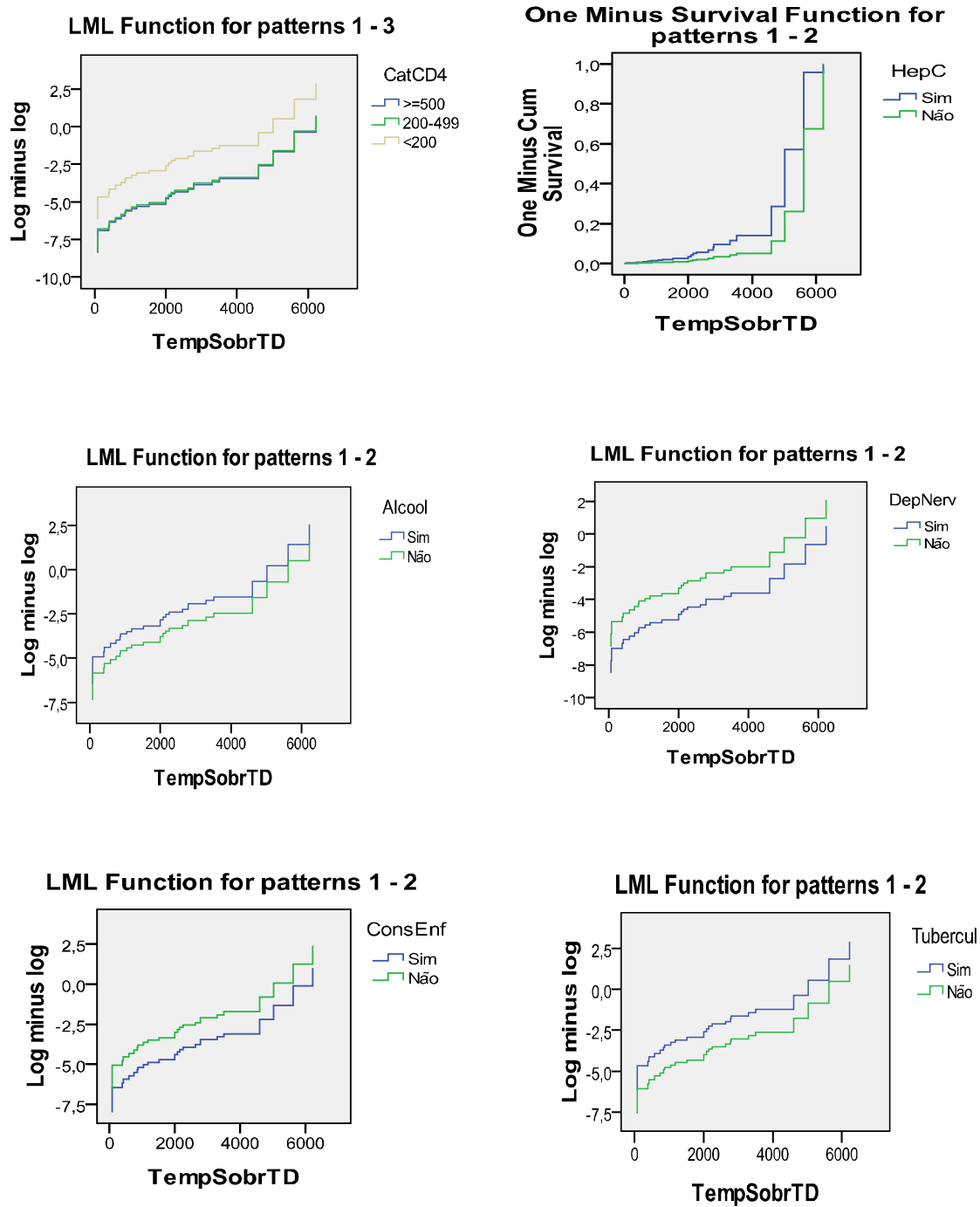


Figura 53 – Validação gráfica do pressuposto da proporcionalidade de riscos para as covariáveis do tipo categórico

	Rho	Chisq	p
<b>CatCD4(1)</b>	0.2065	0.9633	<b>0.3264</b>
<b>CatCD4 (2)</b>	0.4781	4.1105	<b>0.0426</b>
<b>Tubercul</b>	0.0451	0.0692	<b>0.7926</b>
<b>DepNerv</b>	-0.0978	0.2349	<b>0.6279</b>
<b>ConsEnf</b>	0.2256	1.2435	<b>0.2648</b>
<b>IdadeDiag</b>	-0.2670	1.4135	<b>0.2345</b>
<b>IdadeEnt</b>	0.2627	1.3942	<b>0.2377</b>
<b>Alcool</b>	-0.0259	0.0179	<b>0.8935</b>
<b>Hep C</b>	-0.1325	0.5213	<b>0.4703</b>
<b>Global</b>	---	<b>7.2550</b>	<b>0.6106</b>

Tabela 88 – Resultados Teste Rho

De facto, de acordo com os resultados obtidos no Teste de Rho (*vide* Tabela 88), cuja  $H_0$  defende a proporcionalidade de riscos, pode afirmar-se que o pressuposto é válido para o modelo avaliado e respectivas variáveis explanatórias.

Pode por isso afirmar-se que o risco se mantém ao longo de todo o período de acompanhamento, uma vez que não depende do tempo  $t$ .

### 5.6.2 – Análise de Resíduos e Adequabilidade do modelo

A análise de resíduos do modelo ajustado visa apurar se existem observações que possam estar a influenciar significativamente a variância do erro, no caso de se confirmar a existência de *outliers*. E por outro lado verificar também se existem observações consideradas influentes que estejam a condicionar a própria interpretação do risco, tomando dessa forma o modelo mais fraco e portanto menos útil.

Numa análise preliminar, cujos resultados se encontram em anexo (*vide* Anexo 2) detectaram-se determinadas observações cujo impacto no modelo, aquando da sua eliminação da base de dados revelou tratarem-se efectivamente de observações influentes. Tome-se como exemplo o caso da alteração no coeficiente da covariável CatCD4(1) após a eliminação da observação 120. Esta provocou uma alteração de cerca de 500%! Depois de analisar esse indivíduo constatou-se que se trata de um indivíduo falecido, razão pela qual não existe interesse maior em eliminá-lo, e o seu impacto poderia estar relacionado com a forma como as categorias das covariáveis em questão estavam definidas. Perante isso, optou-se por tentar recategorizar a covariável CatCD4. De facto,

a observação 120 correspondia a um indivíduo cuja contagem de células CD4 havia resultado superior a 500 e no entanto o seu estado actual é falecido.

Após várias tentativas conseguiu-se obter um modelo significativo e sem necessidade de eliminar algumas das observações anteriormente detectadas como influentes. A covariável CatCD4 passou a ter apenas duas categorias: 1 -  $\geq 200$  células e 2 -  $< 200$  células (esta última constitui a categoria de referência). O modelo em vigor (cuja Deviance = 188,396) é o que se apresenta na Tabelas 89.

Variables in the Equation								
	B	SE	Wald	df	Sig.	Exp(B)	90.0% CI for Exp(B)	
							Lower	Upper
CatCD4200	-2,230	,530	17,685	1	,000	,108	,045	,257
Tubercul	1,336	,435	9,441	1	,002	3,802	1,860	7,772
HepC	1,026	,566	3,285	1	,070	2,789	1,099	7,077
DepNerv	-1,325	,786	2,839	1	,092	,266	,073	,969
Alcool	1,090	,518	4,419	1	,036	2,973	1,267	6,973
ConsEnf	-1,353	,472	8,225	1	,004	,258	,119	,562
IdadeEnt	-,318	,096	10,922	1	,001	,728	,621	,852
IdadeDiag	,382	,098	15,067	1	,000	1,465	1,246	1,722

Tabela 89 – Estimativas Modelo de Cox

No que respeita ao pressuposto da proporcionalidade de riscos este mantém-se válido de acordo com os resultados obtidos no Teste de Rho (Tabela 90).

	Rho	Chisq	p
<b>CatCD4</b>	0.4962	4.7063	<b>0.0301</b>
<b>Tubercul</b>	0.0448	0.0690	<b>0.7927</b>
<b>DepNerv</b>	-0.0983	0.2376	<b>0.6260</b>
<b>ConsEnf</b>	0.2268	1.2592	<b>0.2618</b>
<b>IdadeDiag</b>	-0.2644	1.4246	<b>0.2327</b>
<b>IdadeEnt</b>	0.2596	1.4036	<b>0.2361</b>
<b>Alcool</b>	-0.0269	0.0197	<b>0.8883</b>
<b>Hep C</b>	-0.1362	0.5673	<b>0.4513</b>
<b>Global</b>	---	<b>7.4311</b>	<b>0.4909</b>

Tabela 90 – Resultados Teste Rho

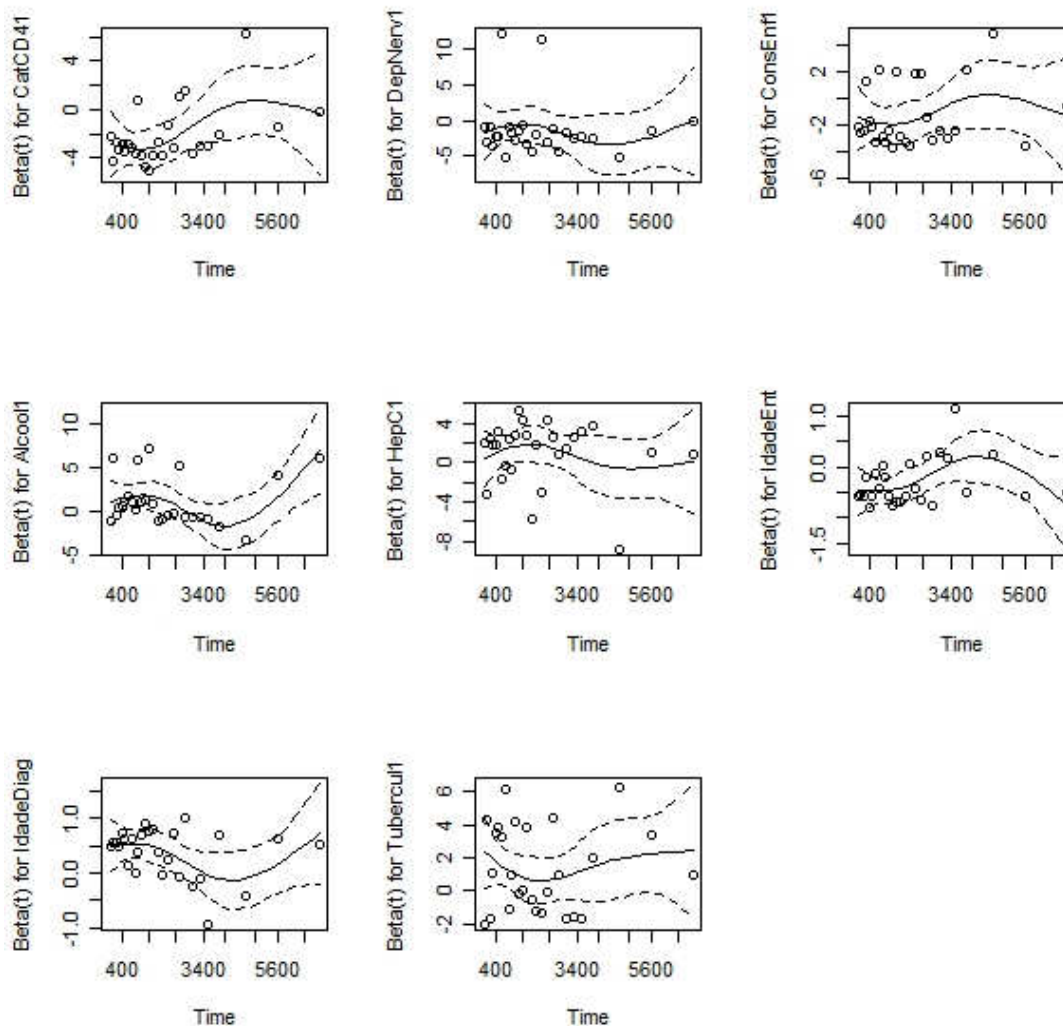
Resíduos de Schoenfeld

Figura 54 – Representação gráfica dos resíduos de Schoenfeld

Da observação aos gráficos da figura 54 não se detectam grandes diferenças na linha *lowess*, tal que não parece existir tendência específica para nenhuma das covariáveis em estudo. Além disso, pode também afirmar-se que, no geral, os valores se distribuem aleatoriamente em torno de 0, o que indicia a validade do pressuposto referido no ponto anterior. Como tal, não se considera que as covariáveis dependam do tempo de sobrevivência e os riscos são aproximadamente proporcionais.

### Resíduos Martingala e Resíduos Deviance

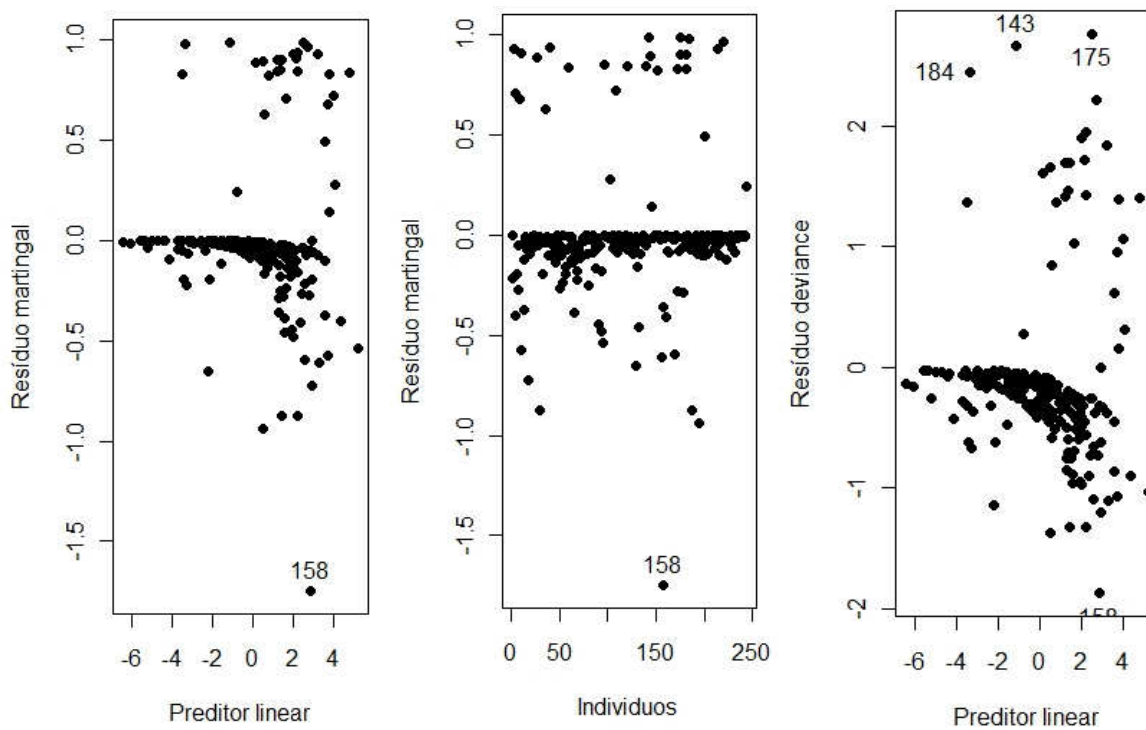


Figura 55 – Representação gráfica Resíduos Martingale e Deviance

Da representação gráfica dos resíduos *martingale* contra o preditor linear verifica-se que as observações estão distribuídas de forma aleatória (pois não se detecta qualquer padrão na sua distribuição), evidenciando apenas um candidato a *outlier* (a observação 158). Da análise ao gráfico dos resíduos *martingale* contra os indivíduos mais uma vez se destaca a mesma observação como possível *outlier*.

No gráfico dos resíduos deviance versus preditor linear, detectam-se algumas observações afastadas (143, 158, 175 e 184) constituindo estas também candidatos a *outliers*. Será analisado o seu impacto em termos das alterações na deviance.



### DfBetas

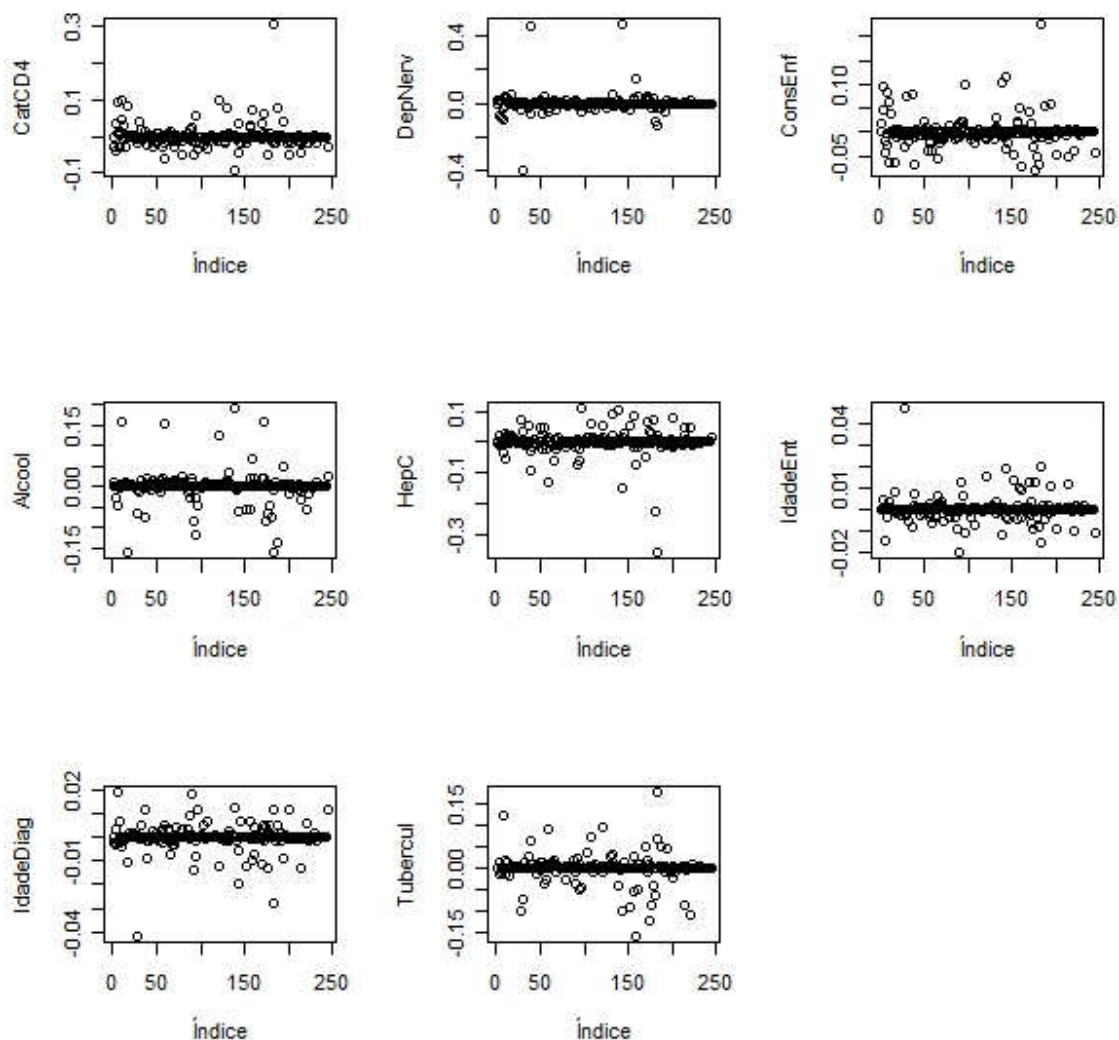


Figura 56 – Representação gráfica DfBetas

Da observação aos gráficos da Figura 56 verifica-se que existem alguns pontos que se destacam dos restantes. Tais pontos constituem possíveis observações influentes, e nesse caso podem estar a distorcer a realidade que o modelo visa representar uma vez que conduzem a uma incorrecta interpretação do risco de morte.

Por covariável detectam-se as seguintes observações possivelmente influentes:

- CatCD4: observação 184;
- DepNerv: observações 30, 40 e 143;

- ConsEnf: observação 184;
- HepC: observações 181 e 184;
- IdadeDiag e IdadeEnt: observação 27 (com ligeira influência);
- Tubercul: observações 158 e 184.

De facto, pela análise do gráfico que representa a medida de influência total (Figura 57) confirmam-se como possíveis observações influentes as observações 30, 40, 120, 143 e 181.

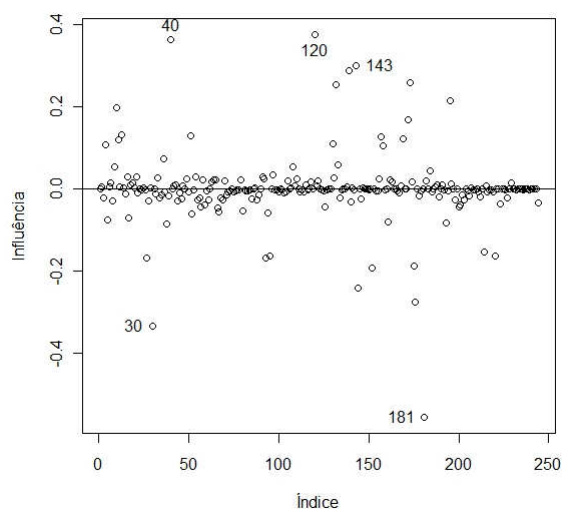


Figura 57 – Representação gráfica da medida de influência total

Todas as observações possivelmente influentes bem como candidatos a outliers devem ser retiradas do modelo de forma a apurar o seu impacto concreto em termos de variação nos valores dos coeficientes, na significância das variáveis e ainda na *deviance*.

Refira-se que, apenas a observação 30 e a 158 correspondem a dados censurados, tal que as restantes dizem respeito a indivíduos para os quais ocorreu o evento de interesse. Esta informação pode ser relevante aquando da decisão de eliminar permanentemente determinadas observações da base de dados actual.

De acordo com os resultados obtidos na análise sobre o impacto das observações detectadas, verifica-se que os indivíduos com a entrada 120, 158 e 175 não parecem estar a influenciar o modelo. De facto, quando retiradas, individualmente, do modelo não se registam quaisquer alterações significativas face ao modelo com todas as observações (*vide* Tabela 91).

Covariáveis	Modelo inicial		sem obs. 120			sem obs. 158			sem obs. 175		
	Coef.	sig.	Coef.	Δ% face ao inicial	sig.	Coef.	Δ% face ao inicial	sig.	Coef.	Δ% face ao inicial	sig.
CatCD4_200	-2.230	0.000	-2.347	5%	0.000	-2.335	5%	0.000	-2.235	0%	0.000
Tubercul	1.336	0.002	1.228	-8%	0.006	1.563	17%	0.000	1.466	10%	0.001
HepC	1.026	0.070	0.971	-5%	0.089	1.139	11%	0.043	0.992	-3%	0.090
DepNerv	-1.325	0.092	-1.328	0%	0.091	-1.563	18%	0.055	-1.350	2%	0.090
Alcool	1.090	0.036	0.943	-13%	0.086	0.994	-9%	0.055	1.181	8%	0.025
ConsEnf	-1.353	0.004	-1.354	0%	0.005	-1.429	6%	0.002	-1.325	-2%	0.005
IdadeEnt	-0.318	0.001	-0.337	6%	0.001	-0.335	5%	0.000	-0.308	-3%	0.001
IdadeDiag	0.382	0.000	0.398	4%	0.000	0.399	4%	0.000	0.378	-1%	0.000
<b>-2LL</b>	188.396		180.441			183.998			178.570		

Tabela 91 – Alterações verificadas no modelo após exclusão de observações 120, 158 e 175

No entanto, para as restantes observações (27, 30, 40, 143, 181 e 184) existem alterações consideráveis ao nível dos valores dos coeficientes das covariáveis e sua significância no modelo. De facto, quando retiradas da base de dados, quatro daquelas observações conduzem à exclusão da covariável DepNerv (Tabela 92), uma vez que o valor do p-value obtido não permite rejeitar a respectiva  $H_0$  no Teste Wald ( $\beta = 0$ ).

Covariáveis	sem obs. 27			sem obs. 30			sem obs. 181			sem obs. 184		
	Coef.	Δ% face ao inicial	sig.	Coef.	Δ% face ao inicial	sig.	Coef.	Δ% face ao inicial	sig.	Coef.	Δ% face ao inicial	sig.
CatCD4_200	-2.206	-1%	0.000	-2.317	4%	0.000	-2.219	0%	0.000	-2.601	17%	0.000
Tubercul	1.460	9%	0.001	1.496	12%	0.001	1.410	6%	0.001	1.128	-16%	0.013
HepC	0.944	-8%	0.092	0.949	-8%	0.088	1.289	<b>26%</b>	0.035	1.433	<b>40%</b>	0.017
DepNerv	-1.289	-3%	<b>0.100</b>	-0.657	<b>-50%</b>	<b>0.405</b>	-1.204	-9%	<b>0.127</b>	-1.180	-11%	<b>0.133</b>
Alcool	1.160	6%	0.027	1.125	3%	0.031	1.175	8%	0.026	1.261	16%	0.017
ConsEnf	-1.327	-2%	0.005	-1.517	12%	0.002	-1.279	-5%	0.007	-1.629	20%	0.001
IdadeEnt	-0.379	19%	0.001	-0.313	-2%	0.001	-0.309	-3%	0.001	-0.347	9%	0.001
IdadeDiag	0.438	15%	0.000	0.383	0%	0.000	0.381	0%	0.000	0.420	10%	0.000
<b>-2LL</b>	180.097			185.721			179.551			178.435		

Tabela 92 – Alterações verificadas no modelo após exclusão de observações 27,30,181 e 184

Além disso, acresce ainda o impacto verificado ao nível do valor do coeficiente da covariável DepNerv quando retiradas as observações 30, 40 e 143. Nestes casos, a alteração registada é de 50% (Tabela 93).

Covariáveis	sem obs. 40			sem obs. 143		
	Coef.	Δ% face ao inicial	sig.	Coef.	Δ% face ao inicial	sig.
CatCD4_200	-2.207	-1%	0,000	-2.271	2%	0,000
Tubercul	1.263	-5%	0,005	1.442	8%	0,001
HepC	1.128	10%	0,054	1.188	16%	0,052
DepNerv	-1.989	<b>50%</b>	0,060	-1.990	<b>50%</b>	0,062
Alcool	1.070	-2%	0,038	1.158	6%	0,028
ConsEnf	-1.279	-5%	0,008	-1.487	10%	0,003
IdadeEnt	-0.326	3%	0,001	-0.314	-1%	0,001
IdadeDiag	0.391	2%	0,000	0.387	1%	0,000
			178.858	174.722		

Tabela 93 – Alterações verificadas no modelo após exclusão de observações 40 e 143

Mediante o acima exposto opta-se por excluir do modelo a covariável DepNerv. Desta forma, para além de ter em consideração os vários indícios que apontam na não explicabilidade efectiva desta covariável, permite-se, por outro lado, que não seja necessário eliminar quaisquer das observações referidas, uma vez que o seu impacto incidia sobre a covariável excluída (os resultados sobre o impacto das observações 30, 40, 143, 181 e 184 sobre o novo modelo constam da Tabela (Anexo 3).

Na sequência da análise de resíduos efectuada resulta um novo modelo de sobrevivência para a explicação do evento em estudo. Este é o **modelo final** (*vide* Tabela 94 e Figura 58), com Deviance igual a 192,143, que explica o tempo de sobrevivência dos indivíduos com VIH/ SIDA acompanhados no Hospital de Beja.

	Variables in the Equation							90.0% CI for Exp(B)	
	B	SE	Wald	df	Sig.	Exp(B)	Lower	Upper	
CatCD4200	-2,216	,534	17,194	1	,000	,109	,045	,263	
Tubercul	1,197	,428	7,816	1	,005	3,309	1,637	6,692	
HepC	1,273	,553	5,306	1	,021	3,571	1,439	8,862	
Alcool	1,263	,522	5,849	1	,016	3,535	1,498	8,345	
ConsEnf	-1,443	,470	9,432	1	,002	,236	,109	,512	
IdadeEnt	-,300	,095	10,048	1	,002	,741	,634	,866	
IdadeDiag	,365	,096	14,288	1	,000	1,440	1,229	1,688	

Tabela 94 – Estimativas Modelo de Cox – Modelo Final

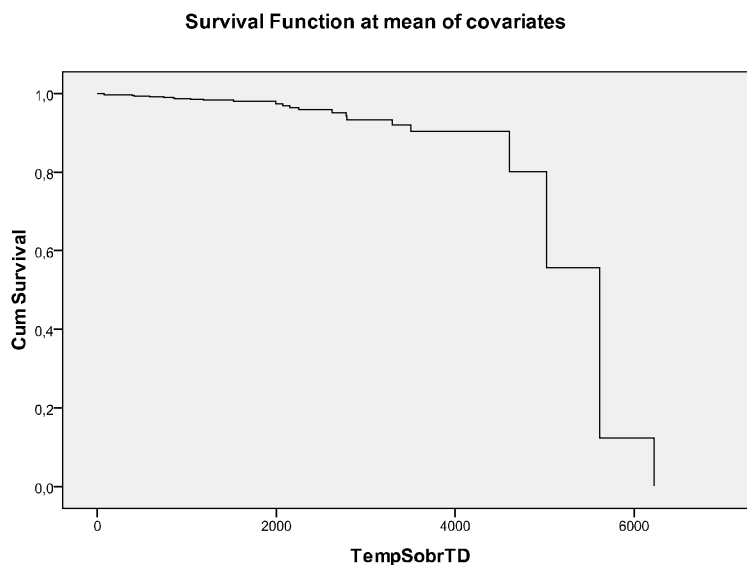


Figura 58 – Representação gráfica da função de sobrevivência estimada pelo Modelo de Cox

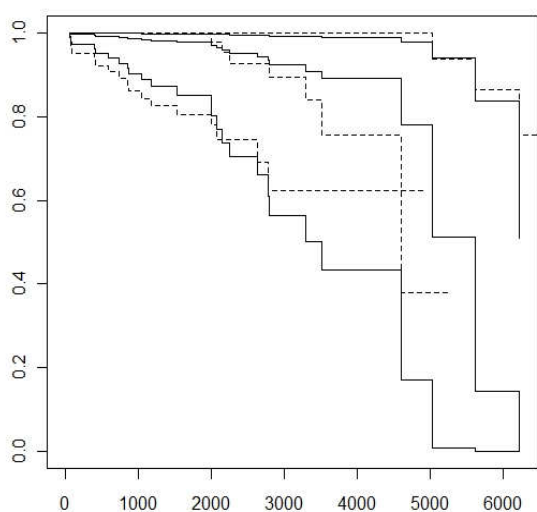


Figura 59 – Representação gráfica do Índice de Prognóstico

No que respeita ao comportamento do modelo, pela observação do gráfico que representa o Índice de Prognóstico na Figura 59 (onde a linha a cheio representa o modelo ajustado e a linha em tracejado representa a estimativa de Kaplan-Meier), admite-se que o modelo tem um bom ajustamento aos dados, uma vez que as curvas ajustadas pelo modelo são próximas das curvas estimadas por Kaplan-Meier.

## 5.7 – Análise paramétrica

Na análise a um possível modelo paramétrico obtiveram-se os seguintes gráficos (Figura 60), cuja observação permite verificar que as distribuições que melhor parecem adequar-se aos dados, tendo em conta a sua distribuição em torno da curva de Kaplan-Meier, são Weibull e Exponencial.

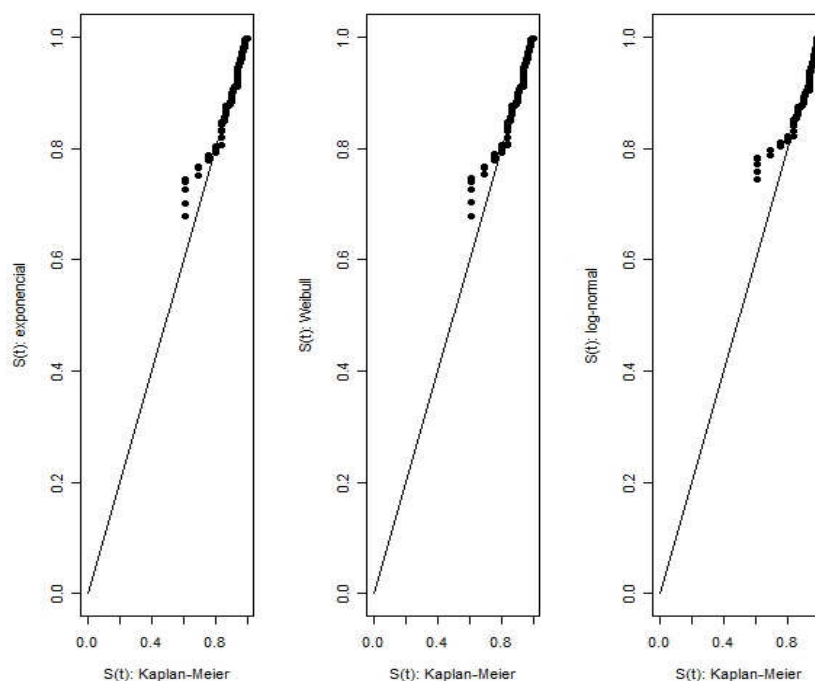


Figura 60 – Representação gráfica do ajustamento aos dados via Modelo Exponencial, Weibull e Log-normal

Sendo a distribuição Exponencial um caso particular da Weibull, calculou-se o Teste da Razão de Verossimilhanças no sentido de apurar se o modelo exponencial apresenta de facto um ajustamento igual ao modelo weibull.

Sendo  $-2LL_{\text{exp}} = 592,1228$  e  $-2LL_{\text{weibull}} = 592,1226$ , o p-value obtido para  $\chi_1^2(0,0000265)$  é igual a 0,9869, donde não se rejeita a hipótese nula ( $H_0: \sigma = 1$ ) de que o modelo exponencial se adequa tão bem aos dados quanto o modelo esperado weibull.

A função de sobrevivência pelo modelo exponencial seria:

$$S(t) = \exp\left\{-\left(\frac{t}{\exp(11.04 + 1.849CatCD4200 - 1.153Tuberc - 0.886HepC - 1.048Alcool + 1.085ConsEnf + 0.115IdadeEnt - 0.162IdadeDiag)}\right)\right\}$$

onde todas as covariáveis resultaram significativas pelo resultado obtido no Teste Wald.

A efectiva adequabilidade do modelo pode ser analisada através dos Resíduos de Cox-Snell. Se o modelo for bem ajustado então os resíduos seguem uma distribuição exponencial padrão ( $\lambda = 1$ ). Da observação à representação gráfica dos resíduos verifica-se que, ainda que o modelo exponencial tenha sido, entre os modelos testados, o que melhor se adequa aos dados em estudo, o seu ajustamento não é óptimo, pois existe um afastamento considerável entre o modelo esperado e Kaplan-Meier (Figura 61).

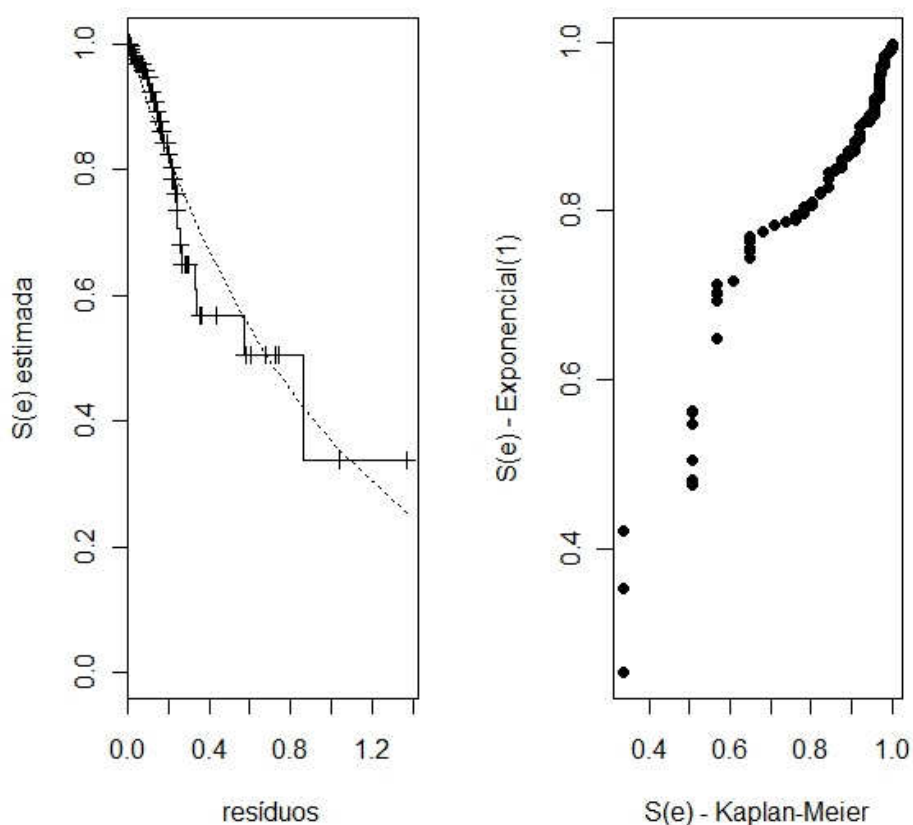


Figura 61 – Representação Gráfica dos resíduos

Pelo exposto se considera que a melhor opção, desconhecendo a verdadeira distribuição que os dados em estudo seguem, é o modelo semi-paramétrico de Cox entretanto já calculado e devidamente validado.

## 5.8 – Interpretação do modelo final

Decorrido o processo de selecção das variáveis que influenciam de forma estatisticamente significativa o tempo de sobrevivência dos indivíduos em estudo, e validados os respectivos pressupostos do modelo de Cox calculado, bem como sua adequabilidade aos dados, tem-se que a função de risco resultante é dada pela expressão:

$$h(t) = h_0(t) \exp (-2.216 \text{ CatCD4200} + 1.197 \text{ Tubercul} + 1.273 \text{ HepC} + 1.263 \text{ Alcool} - 1.443 \text{ ConsEnf} - 0.300 \text{ IdadeEnt} + 0.365 \text{ IdadeDiag})$$

O tempo que decorre entre o diagnóstico da doença VIH/ SIDA e o momento de morte dos indivíduos infectados é principalmente influenciado pela idade com que o indivíduo é diagnosticado, a idade com a qual inicia o acompanhamento hospitalar e a categoria de células CD4 em que o indivíduo se insere pelo seu valor mais baixo obtido até ao momento actual. Além disso, também as restantes patologias e/ou comportamentos de adicção verificados nestes indivíduos parecem determinar a sua longevidade. Se para além da infecção VIH/ SIDA coexistirem as infecções Tuberculose e Hepatite C (conjuntamente ou não ao VIH), bem como o facto do indivíduo apresentar comportamento alcoólico, o seu tempo de sobrevivência é bastante menor quando comparados com indivíduos que não apresentem esses factores.

Uma variável susceptível de análise mais detalhada por parte da equipa multidisciplinar que está actualmente responsável pelo seguimento dos doentes é a frequência destes à Consulta de Enfermagem. Este parece ser um factor que potencia um maior tempo de sobrevivência para os indivíduos que frequentam a referida consulta. Isto poderá estar relacionado com o tipo de abordagem que é feita nesse período de consulta na relação entre enfermeiro e doente e nesse aspecto muito depende da equipa específica de enfermeiros em função pois nada assegura que com quaisquer outros profissionais naquele lugar os resultados seriam os mesmos. Por essa razão esta variável é apenas válida se tudo o que caracteriza a actual Consulta de Enfermagem se mantiver.



Da interpretação à razão de riscos calculada para um nível de significância de 10% para cada uma das covariáveis que constituem o modelo explicativo do tempo de sobrevivência em estudo, considerando nas variáveis do tipo categórico a última categoria como referência, pode afirmar-se que mantendo tudo o resto constante:

- I. Por cada ano a mais com que o indivíduo é diagnosticado com a infecção VIH/SIDA, o seu risco de morte aumenta entre 23% a 69%. A idade de diagnóstico é por isso um factor de risco;
- II. Quanto mais idade o indivíduo tiver aquando da entrada no acompanhamento hospitalar, tanto maior será o seu tempo de sobrevivência esperado. De facto, por cada ano a mais, o risco de morte associado à infecção VIH/SIDA diminui em 26%;
- III. Para um indivíduo cujo valor mais baixo de células CD4 registado esteja acima das 200 células por microlitro de sangue, o seu risco de morte, face a um indivíduo na situação contrária, decresce em cerca de 90%. Esta razão de riscos pode variar entre os 74% e os 96%. Outra interpretação do valor obtido, invertendo a categoria de referência, permite afirmar que o risco de morte aumenta cerca de 9 vezes se o número de células CD4 se situar abaixo das 200 células por microlitro de sangue;
- IV. Se o indivíduo estiver também infectado com Tuberculose, o seu risco de morte é cerca de 3 vezes superior face à situação oposta, de não apresentar essa patologia. O intervalo de confiança associado a esta razão de riscos indica que aquele risco pode resultar até 7 vezes superior num indivíduo tuberculoso;
- V. No caso de ao indivíduo ter sido diagnosticada Hepatite C, o seu risco de morte é 3,6 vezes superior quando comparado com um indivíduo sem essa infecção;
- VI. Se o indivíduo apresentar comportamento alcoólico o seu risco de morte pode ser entre 1,5 a 8 vezes superior face a um indivíduo que não apresente essa adicção;
- VII. A frequência da consulta de enfermagem reduz o risco de morte nos indivíduos acompanhados em cerca de 76%. Este valor pode situar-se entre os 48% e os 90%. Se consideramos como categoria de referência o ir à consulta então pode afirmar-se que o risco de morte aumenta cerca de 4 vezes para os indivíduos que não frequentam a referida consulta.

## 6. Conclusão

O impacto da infecção por VIH/SIDA na sociedade em geral não se limitou apenas aos campos científico e médico, no sentido de desvendar a origem do vírus e consequentemente a forma de controlá-lo ou, no futuro, encontrar a respectiva cura ou vacina. Mesmo assim, têm sido evidentes os enormes avanços em termos de terapêutica, de tal forma que a SIDA passou a ser, em poucos anos, uma doença crónica. No entanto, é também perceptível que o alarme inicial associado ao elevado grau de transmissibilidade do vírus provocou profundas alterações a nível social e cultural na sociedade. A forma como hoje se abordam temas como a sexualidade, a orientação sexual, as tradições religiosas, os valores de família, as assimetrias económico-sociais e o preconceito social são muito diferentes do que há 30 anos atrás, e a infecção por VIH/SIDA teve um importante papel nessa mudança (Faria, 2010). Se os fins justificam ou não os meios ninguém o saberá nesta situação, mas face à realidade da doença o mundo não pode evitar de encará-la e a melhor forma que temos para combatê-la é aceitá-la, e estudá-la. O estudo do vírus e respectivos indivíduos infectados permitirá desenvolver melhores estratégias de prevenção e tratamento da doença.

No presente trabalho de investigação foram recolhidos os dados regionais e estudados os indivíduos infectados e seguidos, via internamento ou em regime de ambulatório, na Unidade Local de Saúde do Baixo Alentejo, ao nível hospitalar, de forma a caracterizar a população infectada da região e simultaneamente averiguar que factores estão a influenciar, e em que medida, o seu tempo de sobrevivência de entre as variáveis para as quais foi possível recolher informação.

A referir que o processo de recolha dos dados foi particularmente moroso e minucioso uma vez que, após a identificação dos indivíduos que vieram a constituir a amostra, as respectivas características foram recolhidas indivíduo a indivíduo para a maior parte das variáveis estudadas. Esta dificuldade na obtenção e validação dos dados pode ser, muitas vezes, responsável pela interrupção ou desistência em determinados objectivos de investigação. No caso isso não aconteceu porque houve apoio por parte da equipa de profissionais da instituição e por outro lado, motivação para efectivamente prosseguir com um estudo inédito sobre os doentes na região.

Uma das possíveis limitações da base de dados utilizada foi o reduzido número de eventos de interesse na análise de sobrevivência (número de mortes) verificado nos indivíduos em estudo. Espera-se que esse não seja um factor de menor potência nos resultados encontrados.

Da análise descritiva aos dados verificou-se que a população infectada e seguida na região do Baixo Alentejo em muito se assemelha à população infectada a nível nacional, na sua distribuição por sexo e idades, tipo de vírus, estadio da doença, comportamentos de risco e patologias apresentadas. Este pode ser um indício de que os resultados e conclusões obtidos poderiam inferir acerca da realidade nacional, e não apenas regional como inicialmente definido. No entanto, tem-se também a consciência de que nem todos os indivíduos potencialmente infectados da região estão diagnosticados. E aqui, assim como nos restantes estudos publicados acerca desta temática, coloca-se a questão sobre a efectiva representatividade da amostra.

Por constituírem uma população dita evasiva é muito difícil aferir o número total exacto de infectados (que incluem os infectados diagnosticados, notificados e não notificados, e os infectados não diagnosticados). O facto de não conhecermos o total geral de indivíduos infectados induz a uma possível subestimação no cálculo das taxas de incidência e prevalência da doença porque o valor comumente utilizado considera os indivíduos infectados e diagnosticados (pois esta é efectivamente a informação disponível).

Uma possível solução a esta questão seria recorrer a métodos estatísticos de estimação da abundância populacional tais como o método de captura-recaptura, indicado nos casos de populações evasivas sobre as quais existe informação escassa ou incompleta. Mas também neste caso há que ter especial cuidado na verificância dos respectivos pressupostos e consequente validade dos resultados obtidos. Este pode constituir um futuro rumo de investigação e continuação do trabalho entretanto desenvolvido.

Na análise de sobrevivência foram aplicados os respectivos métodos de forma a modelar o tempo de sobrevivência dos indivíduos em estudo desde o momento em que são diagnosticados com a doença até à ocorrência do evento morte. O apuramento das variáveis estatisticamente significativas foi realizado por fases, nomeadamente começou-se por uma análise univariada, através de estimativas de Kaplan-Meier e modelo de regressão de Cox, seguida da análise multivariada através do referido modelo semi-paramétrico. De facto, face aos resultados obtidos

na adequabilidade dos modelos paramétricos que se tentou ajustar aos dados, verificou-se que, não conhecendo a verdadeira distribuição destes, a melhor opção na explicação do tempo de sobrevivência dos doentes em estudo é mesmo o modelo de Cox.

Após a validação dos pressupostos inerentes ao modelo considerado e devidos ajustes aquando da análise de resíduos, obteve-se um modelo final com as variáveis que estão a determinar de forma estatisticamente significativa o tempo de sobrevivência dos indivíduos. São elas a categoria de células CD4 resultante da contagem pelo valor mais baixo que o indivíduo obteve até à data (o risco de morte decresce cerca de 90% se a contagem resultar acima das 200 células por microlitro de sangue), a idade com que foi diagnosticado com a doença (neste caso, quanto mais idade tiver maior o risco de morte), a idade com que iniciou o acompanhamento médico (sendo que quanto mais velho o indivíduo começar a ser seguido menor o seu risco de morte – cerca de 26% por cada ano a mais). Além disso, no que respeita às co-infecções e outras patologias apresentadas pelo doente, verificou-se que os indivíduos tuberculosos têm um risco de morte 3 vezes superior aos que não apresentam essa infecção, e, os indivíduos com Hepatite C têm um risco de morte aproximadamente 4 vezes superior face aos que não apresentam essa doença. Estas conclusões são válidas considerando cada uma das variáveis isoladamente e mantendo tudo o resto constante, mas é óbvio que se um indivíduo infectado com VIH/SIDA estiver co-infectado com Tuberculose e Hepatite C, o seu risco de morte é muito superior ao verificado num indivíduo sem essas co-infecções.

Por outro lado, verificou-se também que indivíduos cujo comportamento revele alcoolismo apresentam um risco de morte que pode ser até 8 vezes superior face a outro que não seja alcoólico.

A frequência da consulta de enfermagem, que consiste numa consulta de apoio e informação aos doentes, bem como acompanhamento dos doentes em tratamento, revelou estar a ser, nos indivíduos seguidos na instituição em estudo, um factor protector uma vez que parece reduzir o seu risco de morte em cerca de 76%. Este não poderia ser um resultado aplicado a outra realidade diferente da consulta e equipa específica do Hospital de Beja, uma vez que a única explicação para esta variável se ter revelado significativa é a própria equipa de enfermagem e os profissionais que a constituem pois não existe medida mensurável ou não subjectiva que o justifique de outra maneira. A forma como o (s) enfermeiro (s) estão a abordar e a lidar com os

doentes em acompanhamento é o factor protector dos indivíduos que frequentam (uma ou mais vezes) essa consulta.

De entre as variáveis estudadas na modelação do tempo de sobrevivência foi de alguma forma inesperado o facto da variável TAR (o indivíduo ter iniciado ou não Terapêutica Antiretroviral) não se ter revelado estatisticamente significativa. De certa forma, seria expectável que o facto do indivíduo estar em TAR influenciasse o seu tempo de sobrevivência. A explicação efectiva para a não ocorrência deste facto poderá estar relacionada com a amostra estudada e talvez possa apenas ser desenvolvida pela equipa que lida diariamente com os indivíduos infectados. No entanto, existem duas características da TAR e inerentes ao início do tratamento antiretroviral que são contraditórias e podem estar a anular o suposto efeito da variável no tempo de sobrevivência dos indivíduos. Por um lado, a prática actual é que o doente tem de apresentar situação clínica individual justificável e/ou estar já na fase sintomática com doença definidora de SIDA para poder iniciar o tratamento (o que implica que estes indivíduos apresentem nessa data maior risco de morte e portanto menor tempo de sobrevivência esperado), mas por outro lado, a partir do momento em que o indivíduo inicia o tratamento espera-se também que a doença passe para um nível crónico com o controlo da infecção (aumentando assim nesta fase o tempo de sobrevivência esperado do indivíduo).

Não tendo sido explorado, um futuro rumo de investigação será considerar também modelos de fragilidade. Embora se tenha admitido que os indivíduos são independentes entre si, a introdução de um efeito aleatório na função de risco e respectiva análise crítica poderão vir a explicitar melhor a realidade que se pretende estudar e dessa forma aperfeiçoar o grau de explicabilidade dos dados. O objectivo é tentar captar alguma da heterogeneidade que não está presente nas variáveis significativas do modelo.

Espera-se que os resultados e conclusões obtidas na prossecução deste estudo sejam de alguma forma úteis para os profissionais cujo desafio é a abordagem e o acompanhamento do doente infectado com VIH/SIDA, no sentido de melhorar as suas estratégias e por conseguinte a saúde e a qualidade de vida dos seus pacientes, assim como para os responsáveis ou gestores das instituições susceptíveis de receber e tratar estes indivíduos.

Espera-se também que, para os responsáveis pelo desenvolvimento de campanhas de sensibilização e prevenção da doença, o conjunto de características identificadas nos indivíduos estudados permita uma maior aproximação à situação na comunidade na qual se pretende intervir.

Se existir oportunidade e interesse por parte dos indivíduos infectados e respectivos familiares, espera-se ainda que o presente estudo seja um instrumento de transmissão de conhecimento e por conseguinte de aproximação e compreensão entre eles, respectivas famílias, e a doença em si.

## 7. Referências Bibliográficas

### Referências Literárias

- I. Albino, C. Adesão à terapêutica. Curso “Abordagem ao Doente com VIH Sida”. Unidade Formativa do Centro de Formação Multiprofissional do Centro Hospitalar do Baixo Alentejo. Junho 2008;
- II. Allison, P. *Event History Analysis*, in Hardy M. and A. Bryman (Ed.), *Handbook of Data Analysis*, London: Sage Publications. 2004;
- III. Amaro, F.; Frazão, C.; Pereira, M.E.; Teles, L. *HIV/AIDS risk perception, attitudes and sexual behaviour in Portugal*. International Journal of STD & AIDS 2004; (15: 56-60);
- I. Balows, A.; Hausler Jr., W.J.; Hermann, K.L.; Isenberg, H.D.; Shoadomy, H.J. *Manual of Clinical Microbiology – Fifth Edition*. American Society for Microbiology. Washington D.C. 1991. (Chapter 100 – pag 1011 a 1023);
- II. Bastos, J.; Rocha, C. Análise de Sobrevivência. *ArquiMed*. ISSN 0871-3413. 2006;
- III. Box-Steffensmeier, J.M., e B.S. Jones, *Event History Modeling: A Guide for Social Scientists*. Cambridge: Cambridge University Press. 2004;
- IV. Carvalho, Marília Sá. *Sebenta Análise de Sobrevida*. PROCC/FIOCRUZ;
- V. Carvalho, Marília Sá; Andreuzzi, V.L.; Codeco, C.T.; Barbosa, M.T; Shimakura, S.E. *Análise de sobrevida. Teoria e Aplicações em Saúde*. Rio de Janeiro. Editora Fiocruz. 2005;
- VI. Chaves, Inês. *Sida e Cuidados Paliativos*. 2008;
- VII. Cleves, M.A.; Gould, W.; Gutierrez, R. *An Introduction to Survival Analysis using STATA* (Revised Edition). Texas: Stata Press. 2004;
- VIII. Collet, D. *Modelling Survival Data in Medical Research*. London: Chapman & Hall.. 1994;
- IX. Cox, D.R., *Regression Models and Life-Tables*. *Journal of the Royal Statistical Society. Series B (Methodological)* 34 (2), 187-220. 1992;
- X. Faria, Telo. *Infecção VIH/SIDA, que repercursões sociológicas?* Revista Informação SIDA e outras doenças infecciosas n.º 80 – Maio/Junho. Lisboa. 2010;
- XI. Faria, Telo. *Entrevista publicada no Jornal Diário do Alentejo de 29 de Janeiro, páginas 10 e 11*. Beja. 2010;
- XII. Friedland, G.; Shaltzman, B.; Vileno, J.; Freeman, K.; Schragar, L.K.; Klein, R.S. *Survival Differences in Patients with AIDS*. *Journal of Acquired Immune Deficiency Syndromes*. 4:144-153. Raven Press. New York. 1991;
- XIII. Hosmer, D.W., e S. Lemeshow. *Applied Survival Analysis: Regression Modelling of Time to Event Data*. New York: John Wiley & Sons. 1999;
- XIV. Jahn, A.; Floyd, S.; Crampin, A.C.; Hwaungulu, F.; Mvula, H.; Munthali, F.; McGrath, N.; Mwafilaso, J.; Mwinuka, V.; Mangongo, B.; Fine, P.E.; Zaba, B.; Glynn, J.R. *Population-level effect of HIV on adult mortality and early evidence of reversal after introduction of antiretroviral therapy in Malawi*. *The Lancet* 371: 1603-1611. 2008;

- XV. Jenkins, S. Survival Analysis, *Lecture Notes Manuscript, Unpublished manuscript*. 2005;
- XVI. Kaplan, E.L.; Meier, P. Nonparametric estimation from incomplete observations. *J. of American Statistical Association*. 1958;
- XVII. Kitahata, M.M.; Koepsell, T.D; Deyo, R.A.; Maxwell, C.L.; Dodge, W.; Wagner, E.H. *Physicians' experience with acquired immunodeficiency syndrome as a factor in patients' survival. The New England Journal of Medicine*. 334(11): 701-706. 1996;
- XVIII. Lee, E.T., e J.W. Wang. *Statistical Methods for Survival Data Analysis* (3<sup>rd</sup> Edition). New Jersey: John Wiley & Sons. 2003;
- XIX. Lemp, G.; Payne, S.F.; Neal, D.; Temelso, T.; Rutherford, G. *Survival Trends for Patients with AIDS. The Journal of the American Medical Association*. 263(3):402-406. 1990;
- XX. Marins, J.R.; Jmal, L.F.; Chen, S.Y.; Barros, M.B.; Hudes, E.S.; Barbosa, A.A.; Chequer, P.; Teixeira, P.R.; Hearst, N. *Dramatic improvement in survival among adult Brazilian AIDS patients. Epidemiology & Social*. 17 (11): 1675-1682. 2003;
- XXI. Mata, A. *Imunodeficiências*. Instituto Superior de Ciências da Saúde – Sul, Caparica. 2006;
- XXII. Mauch, S. *Situational Assessment of the HIV/AIDS Notification System – A Portuguese Experience. National Coordination for HIV Infection*. 2009;
- XXIII. Miranda, A. Evolução natural da infecção por VIH – Aspectos Clínicos. *Ver Port Clin Geral*; 19: 587-97; 2003;
- XXIV. Montagnier, L. *Dês vírus et dès hommes*. Paris, Odile Jacob, 1994;
- XXV. Morgan, D.; Mahe, C; Mayanja, B.; Okongo, R.L.; Whitworth, J. *HIV-1 infection in rural Africa: is there a difference in median time to AIDS and survival compared with that in industrialized countries? AIDS* 16: 597-603. 2002;
- XXVI. Nájera, R. *SIDA – Respuestas y orientaciones*. Aguilar. Gráfica Internacional. Madrid. 1997;
- XXVII. Neves, J. História e Epidemiologia – VIH/SIDA. Curso “Abordagem ao Doente com VIH Sida”. Unidade Formativa do Centro de Formação Multiprofissional do Centro Hospitalar do Baixo Alentejo. 2008;
- XXVIII. Oliveira, A. História natural da infecção pelo VIH. Departamento de Doenças Infecciosas dos Hospitais da Universidade de Coimbra. Faculdade de Medicina da Universidade de Coimbra;
- XXIX. Pereira, A. Sida ainda alimenta preconceitos. *Jornal do Centro de Saúde – Ano 4, n.º 47 – Dezembro 2008 – páginas 12 e 13*. Lisboa. 2008;
- XXX. Signorini, D.; Codeço, C.; Carvalho, M. S.; Campos, D.P.; Monteiro, M.; Andrade, M.F.; Pinto, J.F.; Sá, C.A. Effect of *sociodemographic, clinical-prophylactic and therapeutic procedures on survival of AIDS patients assisted in a Brazilian outpatient clinic*. *Revista Brasileira de Epidemiologia*. 8 (3). 2005;
- XXXI. Therneau, T.M.; Grambsch, P.; Fleming, T. *Martingale-Based Residuals for Survival Models, Biometrika* 77 (1), 147-160. 1990;
- XXXII. Ventura, A e Pires, J – SIDA: História de uma epidemia. *Revista Delaofil*. - 15; 1:49-56. 2005;



- XXXIII. Volberding, P. A. Abordagem da infecção VIH – Livro de Bolso da Equipa de Tratamento. Burroughs Wellcome co.São Francisco. EUA;
- XXXIV. Washington C., Winn. Stephen Allen. Willam, Janda. Elmer Koneman. Gary, Procop. Paul, Schreckenberger. Gail, Woods. *Koneman's Color Atlas and Textbook of Diagnostic Microbiology. Sixth edition.* Lippincott Williams & Wilkins. Philadelphia. USA. 2006;

## Referências electrónicas

- [1] Manual Merck – Biblioteca Médica On-line – Edição de Saúde para a Família. Secção 16 Doenças do Sistema Imunitário, Capítulo 167 Biologia do Sistema Imunitário (<http://www.manualmerck.net/> consultado a 10-11-2009);
- [2] Manual Merck – Biblioteca Médica On-line – Edição de Saúde para a Família. Secção 16 Doenças do Sistema Imunitário, Capítulo 168 Doenças devidas a imunodeficiência (<http://www.manualmerck.net/> consultado a 12-11-2009);
- [3] Manual Merck – Biblioteca Médica On-line – Edição de Saúde para a Família. Secção 17 Infecções, Capítulo 187 Infecção pelo vírus da imunodeficiência (<http://www.manualmerck.net/> consultado a 12-11-2009);
- [4] <http://www.roche.pt/sida/virus/> consultado a 23-05-2008 e a 08-12-2009;
- [5] [http://www.who.int/topics/hiv\\_aids/en/](http://www.who.int/topics/hiv_aids/en/) consultado a 09-11-09;
- [6] <http://hivmedicine.aidsportugal.com/> consultado a 14-11-2009;
- [7] [http://diario.iol.pt/noticia.html?id=65328&div\\_id=4071](http://diario.iol.pt/noticia.html?id=65328&div_id=4071) consultado a 19-05-2009;
- [8] <http://www.portaldasaude.pt/> consultado a 23-11-09;
- [9] <http://www.acss.min-saude.pt/> consultado a 23-11-09;
- [10] <http://www.aidsportugal.com/article.php?sid=9875> consultado a 23-11-09;
- [11] <http://linfocit.blogspot.com> consultado a 15-09-2009;
- [12] <http://www.aidsportugal.com/article.php?sid=3170> consultado a 23-11-09;
- [13] <http://www.aidscongress.net/> consultado a 12-01-2010;
- [14] <http://sobrevida.fiocruz.br/> consultado a 09-04-2010;

## Outras Referências

- a. *Aids epidemic update 2007.* Geneva, *World Health Organization/ UNAIDS*, 2007;
- b. *Report on the global AIDS epidemic.* Geneva, UNAIDS, 2008;
- c. Programa Nacional de Prevenção e Controlo da Infecção VIH/SIDA 2007-2010 – Um Compromisso com o Futuro. Lisboa, Comissão Nacional para a Infecção VIH/ SIDA, 2007;
- d. Hospitais SNS – Contrato-Programa 2008 – Metodologia para a definição de preços e fixação de objectivos – Anexo I – Modalidade de Pagamento para a Infecção VIH/ SIDA. Lisboa, ACSS, 2007;
- e. Infecção VIH/SIDA. A Situação em Portugal a 31 de Dezembro de 2008. Doc 140. Departamento de Doenças Infecciosas – Unidade de Referência e Vigilância Epidemiológica – Núcleo de Vigilância

- Laboratorial de Doenças Infecciosas. Instituto Nacional de Saúde Doutor Ricardo Jorge. Ministério da Saúde. Lisboa. 2009;
- f. HIV/AIDS – Surveillance in Europe. End-Year Report 2006, n. ° 75. Eurohiv. France. December 2007;
  - g. Estudo dos casos de SIDA, CRS e PA com residência no distrito de Beja. Vigilância Epidemiológica da Infecção VIH e SIDA. DDI – URVE. 2009;
  - h. Recomendações Portuguesas para o Tratamento da Infecção VIH/sida. CNLCS. Discussão Pública. Fevereiro 2009;
  - i. *Le SIDA parlons-en – Guide de développement dès messages sur les IST/VIH/SIDA*. Première Édition. *Project Santé Familiale et Prévention du SIDA, Communication pour le Changement de Comportement. Côte d’Ivoire*. 2001;

## 8. Abreviaturas e siglas

- ACSS – Administração Central do Sistema de Saúde
- ADN – Ácido desoxirribonucleico
- ARN – Ácido ribonucleico
- CAD – Centros de Aconselhamento e Detecção Precoce do VIH/SIDA
- CAT – Centro de Apoio ao Toxicodependente
- CDC – *Center for Disease Control* (Agência Epidemiológica Federal Americana)
- CNLCS – Comissão Nacional de Luta Contra a SIDA
- CRS – Complexo Relacionado com SIDA (casos sintomáticos não-SIDA)
- CVET – Centro de Vigilância Epidemiológica das Doenças Transmissíveis
- DDI – URVE – Núcleo de Vigilância Laboratorial de Doenças Infecciosas do INSA
- ELISA - *Enzyme Linked Immuno Sorbent Assay*
- FDA - *Food and Drug Administration*
- HAART – *Highly Active AntiRetroviral Treatment*
- HTLV - *Human lymphotropic virus* (Vírus da Leucemia Linfocitária das Células T)
- IDU – *Injecting drug users*
- IF – Inibidor de Fusão
- II – Inibidor da integrase
- INSA – Instituto Nacional de Saúde Dr. Ricardo Jorge
- IP – Inibidor da protease
- KM – Kaplan-Meier
- LAV – *Lymphadenopathy Associated Virus* (Vírus associado a Linfadenopatias)
- LGP – Linfadenopatia generalizada persistente
- MSM – *Men who have sex with other men*
- NITR – Nucleósido inibidor da transcriptase reversa
- NNITR – Não nucleósido inibidor da transcriptase reversa
- OMS – Organização Mundial de Saúde
- ONUSIDA – Programa Conjunto das Nações Unidas para o VIH/ SIDA
- PA – Portador Assintomático
- PCR – *Polymerase Chain Reaction*
- SIDA – Síndrome da Imunodeficiência Adquirida
- TAR – Terapêutica antiretrovítica
- VIH – Vírus da Imunodeficiência Humana

## 9. Glossário

**ADN** – Ácido Desoxirribonucleico. É uma molécula biológica universal presente em todas as células vivas. Nele está contida toda a nossa informação genética, sob a forma de genes;

**ARN** – Ácido ribonucleico. É o responsável pela síntese de proteínas da célula. Formado por moléculas de dimensões muito inferiores às do ADN;

**Célula** – a mais pequena unidade de tecido viva, composta por um núcleo e um citoplasma e rodeada por uma membrana;

**Anticorpo** – proteína, fabricada por linfócitos B, que reage perante um antígeno específico. Também chamado de imunoglobina;

**Antígeno** – qualquer molécula capaz de estimular a resposta imune;

**Cancro** – Tumor maligno formado pela multiplicação desordenada de células de um tecido ou de um órgão;

**Complemento** – grupo de proteínas que ajuda a atacar antígenos;

**Citocinas** – proteínas solúveis, segregadas por células do sistema imunitário, que actuam como mensageiros para ajudar a regular a resposta imune;

**Dislipidemia** – é a presença de níveis elevados ou anormais de lipídios e/ou lipoproteínas no sangue. As anormalidades nos lipídios e lipoproteínas são extremamente comuns na população geral, e são consideradas um factor de risco altamente modificável para doenças cardiovasculares;

**ELISA** – teste imunoenzimático que permite a detecção de anticorpos ou antígenos específicos no plasma sanguíneo;

**Enzima** – substância orgânica de natureza proteica que tem a capacidade de catalisar reacções químicas;

**Gânglios Linfáticos** – ou linfonodos. São pequenos órgãos perfurados por canais que existem em diversos pontos da rede linfática, actuando na defesa do organismo humano e produzindo anticorpos;

**Epidemia** - Doença que, numa localidade ou região, ataca simultaneamente muitas pessoas;

**Genoma** – sequência completa de ADN;

**Imunidade inata** – os componentes do sistema imunitário que nascem com a pessoa, e reagem de forma semelhante perante todas as substâncias estranhas;

**Imunidade aprendida** – o sistema imunitário aprende a responder a cada novo antígeno com que se confronta;

**Incidência** – ocorrência de novos casos de uma dada doença ou condição;

**Leucócito** – glóbulo branco. São exemplos os linfócitos e neutrófilos;

**Linfadenopatia** – crescimento dos gânglios linfáticos, considerado um sinal de doença;

**Linfócito** – célula principal do sistema linfático. Classificam-se em linfócitos B (que produzem anticorpos) e linfócitos T (que ajudam o corpo a distinguir o próprio do alheio);

**Linfoma** – tumores que se iniciam a partir da transformação de um linfócito no sistema linfático;

**Macrófago** – célula que absorve ou ingere micróbios devidamente assinalados pelo sistema imunitário para que sejam destruídos;

**Micro litro** – unidade de volume equivalente à milionésima parte de um litro. Representada pelo símbolo  $\mu\text{l}$ . ( $1 \mu\text{L} = 0,001 \text{ ml} = 10^{-6} \text{ L}$ );

**Molécula** – grupo de átomos quimicamente combinados para formar uma única substância química;

**Neoplasia** – tumor invasivo e metastisante; termo que designa alterações celulares que acarretam um crescimento exagerado dessas células, ou seja, proliferação celular anormal, sem controlo e autonomia, na qual reduzem ou perdem a capacidade de se diferenciar, em consequência de mudanças nos genes que regulam o crescimento e a diferenciação celulares;

**Neutrófilo** – leucócito que ingere antígenos e outras substâncias;

**Nosocomial** – doença infecciosa adquirida por doentes ou profissionais de saúde em meio hospitalar ou qualquer outra unidade de saúde. Também chamada de infecção associada aos cuidados de saúde (IACS);

**Organismo** – o conjunto de órgãos que constituem um indivíduo vivo;

**Pandemia** - Surto de uma doença com distribuição geográfica muito alargada;

**PCR** – reacção em cadeia da polimerase (uma enzima). É um método de amplificação (criação de múltiplas cópias) de ADN ou ARN sem o uso de um organismo vivo. É uma das técnicas mais comuns utilizadas em laboratórios de pesquisas médicas e biológicas para diversas tarefas, entre as quais a detecção de antígenos e anticorpos no diagnóstico de doenças infecciosas;

**Prevalência** – total de casos de uma dada doença ou condição na população;

**Resposta Imune** – a resposta do sistema imunitário perante um antígeno, por via de células ou anticorpos;

**Retrovírus** – tipo de vírus que armazena informação genética como ARN e não como ADN. Quando o vírus entra numa célula hospedeira, liberta o seu ARN e uma enzima (transcriptase reversa) e em seguida cria ADN usando o ARN viral como modelo. Depois o ADN viral é incorporado no ADN da célula hospedeira. Cada vez que uma célula hospedeira se divide, faz também uma nova cópia do ADN viral integrado juntamente com os seus próprios genes. O ADN viral pode controlar as funções da célula (pode tornar-se activo), fazendo com que esta produza novas partículas do vírus. Estes novos vírus são libertados da célula infectada para invadir outras células;

**Sarcoma de Kaposi** – doença sistémica multifocal, maligna, que é originada a partir do endotélio vascular e tem um curso clínico variável. As manifestações mais frequentes da doença são lesões cutâneas, mas as membranas mucosas, os sistemas linfático e visceral – em particular o pulmão e o tracto gastrointestinal – podem estar envolvidos. Estão descritas quatro formas clínicas: Sarcoma de Kaposi clássico; Sarcoma de Kaposi secundário à imunossupressão; Sarcoma de Kaposi Endémico Africano; e o Sarcoma de Kaposi associado ao VIH;

**Seroconversão** – resposta imune contra o vírus (altura em que surgem no organismo os anticorpos contra o VIH);

**Sistema linfático** – rede de gânglios linfáticos ligados entre si por vasos linfáticos. Esta rede filtra, ataca e destrói organismos prejudiciais que causam infecções. Os gânglios linfáticos costumam agrupar-se em zonas em que os vasos linfáticos se ramificam, como o pescoço, axilas e virilhas; Outros órgãos e tecidos corporais (o timo, o fígado, o baço, o apêndice, a medula óssea e pequenas aglomerações de tecido linfático como as amígdalas na garganta e as placas de Peyer no intestino delgado) fazem também parte do sistema linfático;

**Sorologia** – ou Serologia. É o estudo científico do soro sanguíneo. Refere-se ao diagnóstico e identificação de anticorpos e ou antigénios no soro;

**Tumor** – termo genérico que indica um aumento anormal de uma parte ou da totalidade de um tecido. Pode ser benigno, benigno e comportar-se como maligno, ou efectivamente maligno (neste caso também chamado de neoplasia);

**Vírus** – um minúsculo organismo infeccioso que necessita de uma célula viva para se reproduzir. O vírus adere a uma célula, geralmente de um tipo específico. Uma vez dentro dela, liberta o seu ADN ou ARN (que contém a informação necessária para criar novas partículas de vírus) e assume o controlo de alguns processos metabólicos da mesma. Como consequência, os componentes do vírus são fabricados dentro da célula e reunidos adequadamente para que o vírus seja libertado e continue a manter a sua capacidade infectante;

**Vírión** – uma partícula viral completa que está fora da célula hospedeira. É constituída por ADN ou ARN e cercado por uma proteína. Cada espécie de vírus apresenta vírións de formatos diferentes.

## **10. Anexos**

