

UNIVERSIDADE NOVA DE LISBOA  
Faculdade de Ciências e Tecnologia  
Departamento de Engenharia Electrotécnica

# SISTEMA ONLINE DE APOIO À DECISÃO PARA MÉDICOS ONCOLOGISTAS

Pedro Luís Alves

Dissertação apresentada na Faculdade de Ciências e Tecnologia da Universidade Nova de Lisboa para a obtenção do grau de Mestre em Engenharia Electrotécnica e de Computadores

**Orientador:** José Manuel Matos Ribeiro da Fonseca

**Outubro 2010**  
Lisboa



*Aos meus Pais*  
*à minha Irmã*  
*à lembrança dos meus Avós*



# Agradecimentos

Ao Prof. José Manuel Matos Ribeiro da Fonseca agradeço a oportunidade concedida para a realização deste trabalho, bem como toda a motivação e apoio que sempre me proporcionou.

Um especial agradecimento à Eng<sup>a</sup>. Ana Sofia pela total disponibilidade, ajuda, sugestões, correcções e revisões relativas a este trabalho. A sua dedicação foi sem dúvida uma forte motivação para mim neste último período do curso.

Os meus sinceros agradecimentos a todos os meus colegas e amigos de curso que me acompanharam ao longo destes anos, Alexandre Rodrigues, André Eusébio, Magno Guedes, Nelson Alves, David Chaínho, José Belo, Paulo Santos, Gonçalo Cabeçinhas, Mário Salgueiro e Filipe Valente.

À Ana Pena pela sua constante boa disposição, que muitas vezes me distraía e dava alegria quando tudo corria menos bem.

À Eva pela amizade, carinho e paciência que demonstrou para corrigir este trabalho. Obrigado por tudo.

Uma palavra de apreço à minha família por ser uma fonte de apoio e coragem.



# Sumário

Nesta dissertação é proposto um sistema de apoio à decisão para médicos oncologistas e pacientes de cancro da mama.

O sistema desenvolvido, denominado de *BCSurvival*, é um sistema web que tem em consideração diversas características individuais de um paciente com cancro de mama permitindo aos utilizadores identificar o nível de risco em que o paciente se encontra, bem como o intervalo de confiança e as regras associadas a esse nível de risco. O sistema disponibiliza ainda curvas de sobrevivência dos vários níveis de risco identificando a curva onde o paciente se enquadra.

O estudo foi realizado com uma base de dados obtida pelo *Christie Hospital*, Wilmslow, Manchester entre 1990-1994. Estes dados foram utilizados em três modelos de prognóstico para a obtenção do risco, intervalos de confiança e estratificação dos pacientes em grupos de risco. Os modelos de prognóstico incorporados no *web site* são: o *Nottingham Prognostic Index* (NPI), a Regressão de Cox e o *Partial Logistic Artificial Neural Network with Automatic Relevance Determination* (PLANN-ARD).

O sistema *web* oferece também ao médico a capacidade de organizar a informação referente ao prognóstico dos seus doentes, apresentando essa informação numa página *web* especialmente desenhada para esse efeito.

**Palavras-Chave:** cancro da mama, análise de sobrevivência, sistema de suporte à decisão.





# Abstract

This dissertation suggests a web clinical decision support system for clinical oncologists and for breast cancer patients.

The developed system, which was named BCSurvival, is a web-based system which takes into account certain individual characteristics of the patient. It allows users to identify the patients' risk level, the confidence interval and the rules associated with that level of risk, providing an application through which the user can have a broader view of the risk. In an alternative view, the system provides survival curves of the various levels of risk, identifying one where the patient fits.

This study was developed using the data set recruited by Christie Hospital, in Winlmslow Manchester, in this period 1990-1994. This data was used by three prediction models to obtain the risk, confidence intervals and to stratify patients into risk groups.

This system comprises three different prognostic modelling methodologies: the clinically widely used Nottingham prognostic index (NPI); the Cox regression modelling and a Partial Logistic Artificial Neural Network with Automatic Relevance Determination (PLANN-ARD).

The web system also offers the physician the ability to organize information on the prognosis of their patients, presenting this information on a section specially designed for this purpose.

**Keywords:** breast cancer, survival analysis, decision support systems.



# Lista de Acrónimos

| <b>Acrónimo</b> | <b>Descrição</b>  |
|-----------------|---|
| AJAX            | <i>Asynchronous JavaScript and XML</i>                    |
| API             | <i>Application Programming Interface</i>                  |
| ASCII           | <i>American Standard Code for Information Interchange</i> |
| ASP             | <i>Active Server Pages</i>                                |
| ASP.NET         | <i>Active Server Pages .NET</i>                           |
| BCSurvival      | <i>Breast Cancer Survival</i>                             |
| CART            | <i>Classification And Regression Tree</i>                 |
| CERN            | <i>European Organization for Nuclear Research</i>         |
| CLR             | <i>Common Language Runtime</i>                            |
| CSS             | <i>Cascading Style Sheets</i>                             |
| DHTML           | <i>Dynamic HTML</i>                                       |
| DLL             | <i>Dynamic-Link Library</i>                               |
| DOM             | <i>Document Object Model</i>                              |
| GUI             | <i>Graphical User Interface</i>                           |
| HTML            | <i>HyperText Markup Language</i>                          |
| IIS             | <i>Internet Information Services</i>                      |
| IP              | <i>Internet Protocol</i>                                  |
| LINQ            | <i>Language Integrated Query</i>                          |

| <b>Acrónimo</b> | <b>Designação</b>  |
|-----------------|--|
| MCR.            | <i>MATLAB Component Runtime</i>  |
| MIME            | <i>Multimedia Internet Mail Extensions</i>   |
| NCSA            | <i>NATO Communication and Information Systems Services Agency</i>                        |
| NPI             | <i>Nottingham Prognostic Index</i>   |
| PLANN-ARD       | <i>Partial Logistic Artificial Neural Network with Automatic Relevance Determination</i> |
| SOAP            | <i>Simple Object Access Protocol</i>   |
| TMN             | <i>Tumor, Nódulos, Metástase</i>   |
| URL             | <i>Uniform Resource Locator</i>  |
| W3C             | <i>World Wide Web Consortium</i>   |
| WWW             | <i>World Wide Web</i>  |
| WSDL            | <i>Web Services Description Language</i>   |
| XAML            | <i>Extensible Application Markup Language</i>  |
| XML             | <i>eXtensible Markup Language</i>  |
| XHTML           | <i>eXtensible Hypertext Markup Language</i>  |

# Conteúdo

|          |   |           |
|----------|---|-----------|
| <b>1</b> | <b>Introdução</b>   | <b>1</b>  |
| 1.1      | Enquadramento . . . . .                                   | 1         |
| 1.2      | Contribuições da Dissertação . . . . .                    | 2         |
| 1.3      | Estrutura da Dissertação . . . . .                        | 3         |
| 1.4      | Artigo Publicado . . . . .                                | 5         |
| <b>2</b> | <b>Cancro da Mama</b>                                     | <b>7</b>  |
| 2.1      | O que é o cancro da mama? . . . . .                       | 7         |
| 2.2      | Tratamentos para o cancro da mama . . . . .               | 9         |
| <b>3</b> | <b>Modelos de Prognóstico do Cancro da Mama</b>           | <b>11</b> |
| 3.1      | Introdução . . . . .                                      | 11        |
| 3.2      | Análise de sobrevivência . . . . .                        | 12        |
| 3.2.1    | Curvas de <i>Kaplan-Meier</i> . . . . .                   | 14        |
| 3.3      | Modelos de prognóstico . . . . .                          | 16        |
| 3.3.1    | Princípios dos modelos de prognóstico . . . . .           | 17        |
| 3.4      | Descrição e obtenção da base de dados utilizada . . . . . | 18        |
| 3.5      | Modelos prognóstico implementados . . . . .               | 20        |
| 3.5.1    | <i>Nottingham Prognostic Index</i> . . . . .              | 20        |
| 3.5.2    | <i>Cox Proportional Hazards Modelling</i> . . . . .       | 22        |

|          |  |           |
|----------|--|-----------|
| 3.5.2.1  | Selecco das variveis de prognstico . . . . .                | 23        |
| 3.5.3    | <i>PLANN-ARD Modelling</i> . . . . .                           | 25        |
| 3.6      | Estratificao em grupos de risco . . . . .                     | 27        |
| <b>4</b> | <b>WEBSITE - <i>Breast Cancer Survival</i>, Implementao</b>   | <b>29</b> |
| 4.1      | Introduo . . . . .  | 29        |
| 4.2      | Tecnologias utilizadas pelo Site - <i>BCSurvival</i> . . . . . | 29        |
| 4.2.1    | Estrutura Geral do Site . . . . .                              | 29        |
| 4.2.2    | Site . . . . .   | 31        |
| 4.2.3    | HTML . . . . .   | 32        |
| 4.2.4    | CSS . . . . .  | 34        |
| 4.2.5    | ASP.NET . . . . .  | 34        |
| 4.2.5.1  | ASP.NET <i>Master Pages</i> . . . . .                          | 38        |
| 4.2.5.2  | ASP.NET <i>AJAX Control Toolkit</i> . . . . .                  | 39        |
| 4.2.6    | JavaScript . . . . .   | 40        |
| 4.2.7    | jQuery . . . . .   | 41        |
| 4.2.8    | SILVERLIGHT . . . . .  | 43        |
| 4.2.9    | IIS . . . . .  | 46        |
| 4.2.10   | <i>Web Services</i> . . . . .                                  | 47        |
| 4.2.11   | SQL Server 2005 . . . . .                                      | 48        |
| 4.2.12   | LINQ . . . . .   | 49        |
| 4.2.13   | MATLAB . . . . .   | 50        |
| 4.2.13.1 | Matlab <i>Compiler</i> . . . . .                               | 50        |
| 4.3      | Implementao do Site - <i>BCSurvival</i> . . . . .             | 51        |
| 4.3.1    | Base Dados . . . . .   | 51        |
| 4.3.2    | Interface Grfica . . . . .                                    | 52        |
| 4.3.3    | Colocao do Site <i>online</i> . . . . .                       | 56        |
| 4.3.4    | Sistema de suporte  deciso . . . . .                         | 62        |

|          |  |           |
|----------|--|-----------|
| <b>5</b> | <b>WEBSITE - <i>Breast Cancer Survival</i>, Funcionalidades</b>            | <b>65</b> |
| 5.1      | Introdução . . . . .   | 65        |
| 5.2      | Funcionalidades do Site - <i>BCSurvival</i> . . . . .                      | 66        |
| 5.2.1    | <i>Home Page</i> . . . . .   | 66        |
| 5.2.2    | <i>Login</i> . . . . .   | 67        |
| 5.2.3    | Menu de Topo . . . . .   | 68        |
| 5.2.4    | Menu de Utilizador . . . . .   | 69        |
| 5.2.5    | Menu de Gestão de Pacientes . . . . .                                      | 70        |
| 5.2.6    | Menu de Prognósticos . . . . .   | 74        |
| 5.3      | Comparação com outros sistemas de identificação de risco de cancro da mama | 79        |
| <b>6</b> | <b>Conclusões</b>  | <b>87</b> |
| 6.1      | Discussão dos Resultados . . . . .   | 87        |
| 6.2      | Trabalho Futuro . . . . .  | 89        |
| <b>A</b> | <b>Regras para o modelo <i>Cox Proportional Hazard</i></b>                 | <b>91</b> |
| <b>B</b> | <b>Regras para o modelo <i>PLANN-ARD</i></b>                               | <b>93</b> |
| <b>C</b> | <b>Árvore de decisão para o modelo <i>Cox Proportional Hazards</i></b>     | <b>95</b> |
| <b>D</b> | <b>Árvore de decisão para o modelo <i>PLANN-ARD</i></b>                    | <b>97</b> |





# Lista de Figuras

|      |  |    |
|------|--|----|
| 2.1  | Mecanismo de formação do cancro. . . . .   | 7  |
| 2.2  | Mecanismo de formação de metástases. . . . .   | 8  |
| 2.3  | Evolução do Cancro da Mama. . . . .  | 9  |
| 3.1  | Análise de sobrevivência num grupo de 5 indivíduos . . . . .                             | 13 |
| 3.2  | Curva de sobrevivência num período de 60 meses . . . . .                                 | 15 |
| 4.1  | Diagrama do <i>site</i> - <i>BCSurvival</i> . . . . .                                    | 30 |
| 4.2  | Localização dos dados alterados na página . . . . .                                      | 41 |
| 4.3  | Envio de email para um paciente . . . . .  | 43 |
| 4.4  | Aplicação criada em Silverlight . . . . .  | 44 |
| 4.5  | Processo de pedido-resposta de uma página ASP.NET . . . . .                              | 46 |
| 4.6  | Página inicial do <i>site</i> <i>BCSurvival</i> . . . . .                                | 53 |
| 4.7  | Página inicial do <i>site</i> <i>BCSurvival</i> , para utilizadores registados . . . . . | 53 |
| 4.8  | <i>Template</i> por defeito . . . . .  | 54 |
| 4.9  | Menu de Topo . . . . .   | 55 |
| 4.10 | <i>Labels</i> em formato rectangular com os cantos arredondados . . . . .                | 55 |
| 4.11 | Favicon . . . . .  | 56 |
| 4.12 | Home Page do IIS . . . . .   | 57 |
| 4.13 | Criação de um directório virtual . . . . .   | 58 |

|      |  |    |
|------|--|----|
| 4.14 | Propriedade do <i>site</i> - <i>BCSurvival</i> . . . . .                         | 58 |
| 4.15 | Adição de um novo tipo de MIME . . . . .   | 62 |
| 4.16 | Aplicação Silverlight . . . . .  | 62 |
| 4.17 | Diagrama do <i>site</i> - <i>BCSurvival</i> . . . . .                            | 63 |
| 5.1  | <i>Home page</i> do <i>site</i> - <i>BCSurvival</i> . . . . .                    | 66 |
| 5.2  | Área de <i>Login</i> . . . . .   | 67 |
| 5.3  | Página de registo. . . . .   | 67 |
| 5.4  | Página de recuperação. . . . .   | 67 |
| 5.5  | Menu de Topo. . . . .  | 68 |
| 5.6  | Página de FAQ. . . . .   | 68 |
| 5.7  | Página de Contactos. . . . .   | 68 |
| 5.8  | Menu de Utilizador. . . . .  | 69 |
| 5.9  | <i>Tab</i> de Informação pessoal. . . . .  | 69 |
| 5.10 | <i>Tab</i> de publicações. . . . .   | 69 |
| 5.11 | Menu de Gestão de Pacientes. . . . .   | 70 |
| 5.12 | <i>Tab</i> para adicionar e editar pacientes. . . . .                            | 71 |
| 5.13 | <i>Tab</i> para gerir os pacientes. . . . .                                      | 72 |
| 5.14 | <i>Tab</i> para gerir os prognósticos dos pacientes. . . . .                     | 73 |
| 5.15 | <i>Home page</i> de prognósticos para os utilizadores autenticados. . . . .      | 74 |
| 5.16 | Grupo de risco do modelo PLANN-ARD e grupo de risco do NPI. . . . .              | 75 |
| 5.17 | Curvas de sobrevivência de <i>Kaplan Meier</i> . . . . .                         | 77 |
| 5.18 | Interface indicando o índice de risco PLANN-ARD e o índice de risco NPI. . . . . | 78 |

# Lista de Tabelas

|     |   |    |
|-----|---|----|
| 3.1 | Função de sobrevivência num período de 60 meses (método de <i>Kaplan-Meier</i> )      | 15 |
| 3.2 | Descrição das variáveis do conjunto de dados fornecidos pelo <i>Christie Hospital</i> | 19 |
| 3.3 | Parâmetros do modelo de Regressão de Cox . . . . .                                    | 24 |
| 5.1 | Sistemas de identificação de risco de cancro da mama <i>online</i> . . . . .          | 85 |
| A.1 | Regras obtidas pelo CART para o modelo <i>Cox Proportional Hazard</i> . . . . .       | 91 |
| B.1 | Regras obtidas pelo CART para o modelo PLANN-ARD . . . . .                            | 93 |



# Introdução

## 1.1 Enquadramento

Num mundo ideal, assim que um tumor maligno é diagnosticado, o tratamento é administrado para efectuar a cura. Quando a cura não é possível, a estimativa do tempo em que o tumor pode reaparecer ou, mais importante ainda, causar a morte, são questões muitas vezes colocadas.

De um modo geral, depois de ser comunicada a notícia de cancro há lugar à introspecção. A pessoa procura investigar *online* a etiologia, a sintomatologia e ocorrem interrogações sobre as possibilidades de cura. A internet constitui frequentemente um escape, um lugar seguro onde, fora de juízos alheios, se procura dar um novo rumo à situação vivenciada. A internet permite aceder fácil e rapidamente a grandes quantidades de informação a uma escala global possibilitando conhecer os avanços que vão surgindo na área do cancro da mama, quer em termos científicos, quer no âmbito das organizações mundias de saúde. Muitas vezes, uma nova esperança renasce e o desânimo dá lugar à iniciativa de procurar saber mais e melhor [Fonseca et al., 2006].

Perante o exposto, compreende-se a necessidade e o desafio daqueles que pretendem tirar o máximo partido da internet como meio para obter respostas rápidas em tempo útil. É neste contexto que surge a necessidade de criação de uma plataforma que auxilie os médicos

e os pacientes a tomar uma melhor decisão sobre os tratamentos a efectuar. Para os médicos oncologistas e pacientes de cancro da mama qualquer informação fidedigna, seja sobre a doença ou sobre o tratamento a efectuar, é sempre uma mais-valia, especialmente quando essa informação se encontra à distância de um simples *click*.

A internet é portanto, hoje em dia, uma fonte de informação privilegiada, acessível a todos e utilizada para diversos fins. É nessa perspectiva, que a criação de um *web site* que incorpore um sistema de apoio à decisão com capacidade de previsão e organização de toda a informação inerente ao cancro da mama tem suscitado uma atenção especial da comunidade médica. Este sistema de apoio à decisão é sem dúvida uma mais valia uma vez que permite efectuar a previsão de sobrevivência do paciente ou mesmo o comportamento futuro da doença.

A criação de um sistema de apoio à decisão para médicos oncologistas do cancro da mama é o principal objectivo desta tese. Entendeu-se ser vantajoso reunir numa página web três modelos de prognóstico que permitissem visualizar e comparar os diferentes índices de prognóstico de sobrevivência, os respectivos intervalos de confiança e os grupos de risco a que cada paciente pertence. Os modelos matemáticos, juntamente com tecnologias *web*, constituem um meio privilegiado para atingir este objectivo. Procedeu-se então à escolha do *Nottingham Prognostic Index* (NPI), da Regressão de Cox e do *Partial Logistic Artificial Neural Network with Automatic Relevance Determination* (PLANN-ARD) como modelos de prognóstico fiáveis e aceites pela comunidade médica. Estes modelos de prognóstico, por sua vez, assentam num sistema tecnológico web o ASP.NET criado para esse efeito.

## 1.2 Contribuições da Dissertação

Pretendeu-se com esta dissertação criar um *web site* a utilizar quer por médicos oncologistas, quer por pacientes com cancro da mama, de forma a ajudar a tomada de decisão do tratamento a efectuar. O sistema permitirá identificar o nível de risco a 5 anos em que o paciente se encontra, bem como o intervalo de confiança e as regras associadas a esse nível

de risco, tendo em conta as características do paciente. Disponibilizará também as curvas de sobrevivência dos vários níveis de risco, identificando aquela onde o paciente se enquadra.

No *web site* serão incorporados três diferentes modelos de prognóstico: o *Nottingham Prognostic Index* (NPI), a Regressão de Cox e o *Partial Logistic Artificial Neural Network with Automatic Relevance Determination* (PLANN-ARD). Os três modelos permitem calcular diferentes índices de prognóstico que, por sua vez poderão ser comparados e analisados num gráfico de dispersão de forma a ajudar os médicos e pacientes a identificarem um prognóstico mais fundamentado.

Os principais contributos desta tese são os seguintes:

- Desenvolvimento e criação de um *web site* em ASP.NET, tendo em consideração os requisitos descritos e os estudos desenvolvidos anteriormente.
- Desenvolvimento e incorporação no *web site* de um componente, que permita a ligação do Matlab com o mesmo, uma vez que na metodologia PLANN-ARD, o risco de sobrevivência e os intervalos de confiança associados são calculados em ambiente Matlab.
- Criação de uma plataforma de gestão de pacientes que facilite a adição de novos pacientes e de novos prognósticos. Através do historial clínico do doente, o utilizador terá a possibilidade de comparar os prognósticos e tratamentos, melhorando assim a sua capacidade de decisão.
- Tornar a experiência do utilizador mais rica graficamente do que aquela que ele obteria com aplicações desktop, recorrendo ao uso de tecnologias como AJAX, jQuery e *Web Services*.

### 1.3 Estrutura da Dissertação

A dissertação encontra-se organizada da seguinte forma:

*Capítulo 1:* Apresenta em traços gerais os objectivos e as contribuições desta dissertação,

bem como a definição de um sistema de apoio à decisão para médicos oncologistas e a sua importância na sociedade.

*Capítulo 2:* Situa o leitor no universo do cancro da mama descrevendo o que é o cancro da mama e como se desenvolve. Na última secção descrevem-se alguns tratamentos para o cancro da mama.

*Capítulo 3:* Expõe os modelos de prognóstico utilizados em cancro da mama, os dados utilizados por esses modelos e as variáveis de maior significância para a análise e obtenção de prognósticos. Faz uma breve introdução a conceitos de análise de sobrevivência e a sua importância na interpretação de resultados. No final do capítulo é descrita a metodologia usada para estratificar os pacientes em diferentes grupos de sobrevivência.

*Capítulo 4:* Apresenta e descreve em detalhe a estrutura geral do *web site*. Inicialmente este capítulo dá ao leitor um enquadramento sobre o que é um *site*. Seguidamente é feita uma introdução às tecnologias necessárias à criação e aumento de funcionalidades do *web site* de acordo com as especificações desejadas. Por último, é descrita detalhadamente a implementação do *site*, desde o desenvolvimento do *design* gráfico, à base de dados e à sua colocação *online*.

*Capítulo 5:* Identifica as funcionalidades do *web site* - *BCSurvival* e demonstra as suas capacidades de utilização de diversas ferramentas de prognóstico do cancro da mama e da gestão de pacientes e seus prognósticos.

*Capítulo 6:* Faz a análise e discussão em torno das metodologias de prognóstico e da arquitectura web construída. Por último, apresenta algumas sugestões de trabalho a desenvolver no futuro.



*Apêndice A:* Regras obtidas pelo algoritmo de estratificação através do CART para o modelo de prognóstico *Cox Proportional Hazard*.

*Apêndice B:* Regras obtidas pelo algoritmo de estratificação através do CART para o modelo de prognóstico *PLANN-ARD*.

*Apêndice C:* Árvore de decisão para o modelo *Cox Proportional Hazard*.

*Apêndice D:* Árvore de decisão para o modelo *PLANN-ARD*.

*Bibliografia*

## 1.4 Artigo Publicado

Durante o desenvolvimento do presente trabalho, foi apresentado na conferência *DoCEIS'10* um artigo intitulado *A clinical decision support system for breast cancer patients*. Os *proceedings* deste artigo estão disponíveis no livro *Emerging Trends in Technological Innovation* nas páginas [122-129], Springer, 2010.



## Cancro da Mama

### 2.1 O que é o cancro da mama?

O organismo humano é constituído por muitos milhões de células que se reproduzem pelo processo de divisão celular. Em condições normais, este processo é ordenado e controlado, sendo responsável pela formação, crescimento e regeneração de tecidos saudáveis do corpo. Contudo, as células, por vezes, perdem a capacidade de limitar e comandar o seu próprio crescimento passando então, a dividir-se e multiplicar-se muito rapidamente e de forma aleatória (figura 2.1 [Oncologia, 2010]). Estas alterações aumentam com a idade e por isso,

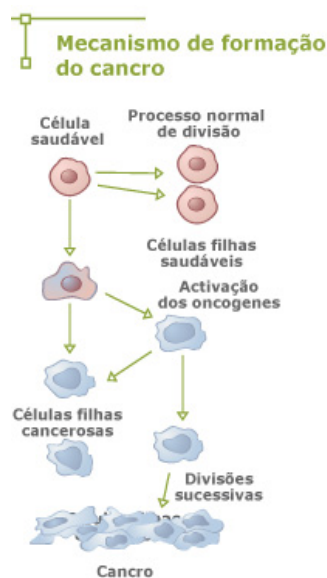


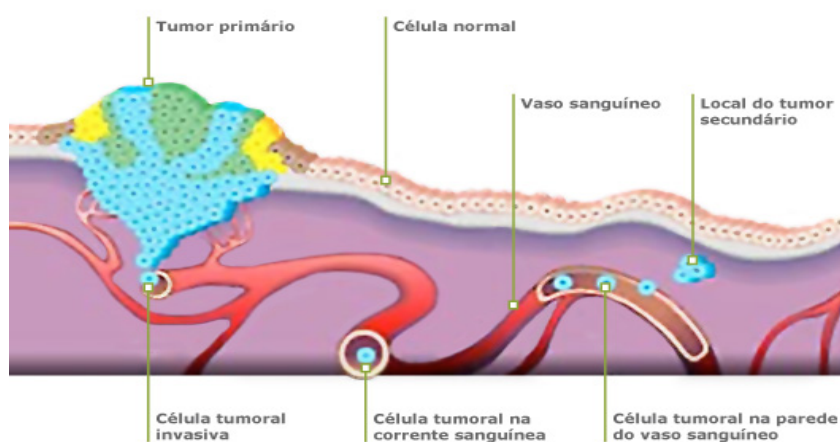
Figura 2.1: Mecanismo de formação do cancro.

## 2.1. O QUE É O CANCRO DA MAMA?

---

o cancro é mais frequente em idades avançadas. Como consequência desse processo de multiplicação e crescimento desordenado das células, ocorre um desequilíbrio na formação dos tecidos do corpo formando um tumor [Oncologia, 2010].

O cancro da mama é um tumor maligno que se desenvolve nas células do tecido mamário. A figura 2.2 (obtida no portal de oncologia [Oncologia, 2010]) ilustra o mecanismo de formação de metástases que ocorre quando as células tumorais “entram” nos vasos linfáticos ou nos vasos sanguíneos, passam à circulação, disseminam-se e “colonizam” outros órgãos. Este me-

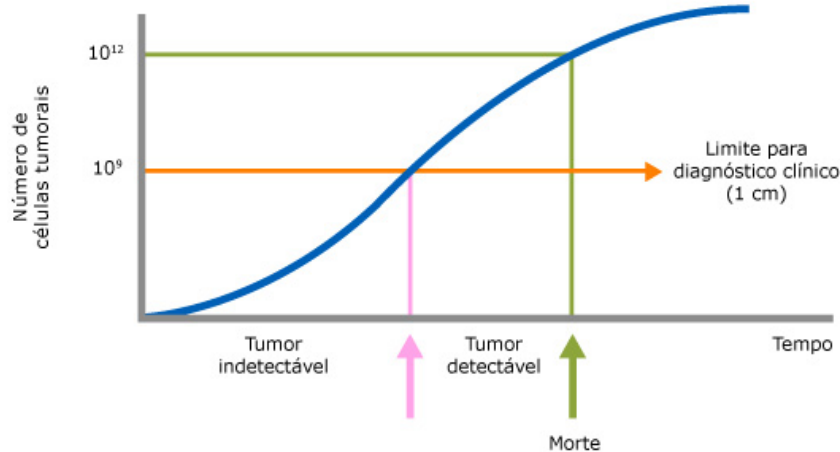


**Figura 2.2:** Mecanismo de formação de metástases.

canismo quando não é diagnosticado a tempo, permite ao tumor “espalhar-se” (metastizar) pelo organismo, tornando o seu “combate” bastante mais complexo [Oncologia, 2010].

Na segunda metade do século XX, a taxa de incidência (número de novos casos, num dado período) do cancro da mama aumentou, sobretudo nos países desenvolvidos. Este tumor, apesar de incidir mais nas mulheres também pode ocorrer nos homens apesar de neste caso ser muito mais raro (1 homem para cada 100 mulheres com cancro da mama) [Oncologia, 2010].

Regra geral, o cancro da mama é um tumor de evolução lenta, podendo ser detectado, através de mamografia, em fases iniciais de desenvolvimento, antes mesmo de ser palpável. Se for detectado e diagnosticado numa fase inicial, apresenta uma taxa de cura superior a 90%. A figura 2.3 representa a evolução do número de células tumorais em função do tempo [Oncologia, 2010]. Os cancros da mama diferem entre si possuindo características específicas, bem como “comportamentos” e evoluções diferentes de pessoa para pessoa. Esta



**Figura 2.3:** Evolução do Cancro da Mama.

diferença de pessoa para pessoa e de cancro para cancro, tem motivado os cientistas à procura de um tratamento específico para cada caso, pois esta parece ser a abordagem mais promissora.

## 2.2 Tratamentos para o cancro da mama

A decisão acerca do tratamento a aplicar só pode ser tomada após uma cuidadosa verificação do perfil individual de cada paciente. O médico oncologista, ao escolher um determinado programa de tratamento tem em conta inúmeros factores, entre os quais a idade, o estado geral da doença, a possível localização das metástases e o tipo de células cancerosas envolvidas. Alguns exemplos de tratamento para o cancro da mama são:

**Quimioterapia:** Quimioterapia é a utilização de fármacos que agem na destruição das células malignas. Podem ser aplicados quer por via oral quer através de injeções intramusculares ou endovenosas [PortalSaúde, 2010]. Normalmente, administram-se vários ciclos de quimioterapia sendo que o período total de tratamento pode prolongar-se por vários meses. Os efeitos secundários mais frequentemente referenciados, que variam de gravidade conforme os fármacos usados, incluem: fadiga, náuseas, vómitos e perda de cabelo. Durante o período de tratamento fazem-se regularmente análises ao sangue para verificar se o número das células sanguíneas se reduziu. Se for necessário pode-se recorrer a transfusões ou tratamen-

tos médicos para repor o número de eritrócitos. Também há medicamentos disponíveis para melhorar o número de leucócitos e, portanto, aumentar a resistência do doente às infecções. Por fim, podem-se administrar outros tratamentos para prevenir a infecção [HER2, 2010].

**Terapêutica hormonal:** A terapêutica hormonal tem como finalidade impedir que as células malignas continuem a receber a hormona que estimula o seu crescimento. O tratamento pode incluir o uso de fármacos que modificam a forma de actuar das hormonas, ou a cirurgia, que remove os órgãos responsáveis pela produção dessas hormonas. Da mesma maneira que a quimioterapia, a terapia hormonal actua nas células do corpo todo e os seus efeitos secundários raramente são graves [PortalSaúde, 2010].

**Radioterapia:** A radioterapia consiste no tratamento do cancro por meio de raios-X ou de outras fontes de radioactividade. As fontes deste tipo produzem radiações ionizantes que, ao passar através do tecido doente, destroem ou abrandam o desenvolvimento das células anómalas. Contudo, a radioterapia pode ter efeitos secundários, tais como danos graves no tecido normal. A radioterapia é frequentemente utilizada em conjugação com outras formas de tratamento do cancro. No cancro da mama, a radioterapia é muitas vezes usada depois da remoção cirúrgica dum cancro da mama maligno, para destruir algumas células remanescentes do tumor. A radioterapia também pode ser usada para reduzir o tamanho de um tumor, ou para destruir células do cancro da mama que se tenham metastizado. A radioterapia, no entanto, só é usada quando os benefícios compensam os riscos de causar danos no tecido ok [HER2, 2010].

**Terapêutica com anticorpos monoclonais:** O aumento do conhecimento acerca dos genes humanos responsáveis pelo crescimento das células cancerosas, conduziu a uma nova fase no tratamento do cancro da mama. Uma nova abordagem dirigida ao tratamento do cancro da mama, envolve o uso de anticorpos monoclonais, concebidos para atingir células cancerosas específicas no organismo [HER2, 2010].

# Modelos de Prognóstico do Cancro da Mama

## 3.1 Introdução

Neste capítulo pretende-se apresentar a metodologia de suporte à construção de modelos de prognóstico utilizados no caso específico de cancro da mama. São apresentados 3 modelos que foram posteriormente integrados no sistema *web* de apoio à decisão, permitindo ao sistema fazer previsões individuais de sobrevivência para um paciente específico e efectuar a sua classificação em grupos de risco ou de gravidade. Um destes modelos é o *Nottingham Prognostic Index*, sendo que os outros modelos correspondem a recentes estudos nesta área.

Na formulação, desenvolvimento e validação de um modelo de prognóstico é necessário seguir alguns passos elementares. De um modo geral estes são: seleccionar as variáveis de resultado, seleccionar a população de pacientes a estudar, seleccionar as variáveis que são factores de risco, recolher e analisar os dados disponíveis, desenvolver o modelo estatístico e, por fim, validar o modelo. Para além de identificar a sobrevivência de um paciente através de um modelo de prognóstico, a estratificação dos pacientes em grupos de risco é uma componente clínica fundamental. Sendo assim, este capítulo apresenta também o algoritmo de estratificação utilizado.

Antes de caracterizar os 3 modelos de prognóstico este capítulo apresenta uma breve introdução aos conceitos da análise de sobrevivência.

## 3.2 Análise de sobrevivência

Nesta secção demonstra-se a utilidade da análise de sobrevivência utilizando casos práticos que exemplificam como se calculam, interpretam e comparam curvas de sobrevivência. São também definidos alguns conceitos básicos, tais como: evento, indivíduo censurado e curvas de *Kaplan-Meier*.

A análise de sobrevivência é uma das áreas da estatística que mais evoluiu nos últimos vinte anos do século passado. A razão deste crescimento deveu-se, sobretudo, ao desenvolvimento de novos métodos estatísticos e ao aparecimento de computadores cada vez mais rápidos. Uma evidência quantitativa deste crescimento, é o número de projectos e aplicações em análise de sobrevivência usados nas mais diversas áreas da medicina, nomeadamente no cancro da mama [Colosimo and Giolo, 2006].

Para efectuar uma análise de sobrevivência começa-se por definir o grupo de indivíduos em estudo (foram considerado para esta tese indivíduos com cancro da mama com as características definidas na secção 3.4). Em seguida, é necessário observar o grupo de indivíduos num período de estudo bem definido. Esta observação corresponde, geralmente, ao período de tempo durante o qual os indivíduos desenvolvem determinado evento. O nome “análise de sobrevivência” é enganador uma vez que, quando se refere tempo de sobrevivência não significa necessariamente tempo até à morte, mas sim tempo até ao evento em questão. Este evento é normalmente a morte, mas poderá ser qualquer outro evento: progressão da doença, efeito lateral ou alguma outra mudança do estado de saúde [Clark et al., 2003]. Nesta tese o evento considerado é a morte devida a qualquer causa.

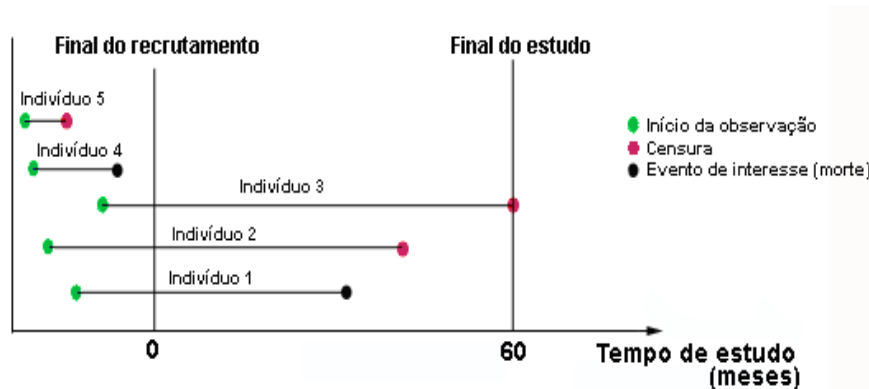
Em muitos estudos onde existem longos períodos de tempo (por exemplo, estudos de 5 ou 10 anos) muitos dos indivíduos não atingem o tempo total de seguimento previsto. Mesmo que não desenvolvam o evento até ao fim do estudo (independentemente do motivo pelo qual acontece ou do que poderá acontecer posteriormente) os dados destes participantes constam na análise final. Tal pode ocorrer porque ou abandonaram o estudo, ou seja, foram perdidos no seguimento (*follow-up*) ou o período de estudo chegou ao fim. Quando tal



acontece, o participante é designado como indivíduo censurado. Assim, censurado significa que o indivíduo não desenvolveu o evento até ao fim da sua observação no estudo. Parte-se do pressuposto que a censura dos participantes não está relacionada com o seu prognóstico.

A grande vantagem da análise de sobrevivência é permitir a utilização da informação de todos os participantes até ao momento em que desenvolvem, ou não, o evento de interesse. Assim, este tipo de técnica é ideal para analisar respostas binárias (ter ou não ter um evento) em estudos longitudinais que se caracterizam por tempos de seguimento diferentes entre os indivíduos e perdas de *follow-up* [Rebasa, 2005].

A figura 3.1 apresenta um exemplo de uma análise de sobrevivência, num grupo de 5 indivíduos num determinado período de estudo. O início de seguimento dos indivíduos em situação pós-operatória é assinalado por círculos a verde e a ocorrência do evento de interesse está representado por círculos a preto. A existência de censura durante o tempo de estudo é apresentada por círculos a encarnado.



**Figura 3.1:** Análise de sobrevivência num grupo de 5 indivíduos

Ao efectuar a análise da figura 3.1 constata-se que, para os indivíduos 1 e 4 são observados os tempos de sobrevivência devido ao aparecimento do evento de interesse. Em relação aos indivíduos 2, 3 e 5 existe censura. Ou seja, o indivíduo 5 é censurado porque abandonou o estudo antes do final de recrutamento de indivíduos. O indivíduo 3 é censurado porque não desenvolveu o evento de interesse (morte) até ao final dos 5 anos de estudo. O indivíduo 2 é censurado porque, por qualquer motivo, deixou de ser seguido antes do estudo ter terminado.

### 3.2.1 Curvas de *Kaplan-Meier*

No âmbito do estudo de sobrevivência existem diversos métodos de análise. As curvas de *Kaplan-Meier* são o método mais utilizado quando existem observações censuradas. Estas curvas de sobrevivência permitem comparar a rapidez com que os indivíduos desenvolvem determinado evento. Trata-se de um método não paramétrico, dado que a estimativa é realizada sem que se faça nenhuma suposição sobre a distribuição de probabilidade do tempo de sobrevivência. A variável dependente (*outcome*) é sempre o tempo até à ocorrência de determinado evento [Kaplan, 1958].

O método de *Kaplan-Meier* divide o tempo de seguimento em intervalos, cujos limites correspondem ao tempo de seguimento em que ocorreram eventos. Este é um dos métodos mais utilizado actualmente, pois utiliza a data exacta do evento, fornecendo, deste modo, resultados mais precisos.

Em seguida, exemplifica-se o cálculo de uma curva de *Kaplan-Meier*, considerando como evento a morte do paciente.

Sejam  $t_1 < t_2 < t_3 < \dots < t_r$  os tempos de morte distintos numa amostra de dimensão  $n$  em que ( $r \leq n$ ),  $d_i$  o número de mortes ocorridas em  $t_i$  e  $n_i$  o número de indivíduos em risco de morte em  $t_i$ . A função de sobrevivência  $S(t)$  determinada através do método de *Kaplan-Meier*, define-se da seguinte forma:

$$\hat{S}(t) = \prod_{i:t_i \leq t} \left( \frac{n_i - d_i}{n_i} \right) = \prod_{i:t_i \leq t} \left( 1 - \frac{d_i}{n_i} \right) \quad (3.1)$$

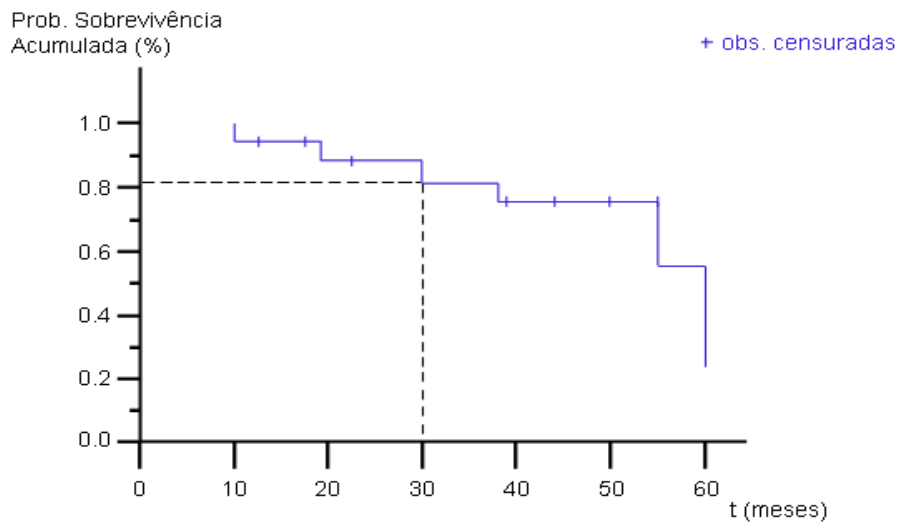
Na equação 3.1, a probabilidade de ocorrer um evento (morte) no intervalo desde  $t_{i-1}$  até  $t_i$  corresponde ao número de indivíduos mortos no tempo  $t_i$  sobre o número de indivíduos ainda vivos mais os indivíduos que morreram no tempo  $t_i$ . A fracção  $1 - \frac{d_i}{n_i}$  vai corresponder assim à probabilidade de sobreviver nesse mesmo intervalo  $t_i$ .

A tabela 3.1 pretende demonstrar um exemplo do cálculo da probabilidade de sobrevivência, num grupo de 18 pessoas, ao fim de 60 meses. Observando a tabela 3.1 verifica-se que ao fim de 60 meses a probabilidade de sobreviver é aproximadamente 22%. Com base

**Tabela 3.1:** Função de sobrevivência num período de 60 meses (método de *Kaplan-Meier*)

| $t_i$         | $n_i$               | $d_i$         | $c_i$              | $s(t_i)$                    | $\hat{S}(t_i)$                  |
|---------------|---------------------|---------------|--------------------|-----------------------------|---------------------------------|
| tempo (meses) | n.º de indiv. vivos | n.º de mortes | observ. censuradas | sobrev. no intervalo $t_i$  | sobrev. acumulada               |
| 0             | 18                  | 0             | 0                  | 1                           | 1                               |
| 10            | 18                  | 1             | 0                  | $1 - \frac{1}{18} = 0.9444$ | $1.0000 \times 0.9444 = 0.9444$ |
| 13            | 17                  | 0             | 1                  | $1 - \frac{0}{17} = 1.0000$ | $0.9444 \times 1.0000 = 0.9444$ |
| 18            | 16                  | 0             | 1                  | $1 - \frac{0}{16} = 1.0000$ | $0.9444 \times 1.0000 = 0.9444$ |
| 19            | 15                  | 1             | 0                  | $1 - \frac{1}{15} = 0.9333$ | $0.9444 \times 0.9333 = 0.8815$ |
| 23            | 14                  | 0             | 1                  | $1 - \frac{0}{14} = 1.0000$ | $0.8815 \times 1.0000 = 0.8815$ |
| 30            | 13                  | 1             | 0                  | $1 - \frac{1}{13} = 0.9230$ | $0.8815 \times 0.9230 = 0.8137$ |
| 36            | 12                  | 1             | 0                  | $1 - \frac{1}{12} = 0.9167$ | $0.8137 \times 0.9167 = 0.7459$ |
| 38            | 11                  | 0             | 1                  | $1 - \frac{0}{11} = 1.0000$ | $0.7459 \times 1.0000 = 0.7459$ |
| 44            | 10                  | 0             | 1                  | $1 - \frac{0}{10} = 1.0000$ | $0.7459 \times 1.0000 = 0.7459$ |
| 50            | 9                   | 0             | 1                  | $1 - \frac{0}{9} = 1.0000$  | $0.7459 \times 1.0000 = 0.7459$ |
| 55            | 8                   | 2             | 1                  | $1 - \frac{2}{8} = 0.7500$  | $0.7459 \times 0.7500 = 0.5594$ |
| 60            | 5                   | 3             | 0                  | $1 - \frac{3}{5} = 0.6000$  | $0.5594 \times 0.6000 = 0.2237$ |

nesta tabela é possível construir a curva de *Kaplan-Meier* ao fim de 60 meses (figura 3.2). Através da curva também se pode estimar directamente a sobrevivência ao fim de qualquer intervalo temporal. Neste caso, deve-se traçar uma linha vertical que passe no momento temporal em que se pretende apurar o nível de sobrevivência. Observa-se na curva que a sobrevivência ao fim de 30 meses é aproximadamente 81%. Da curva obtida é possível retirar algumas ilações, como por exemplo, a morte de 1 paciente significa um degrau decrescente e como consequência um intervalo com mais mortes irá originar um declive mais acentuado.



**Figura 3.2:** Curva de sobrevivência num período de 60 meses

O método de *Kaplan-Meier* é um método de análise que não tem em consideração as variáveis de prognóstico no momento exacto em que ocorreu o evento de interesse. No entanto, quando se pretende modelar a evolução de um determinado evento, mediante uma população torna-se essencial ter em consideração as variáveis que influenciam a sobrevivência. Para tal, é necessário definir novos modelos de prognóstico, bem como as variáveis determinantes para esses modelos.

## 3.3 Modelos de prognóstico

Os modelos de prognóstico do cancro da mama são baseados em princípios matemáticos estatísticos, que consistem na atribuição de um índice de risco ou gravidade (*hazard*) aos pacientes, sendo este normalmente apresentado sob a forma de uma probabilidade. Estes modelos partem, frequentemente, de características do paciente diagnosticadas pelo médico e consideradas importantes para o prognóstico. Neste estudo, seis variáveis fisiológicas foram consideradas importantes: idade no momento do diagnóstico, tamanho patológico (tamanho do tumor em centímetros), estado dos nódulos, estrogénio, tipo histológico e rácio entre nódulos afectados e nódulos removidos. A metodologia utilizada na escolha das variáveis para os modelos será explicada na secção 3.5.2.1. Estas características do paciente são posteriormente relacionadas através da construção de uma equação matemática, sendo o resultado obtido normalmente sob a forma de uma probabilidade de sobrevivência. Os resultados obtidos pelos diferentes modelos vão permitir depois criar grupos de pacientes estratificados pelo valor do índice de sobrevivência [Fernandes et al., 2010].

Para avaliar e comparar a qualidade dos modelos, estabelecendo-se relações entre os valores previstos e os valores observados, é necessário a sua validação através de diversas metodologias. Essas metodologias podem ser consultadas no artigo *Double-blind evaluation and benchmarking of survival models in a multi-centre study* [Taktak et al., 2007].

### 3.3.1 Princípios dos modelos de prognóstico

Segundo Altman e Wyatt qualquer modelo de prognóstico deve ser **credível** (*credibility*), **eficaz** (*effectiveness*) e **generalizável** (*generality*) [Altman and Wyatt, 1995].

Embora os melhores modelos de prognóstico tenham um forte suporte estatístico, os médicos são normalmente relutantes à sua utilização enquanto não acreditarem completamente nas suas previsões. Alguns pré-requisitos para aumentar a **credibilidade** clínica dos modelos de prognóstico são:

- todos os dados clínicos relevantes e referentes aos pacientes deverão ser testados a fim de serem integrados no modelo, nunca devendo ocultar-se dados importantes;
- a obtenção de dados deverá ser um processo simples, que não ocupe demasiado tempo e recursos (computorização de todo o modelo). Os dados deverão ser obtidos a partir de classes de doentes cujas previsões do modelo irão servir futuramente.

Os modelos deverão prever com elevado rigor a ocorrência de um evento para que a confiança da classe médica aumente e para que os modelos possam representar uma ajuda real aos profissionais de saúde.

Quanto à **eficácia** clínica, pode-se referir que os médicos deverão recorrer tanto ao seu senso clínico (ditado por vezes por muitos anos de prática e experiência médica), como a resultados ditados por ensaios clínicos e estudos de comparação. Torna-se de vital importância neste contexto validar as previsões com base em modelos baseados em ensaios clínicos bem delineados (máxima eliminação de factores de erro).

Alguns médicos defendem que os prognósticos baseados em modelos cujos dados provêm de uma determinada população de doentes não são generalizáveis, ou seja, não são aplicáveis a uma população distinta. De acordo com alguns especialistas, os modelos poderão ser **generalizáveis** desde que aplicados com segurança, testados de forma independente e separadamente, recorrendo a várias técnicas estatísticas.

## 3.4 Descrição e obtenção da base de dados utilizada

Tal como descrito na secção anterior, a obtenção dos dados deverá ser um processo simples devendo estes ser gerados a partir de classes de doentes cujas previsões do(s) modelo(s) irão servir futuramente.

Os modelos de prognóstico utilizados nesta tese têm em consideração um conjunto de dados composto por 931 pacientes do sexo feminino, obtido pelo *Christie Hospital* em Wilmslow (Manchester) no Reino Unido, no período entre 1990-1994. Estes dados dizem respeito a pacientes com cancro da mama em situação pós-operatória. Foi utilizado o método de filtragem TMN (Tumor, Nodes, Metastasis), que considera o tamanho do tumor  $< 5\text{cm}$ , *Node stage*  $< 2$  e onde não podem existir registos com sintomas clínicos de metástases. O período de estudo considerado é de 5 anos (60 meses) e o evento de interesse para análise é a morte devido a qualquer causa sendo que, o tempo de ocorrência do evento de interesse é medido em meses após a cirurgia. Os pacientes que após os 5 anos de acompanhamento sobreviverem são censurados.

O registo de cada paciente é constituído por 16 variáveis exploratórias, sendo algumas delas variáveis de prognóstico e outras variáveis com os resultados da aplicação de tratamentos. Do conjunto de dados, a categoria “*others*” da variável “*Histological Type*” foi removida uma vez que, os pacientes desta categoria possuem uma dinâmica de doença diferente dos restantes pacientes e como tal, estes valores podem influenciar negativamente o resultado final dos modelos de prognóstico. A variável “*Histological Type*” resume-se portanto a *lobular* e *ductal*. Dos 931 registos, 2 registos na categoria 4 da variável *Nodes involved* foram removidos por serem considerados *outliers*. Após as considerações descritas anteriormente, o conjunto de dados ficou reduzido a 743 pacientes. De notar que um dos pacientes entrou nas duas condições em simultâneo, tinha tanto *Histological type* igual a 3 como *Nodes involved* igual a 4. A inexistência de valores em alguns registos nas bases de dados médicas (*missing values*) é habitual nas áreas de saúde. Por vezes, após a realização de exames preliminares à mama, determina-se de imediato qual o tratamento a seguir, não sendo necessário a realização de

outros exames de despistagem, e como tal, algumas variáveis de prognóstico não são preenchidas. Os valores em falta associam-se por norma a outros dados do paciente, tendo esta associação o nome de *Missing at Random* (MAR). Normalmente, registos considerados como MAR estão associados a um bom prognóstico de cancro da mama. Para colmatar a inexistência de valores nos registos, estes foram substituídos por outros valores, usando uma metodologia de *Multiple Imputation*, considerada como eficiente [Clark and Altman, 2003], [Haybittle et al., 1982].

**Tabela 3.2:** Descrição das variáveis do conjunto de dados fornecidos pelo *Christie Hospital*

| Variáveis          |   | Frequência | %  |
|--------------------|---|------------|----|
| Age category       | 1 | 68         | 7  |
|                    | 2 | 445        | 48 |
|                    | 3 | 418        | 45 |
| Histological type  | 1 | 633        | 68 |
|                    | 2 | 111        | 12 |
|                    | 3 | 187        | 20 |
| Histological Grade | 1 | 116        | 13 |
|                    | 2 | 306        | 33 |
|                    | 3 | 198        | 21 |
|                    | 9 | 311        | 33 |
| Nodes Ratio        | 1 | 624        | 67 |
|                    | 2 | 28         | 3  |
|                    | 3 | 51         | 5  |
|                    | 4 | 43         | 5  |
|                    | 9 | 185        | 20 |
| Pathological Size  | 1 | 541        | 58 |
|                    | 2 | 390        | 42 |
| Oestrogen          | 1 | 155        | 17 |
|                    | 2 | 301        | 32 |
|                    | 9 | 475        | 51 |
| Menopausal status  | 1 | 237        | 25 |
|                    | 2 | 39         | 5  |
|                    | 3 | 655        | 70 |
|                    | 9 | 0          | 0  |
| Nodes involved     | 1 | 444        | 48 |
|                    | 2 | 203        | 22 |
|                    | 3 | 98         | 11 |
|                    | 4 | 2          | 0  |
|                    | 9 | 184        | 20 |
| Evento             |   | 122        | 13 |
| Total              |   | 931        |    |

O uso desta metodologia *Multiple Imputation* obriga à repetição do método 10 vezes sobre o conjunto de dados dos 731 pacientes e como tal, os resultados obtidos pelos modelos de

prognóstico têm em consideração a média dos 10 conjuntos de dados imputados. A tabela 3.2 apresenta o conjunto de dados do *Christie Hospital* [Fernandes et al., 2008]. Através da análise da tabela 3.2 é possível verificar a frequência das variáveis para as diferentes categorias nos dados. A categoria ‘9’ por exemplo, indica, para a variável correspondente, o número de registos que se encontram vazios. No caso concreto da variável *Oestrogen*, 475 pacientes (51%) dos pacientes tinham este campo vazio.

## 3.5 Modelos prognóstico implementados

Uma forma de tentar encontrar os melhores tratamentos para os pacientes com cancro da mama é colocá-los em grupos de risco de prognóstico. Para tal, é necessário um modelo quantitativo de sobrevivência que está, geralmente, focado nos primeiros 5 anos após a cirurgia. É também crucial e de extrema importância interpretar os grupos de prognóstico criados, tendo em conta as variáveis clinicamente relevantes [Lisboa et al., 2003]. Na prática, os médicos oncologistas utilizam frequentemente o algoritmo conhecido por *Nottingham Prognostic Index* (NPI) [Galea et al., 1992], que foi derivado utilizando o método estatístico clássico Proportional Hazards [Cox, 1972].

### 3.5.1 *Nottingham Prognostic Index*

Historicamente, o propósito dos modelos de prognóstico é o de estratificar os pacientes em grupos de risco com diferentes probabilidades de sobrevivência. Um dos índices mais usado é o TMN. Este índice é puramente clínico, já que depende apenas de exames clínicos e de apalpação. Tendo em conta a propagação metastática da doença, o TMN não é suficientemente detalhado para o cancro da mama numa fase inicial. Uma vez que este índice não é o melhor método para separar os pacientes em grupos de risco, surge a necessidade de recorrer ao índice NPI [Fernandes et al., 2010] que foi criado em 1982 por Marcus H. Galea [Galea et al., 1992]. O NPI é muito usado para pacientes com cancro da mama numa fase inicial, após a realização da cirurgia ao cancro da mama. Este índice é muito utili-



zado para informar a escolha da terapia adjuvante, além de ser um indicador da previsão de sobrevivência do paciente. Essa previsão é calculada usando a seguinte fórmula:

$$NPI = 0.2 \times \textit{pathological size} + \textit{histological grade} + \textit{nodes involved}. \quad (3.2)$$

A partir da fórmula 3.2, os pacientes NPI são colocados em 4 diferentes grupos de prognóstico [Galea et al., 1992] e [Haybittle et al., 1982] da seguinte forma:

Grupo 1: Índice NPI  $< 2.41$   $\rightarrow$  excelente prognóstico

Grupo 2: Índice NPI  $\geq 2.41$  e Índice NPI  $< 3.41$   $\rightarrow$  bom prognóstico

Grupo 3: Índice NPI  $\geq 3.41$  e Índice NPI  $< 5.41$   $\rightarrow$  prognóstico moderado

Grupo 4: Índice NPI  $\geq 5.41$   $\rightarrow$  mau prognóstico

Consideram-se apenas 3 factores de prognóstico, pois ao se considerar mais do que 3 factores o número de possíveis subgrupos torna-se maior e, conseqüentemente, o número de pacientes por grupo diminui [Myers et al., 1966].

Em estudos de Galea [Galea et al., 1992] e Haybittle [Haybittle et al., 1982] são realizadas análises pormenorizadas da influência que cada variável tem para o cálculo do índice de prognóstico. Uma vez que o índice NPI foi derivado utilizando o método estatístico clássico *Proportional Hazards* (que tem em conta toda a informação acerca de um grupo de pacientes - factores de prognóstico), este permite que cada variável de prognóstico seja avaliada de forma independente em termos de significância, tendo em conta o efeito das outras variáveis. Verificou-se que:

- A variável *Lymph node (nodes involved)* é a que tem maior impacto na sobrevivência de um indivíduo. Pelo que, tem sido considerada o maior indicador de prognóstico. Esta variável tem sido o único indicador tido em conta por alguns centros clínicos na escolha da terapia a ser efectuada aos pacientes de cancro da mama. Se o paciente tiver cerca de 1-3 *nodes involved* tem 73% probabilidade de sobreviver em 5 anos, enquanto

que se tiver cerca de 4-12 *nodes involved*, a probabilidade de sobreviver diminui para 43% probabilidade [Galea et al., 1992] e [Haybittle et al., 1982].

- Considerando a variável relacionada com o tamanho do tumor - *pathological size* - os pacientes com tumor de 2 cm têm 65% de hipóteses de sobreviver ao fim de 10 anos comparando com 24% de sobrevivência caso o tumor tenha o tamanho entre 4-5 cm [Galea et al., 1992] e [Haybittle et al., 1982].

### 3.5.2 *Cox Proportional Hazards Modelling*

No campo da análise de sobrevivência, o modelo *Proportional Hazards* ou Regressão de Cox Linear é bastante utilizado. É um modelo semi-paramétrico de regressão linear múltipla da função de risco que assume que toda a dependência temporal é exclusiva a um grupo de pacientes base ou população base  $h_0(t)$ . Para intervalos de tempo discretos, o modelo *Proportional Hazards* parametriza a probabilidade de sobrevivência na proporção de uma população base (*baseline*) [Collect, 1994]. A regressão de Cox permite relacionar a função *hazard* e um determinado número de covariáveis significativas onde o risco (*hazard*) para um paciente é calculado consoante o valor das covariáveis  $x_p$  no tempo  $t_k$ , como indica a equação 3.3, considerando o tempo discreto:

$$\frac{h(x_p, t_k)}{1 - h(x_p, t_k)} = \frac{h_0(t_k)}{1 - h_0(t_k)} \cdot \exp \left( \sum_{i=1}^{N_i} bx_i \right) \quad (3.3)$$

O risco *Hazard Ratio* (HR) tem um significado semelhante ao risco relativo. *Hazard* é a probabilidade de um paciente que não experimentou o evento até um determinado momento, experienciá-lo nesse momento. No limite, o modelo em tempo discreto converge para a parametrização conhecida do *Proportional Hazard*:

$$h(x_p, t_k) = h_0(t_k) \cdot \exp \left( \sum_{i=1}^{N_i} bx_i \right) \quad (3.4)$$

Onde  $h_0$  é a função de risco base. Não existe nenhum procedimento determinístico para seleccionar quais os valores dos atributos a serem preenchidos na população base, mas seguiu-se a prática *standard* de seleccionar os atributos dos registos a partir dos quais se espera obter uma maior sobrevivência [Ravdin et al., 1992] [Christensen, 1987].

Na equação 3.3 a incógnita  $x_i$  corresponde às variáveis do paciente e o índice de prognóstico (PI) é calculado através da expressão  $\beta x$  que é igual a  $(\sum_{i=1}^{N_i} bx_i)$ . Para um paciente com determinadas características de prognóstico o cálculo do índice é dado por:

$$\text{CoxPrognosticIndex} = \text{pathological size} \times \text{beta1} + \text{histological grade} \times \text{beta2} + \text{nodes ratio} \times \text{beta3} + \text{oestrogen} \times \text{beta4} + \text{age} \times \text{beta5} + \text{histological grade} \times \text{beta6}.$$

Este índice é de extrema importância, pois vai permitir definir a colocação dos pacientes em grupos de prognóstico, através do método de estratificação em grupos de risco, que será caracterizado posteriormente, na secção 3.6.

Na descrição dos dados na secção 3.4, os 10 conjuntos de dados para treino foram considerados. Como tal, cada conjunto de treino serviu de base para a criação de um modelo, obtendo-se assim um  $\beta$  para cada categoria de cada variável. O resultado do prognóstico final obtido para o modelo do *Cox Proportionals Hazards* para cada paciente é assim determinado pela média dos 10 índices obtidos por cada conjunto de treino.

### 3.5.2.1 Selecção das variáveis de prognóstico

Em muitos estudos é-se confrontado com problemas em que o objectivo principal é analisar a influência que uma ou mais variáveis (factor de prognóstico) têm sobre a mortalidade ou sobre uma doença qualquer. Como tal, é necessário ajustar os factores de prognóstico já conhecidos para determinar o seu efeito independente dos outros, tal como é necessário calcular a sua significância para efeitos de morte ou doença. De um modo geral, a resolução do problema é feita através do estudo realizado a partir de um modelo de regressão.

Cox formulou uma função de verosimilhança parcial que não depende de  $h_0(t)$ , e que permite a realização de inferências sobre  $\beta$  quando  $h_0(t)$  é desconhecido. Partindo desse pressuposto, uma escolha popular para o cálculo das variáveis mais significativas consiste

### 3.5. MODELOS PROGNÓSTICO IMPLEMENTADOS

---

no modelo de selecção de Regressão de Cox Stepwise Forward sobre os conjuntos de dados de treino imputados [Fernandes et al., 2008]. Em estatística, o método Stepwise Forward considera o seguinte: calcula-se primeiro em cada iteração quais as variáveis já incluídas no modelo que podem ser removidas; se nenhuma variável puder ser removida, determina-se se uma nova variável ainda não adicionada pode entrar. Assim, uma variável pode entrar no modelo numa iteração e ser removida na iteração seguinte.

As variáveis utilizadas para o desenvolvimento dos métodos de prognóstico foram seleccionadas utilizando o método de bootstrap e o método selecção de Regressão de Cox Stepwise Forward. Assim, as seis variáveis de prognóstico foram consideradas como as mais significativas para o cálculo da sobrevivência do cancro da mama e codificadas de 1-até-N: *Pathological size*, *Nodes involved*, *Histological grade*, *Histological type*, *ER status* e *Age category*. As primeiras 3 variáveis reflectem o já conhecido *Nottingham Prognostic Index* (NPI). De notar que na construção do modelo apenas alguns factores de prognóstico após a cirurgia são incluídos para efeitos de sobrevivência e risco, sendo excluídas as variáveis relacionadas com os tratamentos para o cancro da mama.

**Tabela 3.3:** Parâmetros do modelo de Regressão de Cox

| Variáveis                 |   | Categorias                | Betas     |
|---------------------------|---|---------------------------|-----------|
| <b>Age category</b>       | 1 | 20-39                     | 0.617202  |
|                           | 2 | 40-59                     | 0         |
|                           | 3 | 60+                       | 0.461044  |
| <b>Histological type</b>  | 1 | Invasive ductal           | 0         |
|                           | 2 | Invasive lobular          | -0.779706 |
| <b>Histological Grade</b> | 1 | Well differentiated       | -0.494438 |
|                           | 2 | Moderately differentiated | 0         |
|                           | 3 | Poorly differentiated     | 0.516832  |
| <b>Nodes Ratio</b>        | 1 | 0-20%                     | 0         |
|                           | 2 | 20-40%                    | 1.221011  |
|                           | 3 | 40-60%                    | 0.685276  |
|                           | 4 | 60+%                      | 1.652103  |
| <b>Pathological Size</b>  | 1 | < 2 cm                    | 0         |
|                           | 2 | 2-5 cm                    | 0.613669  |
| <b>Oestrogen</b>          | 1 | 0-10                      | 0.607958  |
|                           | 2 | 10+                       | 0         |

A tabela 3.3, mostra os coeficientes *betas* calculados para as categorias de cada variável de prognóstico, após efectuada a média dos  $\beta$  obtidos para as diferentes bases de dados imputados. A escolha do valor de *baseline* (zero) em alguns coeficientes está associado à maior frequência com que estes aparecem no conjunto de dados (*Christie Hospital*). Os coeficientes produzidos pela análise de cada  $\beta$ , mostram como cada factor contribui para o risco, que é inverso à sobrevivência.

### 3.5.3 PLANN-ARD Modelling

O *Partial Logistic Artificial Neural Network* (PLANN) proposto por Biganzoli e seus colaboradores, é um modelo não linear, semi-paramétrico, que se baseia em técnicas de inteligência artificial como as redes neuronais (ANNs) [Biganzoli et al., 1998]. O PLANN é um modelo que permite fazer previsões sobre o risco  $h(x_p, t_t)$  de um determinado paciente para cada tempo  $t$ . Este modelo foi comparado com alguns modelos lineares clássicos como o modelo de Regressão de Cox e demonstrou-se que igualou ou superou os modelos lineares para a análise de sobrevivência na presença de censura [Taktak et al., 2007]. Embora permita retirar ilações sobre a relação entre as covariáveis, não é visto como um modelo exploratório como o *Proportional Hazards Regression*.

Neste modelo é incluída uma estrutura de regularização, o método *Automatic Relevance Determination*, que permite assegurar a generalidade das previsões do modelo para novos dados, isto porque os modelos não-lineares como as redes neuronais podem levar ao *overfitting* dos dados. Daí o acrónimo PLANN-ARD [Lisboa et al., 2003].

Para o cálculo do risco, o modelo é composto por uma estrutura de um perceptrão em multi-camada, com 1 camada escondida, uma função de activação não-linear sigmoideal nos nós da camada de entrada e da camada escondida. As estimativas dos pesos da rede neuronal são feitas através de um termo de verosimilhança que tem em conta o estado do paciente num determinado tempo  $t_k$ , usando um indicador se o paciente está vivo no tempo  $t_k$  ( $l = 0$ ) ou morto ( $l = 1$ ).

Na entrada da rede neuronal são introduzidos como *inputs* o tempo discreto (em meses) e o valor das covariáveis. A saída da rede vai ter como *output* o risco  $h(x_p, t_t)$  calculado para determinado paciente em cada intervalo de tempo definido. O cálculo do risco para um paciente é dado por:

$$\frac{h(x_p, t_k)}{1 - h(x_p, t_k)} = \exp \left( \sum_{h=1}^{N_h} w_h g \left( \sum_{i=1}^{N_i} w_{ih} x_{pi} + w t_k + b_h \right) + b \right) \quad (3.5)$$

A estimativa de sobrevivência é dada por:

$$S(t_k) = \prod_{l=1}^k (1 - h(t_l)) \quad (3.6)$$

O cálculo do índice de prognóstico no modelo PLANN-ARD é dado pela equação:

$$PI(x_p) = (-\ln(1 - CCI(t))) = \ln(-\ln(S(t))) \quad (3.7)$$

A variável *CCI* representa a *crude cumulative incidence* e identifica a probabilidade da ocorrência de um determinado evento de interesse [Fernandes et al., 2010]. A variável  $S(t)$  corresponde à estimativa de sobrevivência para um paciente no fim do período de estudo, ou seja, ao fim de 5 anos.

Este índice, tal como o índice do modelo de Cox, vai permitir calcular o respectivo grupo de risco para um paciente. Como consequência do método de imputação referido na obtenção e descrição dos dados em 3.4, foram criadas 10 redes neuronais usando os 10 conjuntos de dados imputados para treino resultando assim em 10 diferentes índices de prognóstico para cada paciente. O índice final será dado pela média dos 10 índices obtidos.

O cálculo do prognóstico de sobrevivência, para um paciente com os respectivos intervalos de confiança é obtido usando as metodologias PLANN-ARD e de Monte Carlo descritas no artigo *An integrated framework for risk profiling of breast cancer patients following surgery* [Jarman et al., 2008].

A metodologia PLANN-ARD descrita anteriormente, fornece para cada tempo  $t_k$  uma

estimativa de risco. Como tal, para cada tempo  $t_k$  os diferentes pesos da rede neuronal do modelo vão ter uma distribuição Gaussiana  $N(a_i, \sigma_i^2)$  [Bishop, 1995]. Com os respectivos pesos, a sobrevivência calcula-se multiplicando a taxa de risco para cada intervalo discreto  $t_k$  como indica a equação 3.6.

Após o cálculo da sobrevivência para cada intervalo  $t_k$ , a sobrevivência para um paciente é obtido pelo método de Monte Carlo. O método selecciona aleatoriamente uma amostra  $\tilde{a}_i$  do conjunto  $N(a_i, \sigma_i^2)$  e calcula respectivamente o risco:  $\tilde{h}_i = g(\tilde{a}_i)$  e a sobrevivência:  $\tilde{S}(t_k)$ . O passo anterior é repetido  $n$  vezes, para construir uma distribuição. Da distribuição retiram-se os intervalos de confiança e a média de sobrevivência para um paciente.

### 3.6 Estratificação em grupos de risco

Na área da saúde é muito frequente o uso de modelos de sobrevivência para estratificar os pacientes em grupos de risco. Esta divisão, além de permitir especular a influência que certos tratamentos têm, permite também avaliar o impacto que certos factores de prognóstico têm na sobrevivência dos pacientes com cancro da mama. Como consequência, uma metodologia de estratificação foi incorporada no *web site* no sentido de separar os 743 pacientes em diferentes grupos de riscos. Existem várias metodologias de estratificação em grupos de risco [Fernandes et al., 2009]. Uma dessas metodologias é o algoritmo de árvores de decisão CART - “Classification And Regression Tree” proposto por Breiman em 1984 [Breiman et al., 1984]. Este algoritmo foi utilizado nesta tese para a estratificação dos índices de prognóstico em grupos de risco e na obtenção de regras booleanas associadas a esses grupos. Neste algoritmo, os valores dos índices de prognóstico (PI) (obtido nos modelos Cox Proportional Hazards e PLANN-ARD) e as variáveis de prognóstico (*Age, Pathological Size, Histological Type, Histological Grade, Nodes Ratio e Oestrogen*) são utilizados para o cálculo dos grupos de risco. A metodologia do modelo CART é tecnicamente conhecida como partição recursiva binária. O processo é binário porque os nós pais são sempre divididos exactamente em dois nós filhos sendo recursivo porque o processo pode ser repetido

tratando cada nó filho como um novo nó pai. As principais etapas do CART consistem em: definir o conjunto de regras para dividir cada nó da árvore, decidir quando a árvore está completa e associar cada nó terminal a uma classe ou a um valor preditivo no caso da regressão. Após encontrar a árvore completa, efectua-se normalmente a “poda” de alguns ramos da mesma, de modo a aumentar o seu poder de generalização e diminuir o conhecido problema de *overfitting* dos modelos. As sub-árvores obtidas através da “poda” da árvore inicial são examinadas testando taxas de erros sendo a melhor delas a escolhida.

O algoritmo CART foi aplicado nos modelos de sobrevivência ao modelo PLANN-ARD, à rede neuronal para eventos e ao modelo Cox Proportional Hazards. Com isto é possível atribuir a cada paciente um grupo de risco. As figuras nos apêndices C e D representam as árvores de decisão para os dois modelos de prognóstico onde 4 grupos de risco foram considerados. As regras extraídas pelo algoritmo CART para os dois modelos encontram-se nos apêndices A e B.



# WEBSITE - *Breast Cancer Survival*, Implementação

## 4.1 Introdução

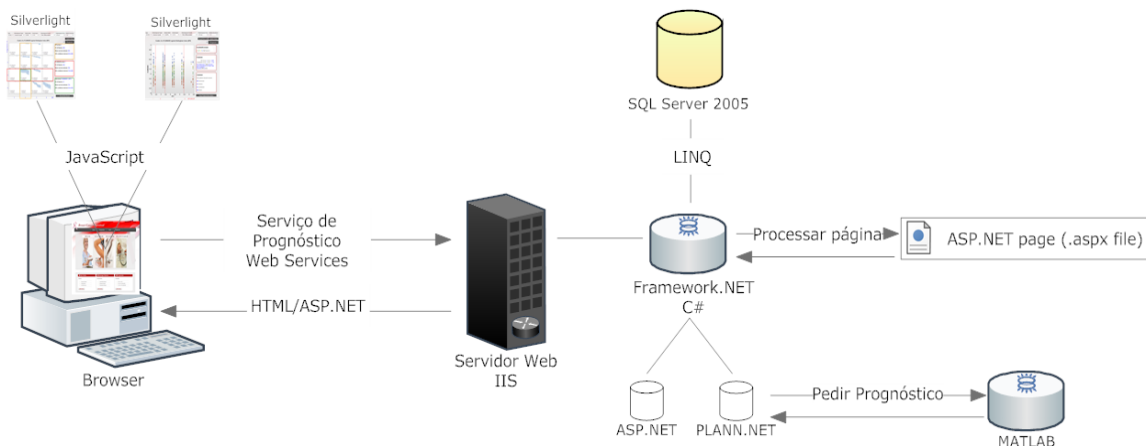
Este capítulo está estruturado em duas partes distintas. Inicialmente faz-se uma breve referência às tecnologias implementadas na construção do *BCSurvival*. Posteriormente, será descrito em pormenor o processo de criação do *site* tendo em conta essas tecnologias, os desafios e as complexidades enfrentadas na sua implementação, bem como a sua resolução até o *site* ser colocado *online*. Pretende-se também demonstrar como as metodologias utilizadas no prognóstico de cancro da mama caracterizadas anteriormente podem ser incorporadas num *site*. A plataforma *web* encontra-se disponível no endereço <http://bcsurvival.pt.1a/>.

## 4.2 Tecnologias utilizadas pelo Site - *BCSurvival*

### 4.2.1 Estrutura Geral do Site

A figura 4.1 ilustra o diagrama do funcionamento do sistema de decisão *BCSurvival*.

A *home page* do sistema permite mostrar os contributos a que o *site* se propõe, isto é, identificá-lo como um sistema *online* de apoio à decisão para *Breast Cancer Survival*, em



**Figura 4.1:** Diagrama do *site - BCSurvival*

que o acesso às funcionalidades do *site* fica dependente de um registo gratuito. Depois de registado, o utilizador passa a estar identificado pela aplicação *web* e poderá usufruir das ferramentas de prognóstico e de todas as funcionalidades existentes.

O sistema de apoio à decisão médica utiliza diversas metodologias de prognóstico de modo a ajudar os médicos oncologistas e todos os utilizadores interessados a efectuarem inferências de prognóstico do cancro da mama, usando diversas características de um paciente. Essas características, essenciais para o prognóstico do paciente, são colocadas pelo utilizador no *web browser*. Por sua vez, o *browser* processa os dados e coloca-os numa mensagem SOAP, encapsulada num pacote HTTP (*Web Services*). Essa mensagem é enviada para o servidor IIS, onde se encontra a plataforma ASP.NET desenvolvida nesta dissertação e as diferentes metodologias para a inferência do prognóstico, sendo uma delas executada em Matlab (PLANN-ARD).

A plataforma ASP.NET tem incorporado três ferramentas diferentes de prognóstico: o esquema de classificação NPI, o modelo de Cox e o PLANN-ARD, onde o utilizador pode escolher entre as várias categorias para cada uma das seguintes variáveis: *Age*, *Histological Type*, *Node Ratio*, *Oestrogen*, *Histological Grade*, *Pathological Size*, *Nodes involved*.

Os índices de risco das metodologias Cox e PLANN-ARD juntamente com as variáveis vão ser utilizadas para a estratificação dos pacientes em grupos de risco. Esta divisão permite avaliar melhor o impacto que certos factores de prognóstico têm na sobrevivência

dos pacientes com cancro da mama. Uma metodologia de estratificação (árvores de decisão - CART), foi incorporada no *site* no sentido de separar os 743 pacientes em diferentes grupos de risco. O resultado dessa separação pode ser vista no gráfico 2D em Silverlight.

Esta aplicação *web* permite também utilizar e guardar a informação referente aos pacientes (dados e prognósticos efectuados). Deste modo pode-se obter um historial de cada paciente, bem como todas as avaliações efectuadas. O acesso aos dados do sistema mantidos no servidor SQL Server 2005 é realizada através da tecnologia LINQ.

De forma a tornar a aplicação mais interactiva foi implementada uma interface 2D em Silverlight, que corre no lado do utilizador para a localização dos diferentes grupos de risco e índices de prognóstico, em que os diferentes modelos de prognóstico são gerados computacionalmente no servidor e o resultado desses modelos é devolvido para o *web browser* em HTML/ASP.NET através de *Web Services* e JavaScript.

As secções seguintes explicam a necessidade da implementação das diferentes tecnologias introduzidas na aplicação *web* desenvolvida nesta dissertação em ASP.NET.

### 4.2.2 Site

Um *web site* ou *site* é um conjunto de páginas *web* no formato de hipertexto, em que a página inicial é designada de *Home Page* [TechTarget, 2010]. O sistema de hipertexto mais conhecido actualmente é a *Word Wide Web* (www). No *browser* é possível aceder-se a estas páginas *web* através do protocolo HTTP (*Hypertext Transfer Protocol*) ou protocolo de transferência de hipertexto.

As páginas *web* no *site* estão normalmente organizadas em pastas, respeitando uma hierarquia que pode ser vista no URL do *browser*, sendo possível através deste URL ter uma percepção da estrutura global do *web site*. Por exemplo, uma empresa ou uma pessoa indica o nome da página *web* dando o endereço URL da *Home Page*. A partir desta página inicial, pode-se obter todas as outras páginas, quer através do URL, quer navegando pelo *web site*.

Muitas vezes, um *site* é confundido com um servidor *web*. Um servidor normalmente é

um computador que contém os ficheiros e páginas *web* para um ou mais sites. É de salientar que, no caso do *site* ser grande, este pode ser dividido por diversos servidores espalhados, se for caso disso, em diferentes partes do mundo. Resumindo, um *site* pode ser visto como um conjunto de páginas interligadas, contendo informações, ficheiros e documentos de uma entidade (pessoa, empresa, instituição etc). Estas páginas estão associadas a um domínio (endereço URL) dentro da Internet, em que o acesso é feito através de um *web browser*.

### 4.2.3 HTML

Um *site* é composto por páginas HTML (HyperText Markup Language). O HTML é uma linguagem destinada à elaboração de páginas web com hipertexto. Essas páginas são posteriormente interpretadas pelos *browsers*. O HTML é uma linguagem de formatação em que uma página HTML é um ficheiro de texto que é formatado através de comandos (*tags*).

A linguagem HTML foi originalmente construída por Tim Berners-Lee, durante o seu trabalho no CERN, tendo sido conhecida através do *browser* Mosaic, desenvolvido na NCSA. No decorrer dos anos 90, o código HTML expandiu-se de várias maneiras, a *web* cresceu, sendo que o seu sucesso ficou a dever-se aos autores de sites e empresas por usarem as mesmas convenções HTML. Isto levou posteriormente a que se realizasse um trabalho conjunto na criação das especificações HTML [W3C, 2010].

No final de 1994, a *Internet Engineering Task Force* (IETF) desenvolveu o código HTML 2.0, com o intuito de padronizar as características principais do HTML. Em 1995, o JavaScript foi criado como uma extensão do HTML para o *browser* Navigator v2.0. Em 1996, CSS1 é apresentada como a primeira folha de estilo, criada para complementar a linguagem HTML. Desde a publicação do HTML que o grupo de trabalho da W3C tem desenvolvido o XHTML que é uma especificação HTML baseada em XML, considerada pela W3C como um sucessor do HTML. O XHTML faz uso de uma sintaxe mais rigorosa e menos ambígua que se torna mais simples de ser processada. No início de 2008, o W3C anunciou a primeira especificação do HTML 5, que inclui APIs para o desenvolvimento de gráficos bidimensio-

nais.

Em geral, há necessidade dos documentos HTML funcionarem de uma forma equivalente nas diferentes plataformas e *browsers*. Conseguindo esta interoperabilidade reduzem-se custos, como por exemplo, o facto de as pessoas que desenvolvem páginas não necessitarem de criar vários tipos de documentos *web* consoante o *browser* a que se destinam. Se não se efectuar um esforço nesse sentido existe o risco de na *web* se desenvolverem diversos formatos incompatíveis, reduzindo assim o seu potencial comercial para todos os utilizadores. Cada versão do HTML tem procurado levar em consideração os pedidos feitos pela indústria de *software*, para que o investimento efectuado pelos autores das páginas *web* não seja desperdiçado, nem os documentos se tornem inválidos num curto período de tempo.

A estrutura de um documento HTML apresenta as seguintes componentes principais:

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01//EN" "http://www.w3.org/TR/...">
<html> <head> <title>Título do Documento</title> </head>
  <body> Hello World </body>
</html>
```

As etiquetas básicas de HTML, cuja presença é altamente recomendada nas páginas são:

- `< html >`: define o início de um documento HTML e indica ao *browser* que todo conteúdo posterior deve ser tratado como uma série de códigos HTML.
- `< head >`: define o cabeçalho de um documento HTML que contém informações sobre o documento que está aberto.
- `< body >`: define o conteúdo principal, o corpo do documento. No corpo podem-se definir propriedades comuns a toda a página, como a cor de fundo, as margens e outras formatações. Esta é a parte do documento HTML exibida no *browser*.

Para saber mais sobre o HTML consultar por exemplo o artigo da W3C em [W3C, 2010].

### 4.2.4 CSS

As CSS ou folhas de estilo em cascata são um mecanismo simples para adicionar estilos (por exemplo: fontes, cores, espaços) aos documentos *web*. Têm por finalidade devolver ao HTML o propósito inicial da linguagem, ou seja, o HTML foi inicialmente criado para ser uma linguagem de marcação e de conteúdos. No entanto, com a sua popularidade, começou a surgir uma mistura no HTML tanto de conteúdos como de *tags* associadas ao aspecto de um *site*, tornando-se o código cada vez mais ilegível e a manutenção de um *site* cada vez mais difícil. Assim, cabem às CSS todas as funções de apresentação de um documento *web* [Silva, 2007]. Daí o aparecimento da frase acerca do CSS + HTML:

“HTML para estruturar e CSS para apresentar”

As CSS vieram proporcionar “limpeza” no código HTML, ao mesmo tempo que permitem a gestores de projectos de sites a vantagem de separar um *site* em vários níveis havendo uma melhor distribuição de competências, podendo por exemplo os programadores ficar com a gestão de software a ser implementado no *site* e os *web designers* com a respectiva imagem do *site* [Silva, 2007]. O exemplo seguinte indica uma simples formatação CSS num código HTML, onde vai ser indicado ao *browser* que qualquer conteúdo na página *web* associado a parágrafos irá aparecer uma cor de letra a preto, em itálico e com um fundo branco.

HTML:

```
< p > Um parágrafo com texto < /p >
```

CSS:

```
p {color: #000; background-color: #FF0; font-style: italic;}
```

### 4.2.5 ASP.NET

ASP.NET é uma tecnologia gratuita da Microsoft para o desenvolvimento de aplicações *web*, sendo a sucessora da tecnologia ASP.

No início da Internet, as tecnologias para *design* e desenvolvimento de *web sites* eram inexistentes. As páginas estáticas de HTML criadas pelos utilizadores eram usadas para troca

de informação dirigida a um determinado público-alvo. Actualmente, porém, é importante haver sites dinâmicos em que haja bastante interacção. O crescimento do *e-commerce*, ou do *e-bussiness*, levou à criação de bases de dados para a integração de conteúdo nos sites. Para atender a esta necessidade, a Microsoft criou inicialmente uma linguagem de desenvolvimento para *web*, o ASP [Joe and Tomson, 2002] e mais tarde o ASP.NET.

O ASP embora fosse uma linguagem acessível para se criar páginas *web* era limitada em alguns aspectos, por exemplo, a simples validação de dados, implicava uma codificação excessiva, tanto no lado do cliente como no lado do servidor. O resultado desta mistura originou sérios problemas de manutenção de sites em ASP. Para ultrapassar estes e outros problemas a Microsoft introduziu o ASP.NET.

O ASP.NET para quem desenvolve páginas *web*, fica mais fácil não somente na criação de sites simples, mas, também, na criação de sites com alto nível de complexidade. É um componente do IIS que permite, através de uma linguagem de programação integrada na plataforma .NET, criar páginas *web* dinâmicas ou estáticas. Para quem desenvolve em .NET (VB, C#, JScript. NET e J#) devido ao CLR tem tudo para criar uma página ASP.NET. O CLR providencia um ambiente de máquina de aplicação virtual que, garante ao programador não se preocupar com as características específicas do hardware onde a aplicação *web* vai ser executada [ASP.NET, 2010].

A escolha da plataforma ASP.NET para a construção do *site BCSurvival* desenvolvido nesta dissertação foi devido a diversos factores:

- A programação em ASP.NET ser orientada a objectos.
- Os elementos de uma página HTML passam a ser componentes que podem ser controlados: *Web controls*, *HTML controls*, *User controls* (para reaproveitar código), *Custom controls*, *Calendar*, *DataGrid*, *DataList*, *Repeater* etc.
- O facto de herdar todas as características oferecidas pela plataforma .NET (tais como conjunto de classes bastante completas e com um conjunto de operações de enorme utilidade).

- O ambiente para desenvolvimento Visual Studio 2008 é bastante simples e intuitivo.
- A fácil integração com o servidor SQL Server 2005, também este desenvolvido pela Microsoft.
- A experiência na linguagem de programação, nomeadamente, C#.

A Microsoft, ao criar esta nova tecnologia, teve como principal missão diminuir a barreira no desenvolvimento de sites. Esta desenvolveu um modelo de programação “orientado a eventos” no qual um programador adiciona, por exemplo, um evento “onbuttonClick” a um formulário e no lado do servidor escreve o código para controlar esse evento. Para escrever uma página ASP.NET começa-se por escolher uma linguagem compatível com .NET. Optou-se na realização desta tese pela linguagem C#. O ambiente de trabalho escolhido para desenvolver o *web site* foi o Visual Studio 2008.

Em ASP.NET existem dois modelos de programação: *Code Behind* ou *Server Code* e o *Code in Page*.

**Code Behind** - Neste modelo existe uma separação entre o código HTML e código C#. Para cada ficheiro aspx, existe um ficheiro aspx.cs onde será feito o código C#. O ficheiro aspx, contém o conteúdo HTML e o conteúdo dos componentes do ASP.NET (< *asp* :). O recurso ao código nesse modelo é muito grande, facilitando a vida ao programador.

**Code in Page** - Este modelo é semelhante ao modelo do ASP clássico. A diferença é que existem outras *tags* que não existem no ASP clássico. As *tags* permitem fazer herança, implementar uma interface, importar outras classes. Aqui o código fica um pouco menos nítido devido à mistura de código HTML com código C#, JavaScript, CSS, etc.

O código seguinte mostra um pequeno exemplo dos dois modelos de programação descritos anteriormente:

**Exemplo1: Código HTML do ficheiro aspx (Code in Page):**

```
<%@Page Language="C#" AutoEventWireup="true" CodeBehind="Default.aspx.cs"
    Inherits="Example._Default"%>
```



```
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Transitional//EN"
    "http://www.w3.org/TR/xhtml1/DTD/xhtml1-transitional.dtd">
<html xmlns="http://www.w3.org/1999/xhtml"> <head runat="server"></head>
<body>
    <form id="form1" runat="server">
        <div>
            <asp:Button ID="Butt1" runat="server" Text="Hi" onclick="Butt1_Click"/>
            <asp:TextBox ID="TextBox1" runat="server"></asp:TextBox>
        </div>
    </form>
</body>
</html>
```

**Exemplo2:** Código C# do ficheiro aspx.cs, está ligado ao ficheiro Default.aspx (Code Behind):

```
using System;
using System.Collections.Generic;
using System.Linq;
using System.Web;
using System.Web.UI;
using System.Web.UI.WebControls;
namespace Example
{
    public partial class _Default : System.Web.UI.Page
    {
        protected void Page_Load(object sender, EventArgs e){
        }
        protected void Butt1_Click(object sender, EventArgs e){
```

```
        TextBox1.Text = "Hello World!!";  
    }  
}  
}
```

O exemplo 2 é uma página *web* de extensão `Default.aspx`, onde existem dois elementos HTML: um “button” e uma “textbox”. Sempre que se carrega no botão ele faz uma ligação ao servidor (`Default.aspx.cs`), que por sua vez carrega a caixa de texto com o texto Hello World!.

### 4.2.5.1 ASP.NET *Master Pages*

*Master Pages* é um recurso da framework.NET que facilita muito o desenvolvimento de um *web site* em ASP.NET, uma vez que não é necessário repetir passos de desenho em cada uma das páginas *aspx*. Com as *Master Pages* é possível criar uma página padrão (modelo) de extensão `.master` sendo que as restantes páginas *web* criadas posteriormente herdam a aparência visual da *Master Page*. É como se fosse uma página *default* com menus, cabeçalhos e rodapés. Qualquer outra página *aspx* criada, pode herdar da *Master Page* e o programador só utiliza a área que não seja da *Master Page*, normalmente designada por *content* onde o conteúdo normalmente é variável. As páginas de extensão `.aspx` que herdam da *Master Page* são, tipicamente páginas HTML mas divididas em *contents*.

O cabeçalho, a secção de navegação e o rodapé na página *web* são elementos comuns dentro do *site*, mudando apenas o conteúdo à medida que se navega pelas páginas. Antes das *Master Pages* havia formas de colocar estes elementos em cada página *web* através de vários de meios, mas na maioria dos casos, era uma dificuldade acrescida. Os programadores limitavam-se a copiar o código HTML em cada página *web*, o que, embora funcione dava bastante trabalho de manutenção, especialmente no caso de sites de dimensões médias/grandes.

O uso de *Master Pages* aumenta a produtividade e facilita a manutenção de *web sites*, permitindo criar uma equipa de trabalho, com um padrão de desenvolvimento a ser usado por todos os programadores. As aplicações *web* tornam-se assim, mais fáceis de construir

e mais simples de gerir. É importante destacar que o uso desta técnica nesta dissertação proporcionou páginas *web* mais elaboradas e consistentes. Para mais informações acerca das Master Pages, sugere-se a consulta do livro *Professional ASP.NET 3.5 In C# and VB* [Evjen et al., 2008].

#### 4.2.5.2 ASP.NET AJAX *Control Toolkit*

O ASP.NET AJAX Control Toolkit é um projecto *open-source*, construído no topo da *framework* Microsoft ASP.NET AJAX. É um esforço conjunto entre a comunidade AJAX e a Microsoft, no sentido de disponibilizar uma infraestrutura rica em controlos e módulos em ASP.NET AJAX, que possam ser reutilizados e customizados de forma a criar uma experiência *web* interactiva.

Como o nome indica, o ASP.NET AJAX é uma plataforma que se baseia na tecnologia AJAX. O AJAX trata-se de uma técnica desenvolvida por Jessé James Garret que, além de ser executada em qualquer *browser* moderno, permite transferir pequenas quantidades de informação entre o servidor e o *browser* sem que o utilizador tenha essa percepção. Utiliza o objecto XMLHttpRequest para a comunicação assíncrona com o servidor, efectua um pedido HTTP ao servidor, recebe os dados do servidor em XML, processa esses dados e altera o *site* de acordo com os dados recebidos. Esta nova tecnologia, vista como uma fusão de tecnologias (XML, JavaScript, CSS e DOM), pretende mudar o paradigma de como os utilizadores interagem com a Internet na medida em que dá um grande passo em termos de riqueza gráfica quando comparada com as aplicações *desktop* tradicionais [Kanjilal and Putrevu, 2009].

A utilização de ASP.NET AJAX Control Toolkit no *BCSurvival* permitiu proceder a alterações na página *web* com muito menos transferência de informação, pois apenas foram transferidos do servidor os dados necessários, sendo o restante HTML do *site* reutilizado. Por outro lado, como apenas partes das páginas são actualizadas, e essa actualização é normalmente muito mais rápida do que se toda a página fosse actualizada, o utilizador chega a ter a sensação de que tudo acontece em tempo real, localmente no seu computador. Assim sendo, com esta tecnologia, só uma porção da página *web* é actualizada e não a

página toda, reduzindo portanto o consumo de largura de banda nos pedidos efectuados ao servidor [AJAXToolkit, 2010]. No sistema desenvolvido é notória a presença de módulos ASP.NET AJAX Control Toolkit. Estes permitiram dar uma outra riqueza gráfica ao *site*, como por exemplo a colocação de marcas de água nos controlos ASP.NET, o uso de um editor de HTML, os cantos arredondados nos controlos e o uso de *tabs* personalizadas.

### 4.2.6 JavaScript

O JavaScript é uma linguagem de *script* interpretada, desenvolvida em 1995 por Brendan Eich como uma extensão do HTML. É suportada por qualquer *browser* moderno e tanto pode ser usada no lado do cliente como no lado do servidor. O JavaScript pode ser usado facilmente para interagir com os elementos HTML, validar o conteúdo inserido pelos utilizadores nos formulários, gerir as definições da página *web* (como as cores de diferentes controlos, tratar de eventos “onclick” ou “onmouseover”, etc). Como qualquer linguagem de programação, permite ao autor de *scripts*, aceder, usar e modificar atributos de objectos que surjam numa página de internet manipulando assim, os elementos de uma página. Embora o HTML por si só não seja dinâmico, o JavaScript, ao correr eventos do tipo “onclick” ou “onmouseover”, consegue transformar os elementos estáticos de uma página HTML em elementos dinâmicos. Assim, surgiu a tecnologia DHTML (União de JavaScript e CSS) [Kanjilal and Putrevu, 2009].

Nesta dissertação a tecnologia escolhida para fazer de “ponte” entre o Silverlight e o servidor IIS foi o JavaScript pela facilidade de se fazer *debug* em tempo real com projectos *web* desenvolvidos no Visual Studio e no *browser* IE. Outro factor de decisão foi que este garante apenas que alguns dados (idade, tamanho do tumor) indicados pelo utilizador sejam transferidos entre o *browser* e o servidor IIS. Assim, garante-se que só algumas partes do *site* sejam actualizadas, como por exemplo os dados do paciente inseridos, os intervalos de confiança, os grupos de risco e o gráfico de dispersão com o respectivo índice de prognóstico calculado. A figura 4.2 indica os locais onde os dados são alterados na página:

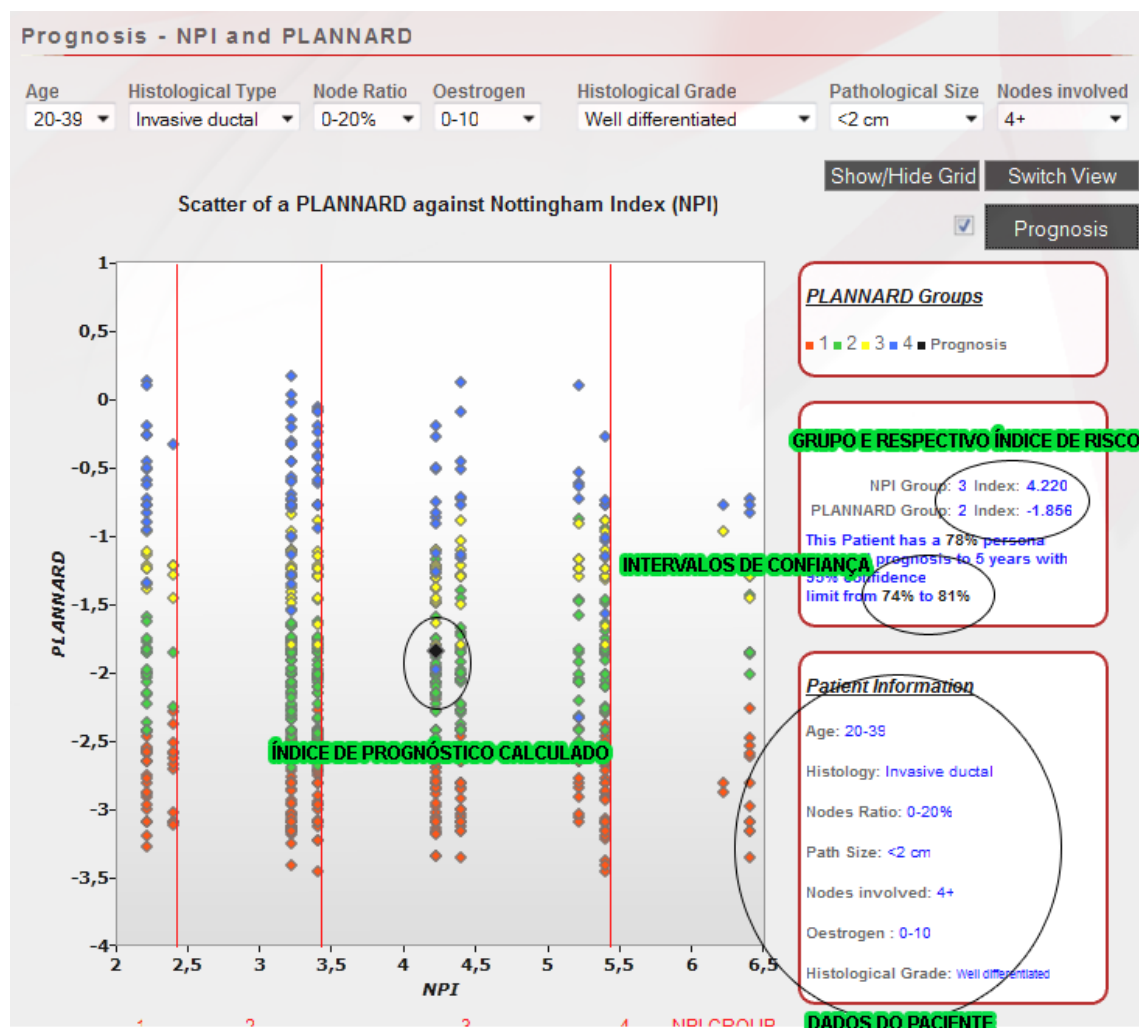


Figura 4.2: Localização dos dados alterados na página

### 4.2.7 jQuery

O jQuery é um novo tipo de biblioteca JavaScript grátis, simples e rápida criada por John Resig, em Janeiro de 2006 [jQuery, 2010] baseando-se no lema:

“Escrever menos, fazer mais”

Esta biblioteca foi desenhada para simplificar a manipulação de eventos em *web sites*, sendo usada em cerca de 20% dos 10.000 maiores sites *onlines* existentes [Battles, 2010]. Este tipo de biblioteca tem como objectivo facilitar e uniformizar o desenvolvimento com Javascript, abstraindo os programadores de problemas como incompatibilidade entre *browsers*, focando-os apenas no desenvolvimento puro. A biblioteca jQuery é constituída por um

ficheiro JavaScript único de extensão .js ocupando apenas 14KB, onde tenta criar um modelo de abstração que permite programar de uma forma intuitiva. O exemplo que se segue indica um simples exemplo de jQuery numa página HTML:

```
<div>
    <a href="#">Clica Aqui</a>
    <div id="divTexto">
        Texto...
    </div>
</div>
```

Na header da página HTML insere-se o código:

```
<script type="text/javascript" language="javascript">
    $(document).ready(function() {
        $("a").toggle(function() {$("#divTexto").hide('slow');}, function() {
            $("#divTexto").show('fast');});});
</script>
```

Esta função esconde ou mostra a div “divTexto” (`$("#divTexto")`) sempre que se clica em qualquer *link* da página *web*. Para se obter o objecto divTexto utiliza-se `$("#divTexto")` que é muito semelhante ao `document.getElementById("divTexto")` mas, muito mais simples e sem problemas de compatibilidade entre *browsers*.

O jQuery proporciona a construção de *scripts* de uma forma simples e intuitiva conseguindo com meia dúzia de linhas de código os mesmos efeitos de um *script* de 30 a 40 linhas de código no JavaScript tradicional [Silva, 2008], o que constitui uma grande vantagem desta biblioteca. Ao permitir interagir com os elementos de uma página HTML, cria-se uma certa dinâmica nos sites. Engloba ainda um conjunto de diversas funções comuns para tratamento de eventos assim como, uma API simples e inteligente, tirando benefício dos elementos de uma página HTML a um nível bastante abstrato. Ao longo da dissertação usou-se o jQuery, exemplos disso são a forma como se alterna entre as “caixas” (divs) de envio de mensagens

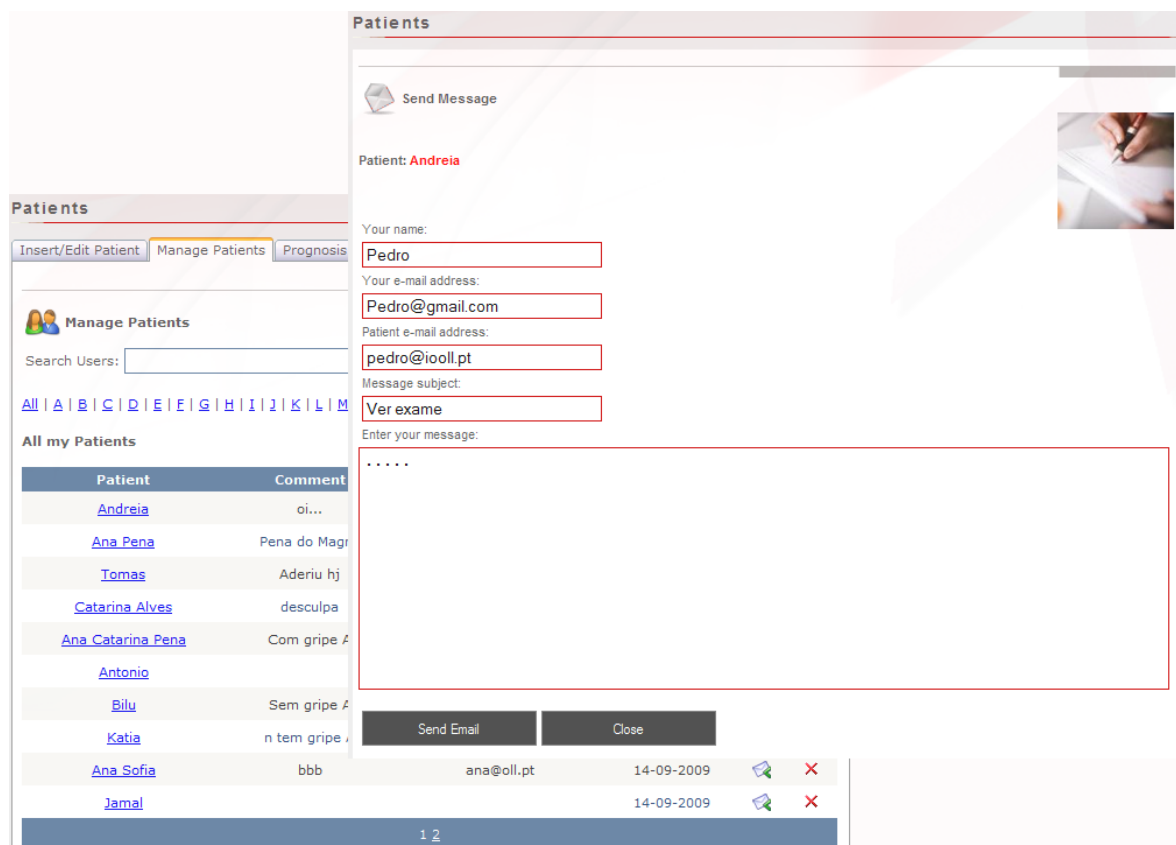


Figura 4.3: Envio de email para um paciente

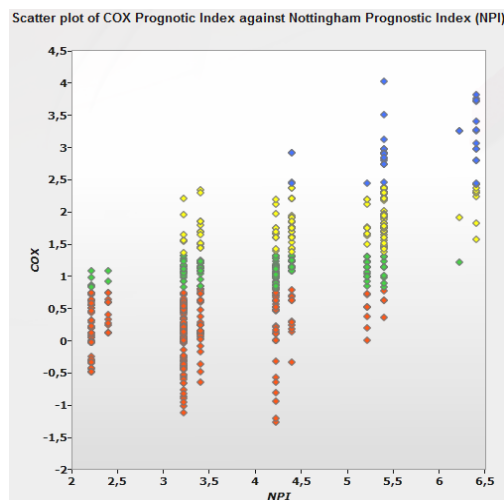
para os vários paciente e na “caixa” onde se encontra toda a informação dos pacientes, como é representada na figura 4.3.

#### 4.2.8 SILVERLIGHT

O Silverlight é um plug-in para *browsers* cujas potencialidades permitem às empresas de *design* e desenvolvimento a criação de experiências graficamente ricas para *web*, *desktop* e PDA(s). O Silverlight é desenvolvido em plataformas como o Visual Studio ou o Blend 2, que propõe combinar o visual e a funcionalidade, ou seja, pretende criar uma ponte entre *designers* e programadores. Esta tecnologia é suportada pela plataforma .NET, que pretende juntar num único ambiente de trabalho as funcionalidades existentes no Adobe Flash, a facilidade de integração com multimédia e as animações [Silverlight, 2010]. Uma aplicação *web* não tira grande partido do lado do cliente em termos gráficos e fazer uma aplicação usando apenas HTML e JavaScript é possível, mas leva a um excesso de código, sendo a sua

manutenção uma tarefa com dificuldades acrescidas. Até agora a única solução credível para páginas ricas graficamente tem sido o Macromedia Flash. A dificuldade de usar este tipo de tecnologia prende-se fundamentalmente com o tempo necessário para se aprender a desenvolver numa plataforma Flash, nomeadamente para aprender a linguagem de programação ActionScript. Para tentar contornar este facto, a tecnologia Silverlight é suportada pela plataforma .NET. Assim, um programador de aplicações *web* com conhecimentos de linguagem C# tem a capacidade de criar aplicações Silverlight mais rapidamente e com as mesmas funcionalidades de uma aplicação em Flash. O Silverlight pretende portanto competir com a Adobe Flash [Swift et al., 2009].

A figura 4.4 representa a aplicação Silverlight desenvolvida nesta tese para correr no lado do cliente, que é responsável por mostrar a localização dos diferentes grupos de risco e índices de prognóstico dos pacientes do *Christie Hospital*.



**Figura 4.4:** Aplicação criada em Silverlight

A criação deste gráfico foi feita a partir de dois ficheiros: o ficheiro XAML e o ficheiro *code-behind*. O ficheiro de extensão XAML suporta a componente gráfica. Neste, define-se o tipo de gráfico, as cores e tipos dos pontos assim como comprimento máximo e mínimo dos eixos e as suas orientações. O exemplo seguinte é mostrado código parcial da componente gráfica:

```
<DVC:ScatterSeries.DependentRangeAxis>
<DVC:LinearAxis Title="" ShowGridLines="False" Orientation="Y" Minimum="-2">
```



```

<DVC:LinearAxis.MajorTickMarkStyle>
    <Style TargetType="Line">
        <Setter Property="Visibility" Value="Collapsed" ></Setter>
    </Style>
</DVC:LinearAxis.MajorTickMarkStyle>
<DVC:LinearAxis.AxisLabelStyle>
    <Style TargetType="DVC:AxisLabel">
        <Setter Property="StringFormat" Value="."/ >
        <Setter Property="Opacity" Value="0" ></Setter>
    </Style>
</DVC:LinearAxis.AxisLabelStyle>
</DVC:LinearAxis>
</DVC:ScatterSeries.DependentRangeAxis>

```

No ficheiro em *code-behind* define-se para cada ponto (1 paciente) os índices de risco, usando os três modelos descritos anteriormente: o NPI, a Regressão de Cox e o PLANN-ARD. O exemplo abaixo indica os dois primeiros pontos correspondentes a dois dos 743 pacientes do *Christie Hospital* associados ao Grupo 1. É repetido o processo tanto para o grupo 2, 3 e 4 do modelo Cox ou PLANN-ARD.

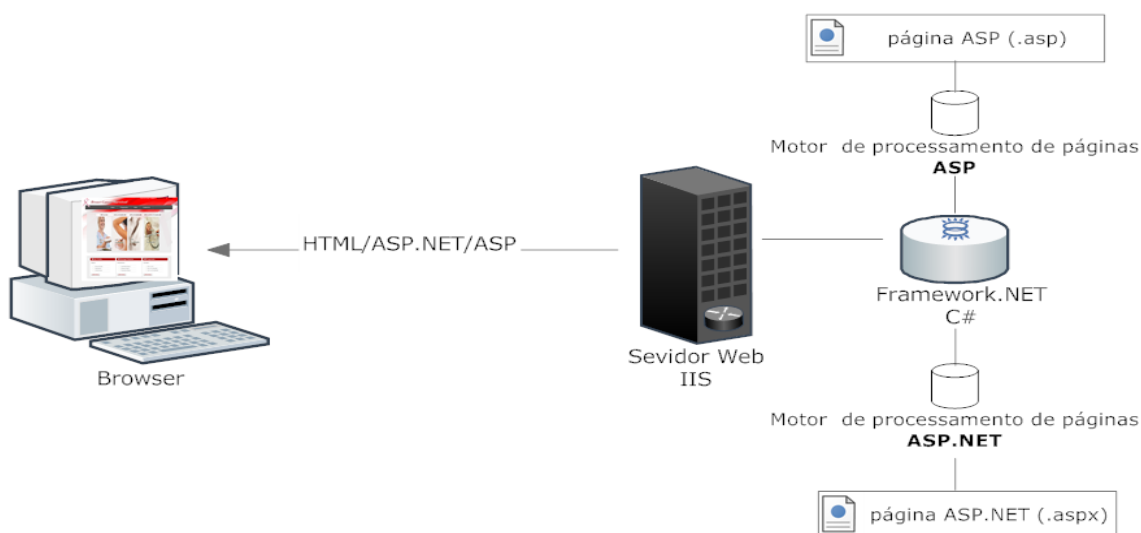
```

(ScatterSeries)mcChart.Series[0]).ItemsSource =
    new KeyValuePair<double, double>[] {
        #region ValoresGRUP01
            new KeyValuePair<double, double>(4.22,-0.81526),
            new KeyValuePair<double, double>(3.22,0.24773),
            ...
        }

```

### 4.2.9 IIS

O IIS foi o servidor escolhido nesta dissertação, que integrado com o próprio sistema operativo Windows é um servidor capaz de gerir o conteúdo de uma página ASP.NET. Este é formado basicamente por dois tipos de aplicações: os *Web Services* e as páginas *web*. Oferece como funcionalidades: correio electrónico, módulo para a criação de servidores FTP, autenticação, cache, CGI, filtros de conteúdo e de páginas, páginas de erro, regras de autorização e restrição de domínio e suporta diferentes tipos de MIME. Além do mais, o IIS oferece todas as bases de interoperabilidade com a plataforma .NET, mantendo-se assim todas as tecnologias dentro da mesma plataforma, facilitando a sua integração e interacção [IIS, 2010].



**Figura 4.5:** Processo de pedido-resposta de uma página ASP.NET

O servidor IIS aceita e responde a pedidos usando o protocolo HTTP (Hypertext Transport Protocol), sendo o resultado a essa resposta uma página HTML. Como as páginas ASP.NET supõem conteúdo dinâmico, o processo de pedido-resposta é um pouco mais complexo. As páginas de extensão aspx são feitas de *tags* ASP.NET (<asp:button). Quando é solicitada uma página de extensão aspx, o servidor IIS não sabe processar a informação contida nas páginas .aspx. O que o servidor IIS tem, é uma lista de motores específicos responsáveis por processar cada tipo de página seja esta ASP, ou ASP.NET. A figura 4.5

representa um processo de passagem de pedido para o motor responsável. O servidor IIS envia então o pedido à Framework.NET que é quem realmente processa o pedido de página. Desta forma, as aplicações *web* em ASP.NET beneficiam da vantagem de execução de código na Framework.NET já que o código é compilado e executado pela Framework.NET sendo devolvido ao IIS para que este envie a página ao utilizador.

#### 4.2.10 *Web Services*

Os *Web Services* são uma tecnologia que fornece serviços de dados através da internet, geralmente sobre o protocolo HTTP. O serviço é descrito no formato WSDL; os clientes que usem o *Web Service* acedem ao servidor utilizando mensagens SOAP sendo os dados veiculados no formato XML. Com a utilização destas APIs de *web* na rede, qualquer software utilizando os mesmos padrões poderá trocar dados entre si [Pereira and Poupa, 2005].

No *site* desenvolvido, o plugin em Silverlight (gráfico de dispersão), interage com os 3 modelos de prognóstico, recorrendo a *Web Services* colocados à disposição no servidor IIS. O servidor IIS que se encontra englobado na plataforma de aplicações *web* ASP.NET na *framework* .NET providencia deste modo um mecanismo que permite que vários pedidos de prognóstico sejam realizados simultaneamente, garantido um processo transparente para os utilizadores.

A adopção de uma tecnologia baseada em *Web Services* garante uma via secundária de comunicação entre cliente e servidor. Isto é, no caso da aplicação *web* desenvolvida, o gráfico em Silverlight com os respectivos índices de prognóstico, usa a linguagem JavaScript como “ponte” para pedir um serviço de prognóstico ao servidor IIS. O objectivo destes serviços é permitir que seja feito um pedido ao servidor, e como resposta, este apenas actualize algumas partes do *site* incluindo o gráfico com o respectivo índice de prognóstico calculado. A não adopção deste método aumenta o fluxo de dados entre cliente e servidor uma vez que, em cada interacção são pedidos ao servidor todos os dados necessários para recriar a página ASP.NET incluindo o gráfico Silverlight. O ficheiro *Web Service* criado pelo ambiente

de desenvolvimento Visual Studio 2008 tem como objectivo disponibilizar de uma forma transparente ao utilizador três serviços de prognóstico (NPI, COX e PLANN-ARD), com o cuidado de utilizar o atributo *WebMethod* para cada método, indicando que o método é um *Web Service*. Temos como exemplo o serviço de prognóstico NPI:

```
[WebMethod]
public double[] LoadNPI(int pathSize, int histGrade, int nodesInvolved){
    valuesNPI[0] = npi.SetNPI(pathSize, histGrade, nodesInvolved);
    valuesNPI[1] = npi.GetGroupNPI();
    ...
    return valuesNPI;
}
```

A capacidade de encapsulamento e de interoperabilidade faz dos *Web Services* uma mais valia para as aplicações *web* para além de permitir um processo transparente para o utilizador.

### 4.2.11 SQL Server 2005

O SQL Server 2005 foi o servidor escolhido para gerir a base de dados desta dissertação. Oferece a aplicação de gestão de base de dados, o *SQL Server Studio Management Express*, que permite guardar e gerir dados de uma forma segura, dando ao utilizador garantias, de que as alterações sobre os dados são aplicadas eficientemente, sejam estas de criar, alterar ou apagar dados no servidor. Além disso, o SQL Server combina recursos de análise, geração de relatórios, integração e notificação. Isso permite a um administrador de um *web site* ter a oportunidade de criar e implementar soluções económicas que ajudem, por exemplo, uma equipa a distribuir os dados através de dispositivos *web* e PDA(s) [SQLServer, 2010].

A facilidade de integração do SQL Server com o Microsoft Visual Studio .NET, foi um factor decisivo, pois torna o desenvolvimento da aplicação *web* na depuração do código bem mais simples quando em comparação com outras ferramentas, nomeadamente o MySQL, phpMyAdmin ou heidiSQL.

#### 4.2.12 LINQ

A tecnologia LINQ veio revolucionar por completo a forma como são tratados os dados das aplicações. Desenvolvida pela Microsoft contribui para aumentar a produtividade dos programadores, isto é, torna possível utilizar expressões semelhantes para obter dados a partir de uma base de dados SQL Server ou a partir de um documento XML. Hoje em dia existe uma grande diversidade de dados que alimentam uma aplicação. Por exemplo, para além das tradicionais bases de dados SQL Server, a interacção com dados XML é muito comun. Se tivermos em conta apenas estas duas “origens de dados”, é fácil concluir que o código necessário à interacção com ambas é bastante diferente. Assim, o LINQ contribui para diminuir a complexidade inerente à pesquisa de dados oriundos de fontes de dados heterogéneas. Estando integrada na plataforma .NET fornece uma nova sintaxe sobre qualquer fonte de dados. Para além de permitir uma maior legibilidade, garante funções de pesquisa sobre várias fontes de dados, reduzindo significativamente as linhas de código necessárias para interagir com os dados. Do ponto de vista da produtividade é das funcionalidade mais importante introduzida pela Microsoft nos últimos anos [Abreu and Morgado, 2009] [LINQ, 2010]. No código seguinte pode-se observar um exemplo de um comando LINQ implementado em C# nesta tese. Primeiro define-se o objecto LINQ com todos os métodos e propriedades necessários para se ligar à base de dados SQL Server. De seguida, faz-se uma *query* que vai devolver o nome de um paciente com determinado ID associado a um médico.

```
BreastCancerLinqToSqlDataContext bc=new BreastCancerLinqToSqlDataContext();
Guid UserDoctorId=(Guid)Session["User"];
var allMyPatients=(from Doc_Pat_Diag in bc.Doctor_Patient_Diagnostics
                   join Pat in bc.Patients
                   on
                   Doc_Pat_Diag.PatientId equals Pat.PatientId
                   where Doc_Pat_Diag.UserId == UserDoctorId
                   select new
```

```
{  
    Name_Patient = Pat.Name,  
}).Distinct();
```

### 4.2.13 MATLAB

O Matlab (nome derivado de *matrix laboratory*) é visto como uma linguagem técnica de programação de alto nível e como uma interface gráfica intuitiva para o desenvolvimento de algoritmos, tratamento de dados e cálculo numérico. Esta interface gráfica permite aos utilizadores resolverem problemas através da notação matemática, cujo elemento básico de informação é uma matriz que não requer qualquer tipo de dimensionamento.

A capacidade de integração de código Matlab com outras linguagens de programação, teve um papel importante nesta dissertação. Essa capacidade permite a sua distribuição em aplicações web ou desktop. Pois ao oferecer uma interface própria para a criação de classes .NET - Matlab *Compiler*, garantiu a integração do modelo PLANN-ARD realizado em Matlab com a plataforma ASP.NET desenvolvida.

#### 4.2.13.1 Matlab *Compiler*

O Matlab *Compiler* é um componente do Matlab que permite integrar as aplicações Matlab com as aplicações .NET. Para isso converte as aplicações de extensão .mat em executáveis ou dll's. Anexado aos executáveis e dlls's pode-se incluir o *engine* MCR e com isso é possível correr a aplicação fora do ambiente trabalho do Matlab. Normalmente, se o utilizador gerar uma aplicação independente cria-se um executável. Caso o utilizador queira integrar o seu código numa aplicação em C#, através da interface cria-se uma dll com procedimentos e funções, permitindo o seu uso no código C#.

Uma vez que o cálculo do prognóstico de sobrevida e respectivos intervalos de confiança para um paciente é realizado a partir de ficheiros Matlab, a integração da aplicação *web* com estes ficheiros torna-se inevitável. Existem 3 métodos de integração do Matlab com a

linguagem C#:

1. Usar o Matlab como servidor ao mesmo tempo que o C#. A comunicação é realizada via COM. Permite fazer *debug* (processo de encontrar e reduzir erros numa aplicação) em simultâneo tanto do lado do servidor, como no lado do C#.
2. Criar uma dll com o Matlab *Compiler*. Com esta dll as funções em Matlab são vistas como métodos de uma classe. Estes métodos podem ser usados directamente no C#. Este método garante *garbage collection* (gestão de memória), *intellisense* (codificação inteligente, isto é completa as palavras à medida que se programa), etc. Além disso, não é necessária a instalação do Matlab nos computadores, podendo ser usado o MCR.
3. Criar uma biblioteca partilhada em C com o Matlab *Compiler*. Este método é considerado muito frágil não sendo aconselhado para integrações em .NET. Não usa o *garbage collection* nem o *intellisense*.

Pelas opções indicadas, a escolha recaiu sobre o segundo método. A utilização deste método garante a interoperabilidade das aplicações deixando de ser necessário a tradução de código Matlab para código .NET [MatlabCompiler, 2010].

## 4.3 Implementação do Site - *BCSurvival*

### 4.3.1 Base Dados

Um sistema *online* de suporte à decisão tem como objectivo possibilitar a classificação do estado de saúde de um paciente baseando-se para isso em determinadas variáveis dos pacientes. Depois de efectuado o cálculo de prognóstico, torna-se fundamental que essas variáveis médicas sejam guardadas numa base dados para futuras consultas de modo a ser possível verificar a evolução do paciente e comparar com prognósticos de outros pacientes. Pretende-se portanto, com esta base dados, auxiliar a tomada de decisão de um médico oncologista sobre os pacientes com cancro da mama.

A base dados eleita para este sistema foi o SQL Server 2005. A escolha desta base dados teve em conta não só a ligação que existe com a plataforma ASP.NET 3.5, mas também o facto do SQL Server ser um produto Microsoft, tal como o ASP.NET, evitando assim qualquer problema de incompatibilidade entre as duas plataformas.

Uma das grandes vantagens da utilização da base dados SQL Server consiste no facto de, quando implementada na plataforma ASP.NET 3.5, criar por defeito um conjunto de tabelas com relações entre si, funções e procedimentos de elevada utilidade para os programadores. Para uma aplicação *web* suportar um sistema de *login*, torna-se indispensável implementar uma base dados com informação acerca dos utilizadores que permita validar as credenciais no *login* do *site*, assim como suportar novos registos ou mesmo recuperar a password. Para contornar estas dificuldades iniciais, a equipa responsável pelo ASP.NET introduziu um conjunto de classes num módulo designado por *Membership* na *.NET framework*. Desta forma, a produtividade dos programadores aumentou, já que estas classes identificam todos os utilizadores com acesso à aplicação, se o utilizador está bloqueado, qual a hora da última visita ao *site*, etc.

#### 4.3.2 Interface Gráfica

A interface gráfica do *web site* tem a finalidade de facilitar o acesso às diferentes funcionalidades do *site*, tornando a navegação o mais intuitiva possível aos utilizadores.

Uma vez que se trata de uma aplicação onde o utilizador tem ao seu dispor graficamente todos os dados de pacientes é fundamental que os dados sejam confidenciais, pois o seu acesso só diz respeito ao utilizador que os criou. Estes dados, só poderão portanto ser acedidos por quem tem permissão para entrar no *site*, pelo que, foi criado um sistema de *login* em que cada utilizador tem associado vários pacientes. Na figura 4.6 é apresentada a *home page* do *web site - BCSurvival*, para qualquer utilizador quer esteja registado quer seja anónimo. Após o utilizador ter efectuado o *login* aparece a página principal apresentada na figura 4.7.

O primeiro passo para a criação da interface gráfica do *web site* foi a definição do *layout*



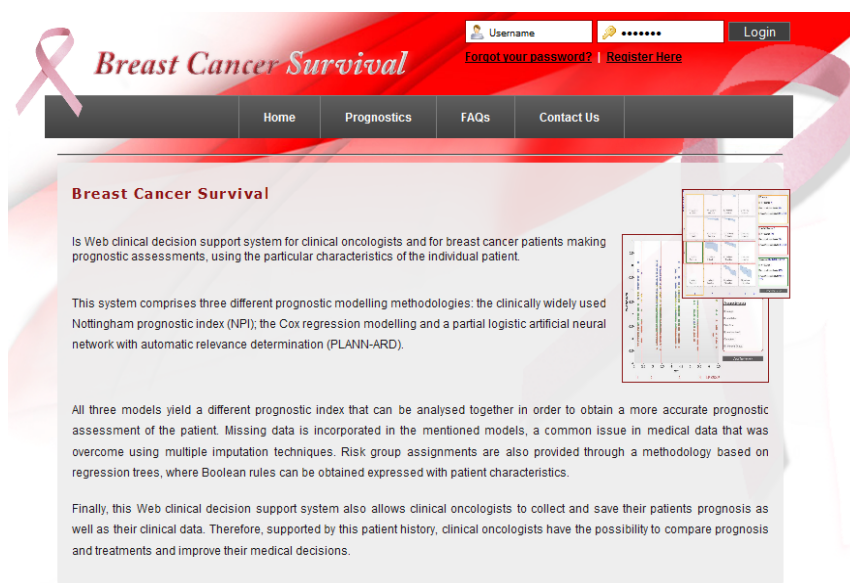


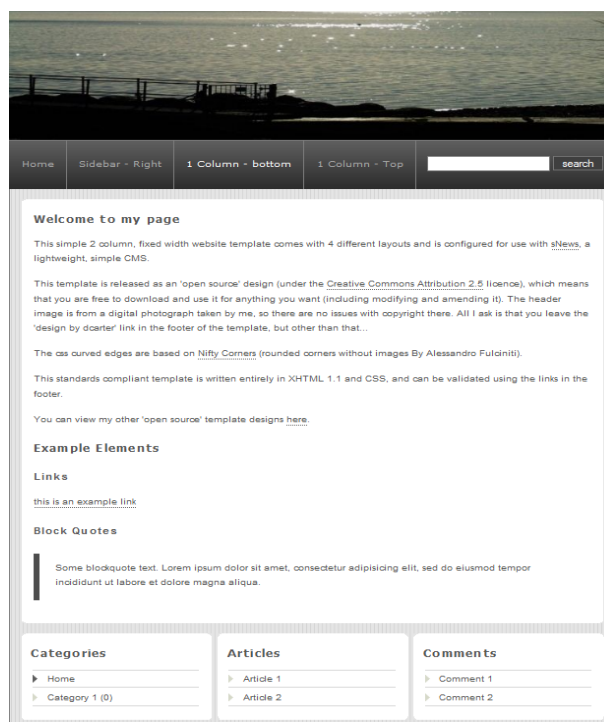
Figura 4.6: Página inicial do site *BCSurvival*

e a disposição dos diferentes objectos pela página. Para tal, foi efectuada previamente uma pesquisa pelos vários *templates* fornecidos gratuitamente pela internet, de forma a encontrar um *template* com a estrutura mais adequada para aplicações médicas. O *template* utilizado como base para o site *BCSurvival*, é apresentado na figura 4.8 [DCarterDesign, 2010].



Figura 4.7: Página inicial do site *BCSurvival*, para utilizadores registados

De seguida, foi necessário alterar o código do *template* que faz parte do ficheiro *MainStyles.css*. O ficheiro *MainStyles.css*, armazenado na pasta `\App-Themes\Red\`, é utilizado para definir o *layout* de todas as páginas do *site*, ou seja, este ficheiro trata das fontes, cores



**Figura 4.8:** *Template* por defeito

e outras características da apresentação gráfica. As páginas do *web site*, excepto as páginas da administração, antes de serem carregadas são formatadas segundo as regras designadas no ficheiro MainStyles.css. O uso de ficheiros de extensão .css, está relacionado com a necessidade de elementos ou objectos associados a uma página HTML usarem *tags* múltiplas. Para chamar uma regra do ficheiro .css o objecto HTML usa a *tag* “class”, sendo possível haver repetições pelo código HTML. Caso não se queira repetir, usa-se simplesmente a *tag* “id” do elemento HTML. Assim, o ficheiro MainStyles.css trata do grafismo (cores, tamanho de letras, margens...) e os ficheiros BreastCancerMainPage.Master e Default.aspx definem o aspecto, ou seja, são responsáveis pela colocação dos diferentes elementos na página.

No *site* estão disponíveis várias funcionalidades de navegação tais como: Menu Principal, Menu de Topo, Área de *Login*. O Menu de Topo (figura 4.9), que tem como referência [CSSMenuMaker, 2010], foi modificado e alterado de forma a ter uma cor idêntica ao *site* em geral, a largura e comprimento foram ajustadas à largura do *web site* com 773px e 43px, respectivamente.

Alguns botões têm como referência [CSSLibrary, 2010]. As cores do texto foram também

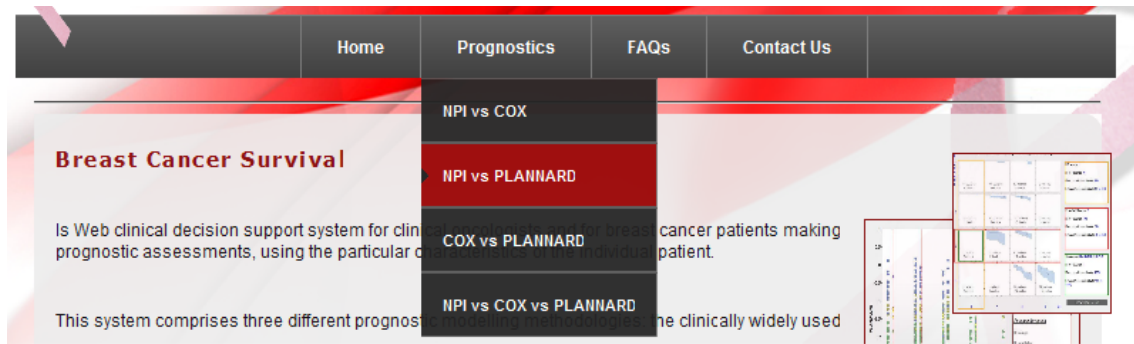


Figura 4.9: Menu de Topo

alteradas assim como o tamanho foi modificado de forma a ficar mais pequeno. As *labels* em formato rectangular com os cantos arredondados foram obtidas em [RoundedCorner, 2010]. As cores, margens, tamanho da letra, no interior das *labels*, foram modificadas para terem uma cor homogénea em relação aos outros elementos disponíveis na página *web*. A figura 4.10 exemplifica uma das *labels* existentes.

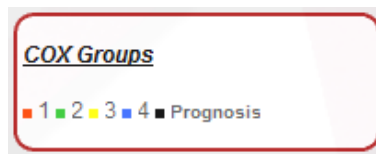


Figura 4.10: *Labels* em formato rectangular com os cantos arredondados

Foi criada uma imagem de fundo em tons de vermelho e cor-de-rosa, tendo em conta o objectivo do prognóstico, que é o cancro da mama, juntando o símbolo do cancro da mama - laço em tom cor-de-rosa. O esquema de cores das restantes imagens, pertencentes aos vários menus disponíveis, foram alterados de forma a combinar com a cor cor-de-rosa que passou a ser a cor predominante do *BCSurvival*. O programa utilizado para a edição e criação das imagens no *site* foi o Adobe Photoshop CS4.

Foi feito um favicon (imagem no formato .ico de 16x16 pixéis), como indicado na figura 4.11, para este *web site*. Os favicons são imagens em miniatura do logótipo do *site* que representam, aparecem antes dos URLs, nos favoritos, ou então nas *tabs* do *browser*. Os favicons ajudam o utilizador a identificar uma página quando esta está guardada nos favoritos. Para gerar o favicon do *BCSurvival* recorreu-se a uma ferramenta disponível *online* (<http://www.favicon.cc/>), utilizando a imagem do logótipo.



**Figura 4.11:** Favicon

As restantes modificações gráficas tiveram em conta a seguinte referência [Duckett, 2008]. Em suma, pretende-se com esta interface proporcionar ao utilizador uma navegação intuitiva pelo *web site*, assim como fornecer ferramentas gráficas de acesso às várias metodologias de prognóstico do cancro da mama.

#### 4.3.3 Colocação do Site *online*

Esta subsecção pretende enumerar e descrever os processos necessários para colocar o *site online*. O primeiro passo foi a instalação e configuração do servidor IIS e Matlab. O IIS foi considerado o servidor mais apropriado para a implementação deste projecto. O segundo passo explica a importância do ficheiro *web.config* na ligação entre o servidor, o *site* e a base de dados. O terceiro passo demonstra de que forma é possível correr as aplicações Matlab e Silverlight no *web site* a partir do servidor IIS. Por último, é identificado o nome e URL atribuído ao *site* para sua visualização no *web browser*.

Para correr um *site* ASP.NET e ficheiros Matlab são necessários, no entanto determinados requisitos mínimos:

- Instalar o servidor IIS 5.0 ou superior.
- Configurar um directório virtual associado à aplicação *web*.
- Instalar a *Framework.Net*.
- Incluir os ficheiros *.aspx* e a Master Page (*.master*).
- Instalar o SQL Server 2005.
- Instalar o Matlab ou o MCR.

O IIS (*Internet Information Services*) é um componente do Windows incluído nas versões Windows 2000, XP, Vista, 7 e Server 2003. Por pré-definição no Windows, este componente não se encontra disponível e por isso, a sua instalação é realizada mediante o ícone de “Adicionar ou remover programas” no painel de controlo escolhendo a opção: “Adicionar ou remover componentes do Windows”. Depois, é necessário seleccionar a *check box* correspondente ao IIS e por fim, escolher a opção para “instalar”. O acesso ao IIS é feito mediante o ícone de “Serviços de Internet Information Server”, que se encontra em “Ferramentas Administrativas”, dentro do Painel de Controlo. Por padrão, qualquer directoria dentro da pasta de raiz do IIS é um potencial *site* de Internet, sendo o seu caminho físico: C:\inetpub\wwwroot\.

A instalação foi efectuada com sucesso pelo *web browser* se ao aceder-se ao `http://localhost/` ou `http://127.0.0.1/`, surgir no *browser* a home page do IIS (figura 4.12). Após



Figura 4.12: Home Page do IIS

a instalação do IIS, o *BCSurvival*, construído pela tecnologia ASP.NET 3.5 no Visual Studio 2008, é colocado numa directoria dentro da raiz: C:\inetpub\wwwroot\BreastCancer\ podendo ser acedido pelo *web browser* em `http://localhost/BreastCancer/`. Apesar do *site* se encontrar dentro de uma pasta da raiz do servidor, por defeito este não o reconhece como um *site*. A essa pasta é necessário convertê-la numa aplicação para que os ficheiros ASP.NET nessa directoria sejam executados automaticamente, utilizando um conjunto de dados próprios de sessão.

De seguida cria-se um directório virtual (caminho físico na *web*), a partir de um directório físico, como indicado na figura 4.13. A pasta onde se encontram as páginas *web* do *BCSurvival*

### 4.3. IMPLEMENTAÇÃO DO SITE - BCSURVIVAL

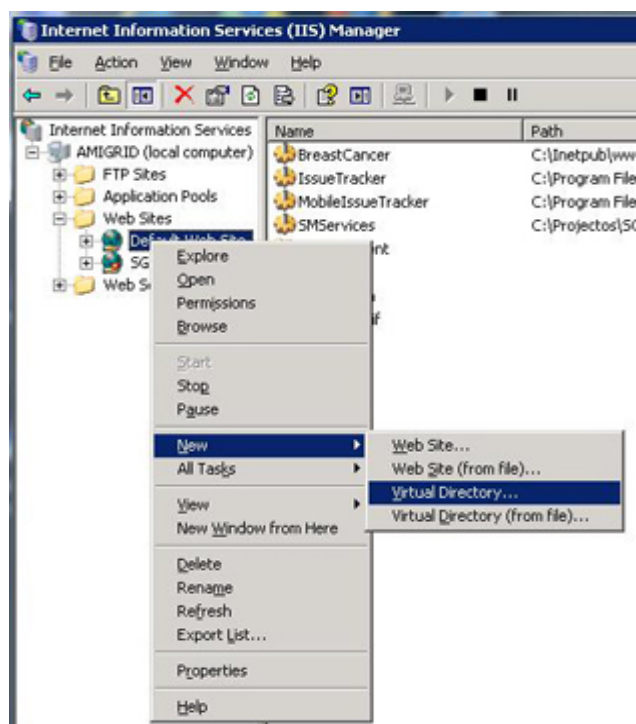


Figura 4.13: Criação de um directório virtual

tem como caminho físico: `C:\Inetpub\wwwroot\BreastCancer\` e como directório virtual *BreastCancer*. Para aceder ao *site* basta colocar no URL do *browser*: `http://localhost/BreastCancer` ou então `http://IPDaMaquina/BreastCancer/`. A criação destes directórios virtuais garante que ao aceder a este *link* apenas se consegue visualizar o *site* e não o conteúdo da pasta *www*. Após a criação do directório virtual, algumas propriedades deste directório, podem ser alteradas, como indicado na figura 4.14. Por exemplo, se o autor do *web site*

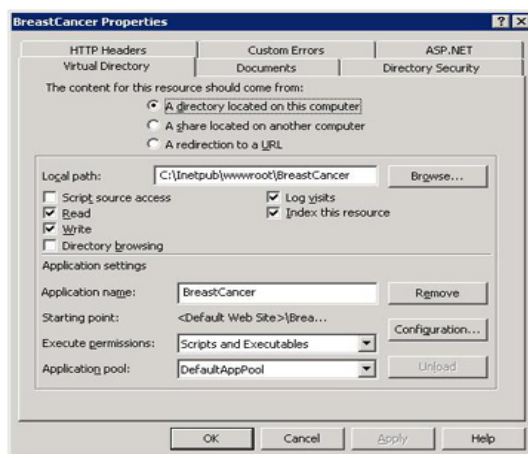


Figura 4.14: Propriedade do *site* - *BCSurvival*

não tiver intenção que se visualize o conteúdo (ficheiros) de uma pasta, a opção “Directory browsing” não pode estar seleccionada. Outras propriedades referentes ao projecto podem ser alteradas, tais como a privacidade da pasta onde se encontra o *web site* - “Directory Security”. Esta *tab* indica se utilizadores anónimos podem aceder ao *web site*. A *tab* “Custom Errors”, por exemplo, lida com os erros mais comuns: 500, 401, etc.

O ficheiro de configuração do *site* é o *web.config* que contém todas as regras de configuração da aplicação ASP.NET num ficheiro de extensão .xml. Ao se utilizar o ficheiro xml não é necessário recompilar a aplicação caso se faça alguma mudança pois, a própria framework detecta a mudança no ficheiro xml e reinicia a aplicação *web*. O ficheiro *web.config* contém as directivas de implementação, guarda os valores e parâmetros (informações úteis) que sejam comuns em toda aplicação ASP.NET. Pode-se por exemplo, indicar o tipo de cultura, o tipo de ligação ao servidor, as páginas de erro, fazer *trace*, indicar os vários modos de compilação, etc. O tipo de ligação ao servidor é bastante importante pois a mesma aplicação *web* pode-se ligar a vários clientes e servidores diferentes. Nos exemplos seguintes estão as principais modificações efectuadas no ficheiro *web.config*.

```
<!--Ligação aos Dados -->
<connectionStrings>
  <add name="BreastCancerDBConnectionString1"
    connectionString="Data Source=AMIGRID\;Initial Catalog=BC;
    Integrated Security=True"
    providerName="System.Data.SqlClient"/>
</connectionStrings>
```

A entrada no exemplo anterior, começa por definir o nome da ligação à base de dados “BreastCancerDBConnectionString1”. Em seguida indica o nome do servidor ao qual se pretende ligar no SQL Server “AMIGRID\”. Depois de efectuada a ligação ao servidor, é indicado o nome da base de dados onde se encontram as tabelas com as informações acerca dos utilizadores “BC”.

```
<!--Mandar Email -->
<system.net>
  <mailSettings><smtp><network host="mail.uninova.pt"/></smtp></mailSettings>
</system.net>
```

A entrada anterior “mail.uninova.pt” diz respeito às definições de envio de emails. Como o servidor do *web site* está localizado dentro do domínio da Uninova, a configuração e a transferência de emails é tratada pelo servidor SMTP da Uninova.

```
<membership defaultProvider="CustomizedMembershipProvider">
  <providers>
    <add connectionStringName="BreastCancerDBConnectionString1"
      name="CustomizedMembershipProvider" applicationName="/"
      type="System.Web.Security.SqlMembershipProvider"
      requiresQuestionAndAnswer="false"
      requiresUniqueEmail="true"
      minRequiredPasswordLength="3"
      minRequiredNonalphanumericCharacters="0"/>
  </providers>
</membership>
```

A entrada “CustomizedMembershipProvider” está associada ao registo de utilizadores no *web site*. Por *default*, o componente “CreateUserWizard” da plataforma ASP.NET pede que a *password* contenha no mínimo 6 letras e para contornar esta requisito definiu-se a *password* para 3 letras. Para além desta modificação, também se retirou a “pergunta”, caso um utilizador se esqueça da sua *password* e também se retiraram os “caracteres obrigatórios” (@,%,&, etc). Tornou-se um requisito ter um email único na base de dados.

```
<system.web>
<identity impersonate="true" userName="CompName\UserName" password="****"/>
</system.web>
```



A entrada “impersonate” é bastante importante para o funcionamento do algoritmo de prognóstico feito em Matlab. Durante a execução do algoritmo no servidor IIS, este vai necessitar de criar, editar e apagar alguns ficheiros temporários de extensão .mat numa directoria com acesso reservado no disco C:\. É preciso dar um acesso especial aos utilizadores autenticados pelo *site* temporariamente. Além do “impersonate”, aos utilizadores autenticados pelo *web site* é preciso adicionar alguns privilégios de rede ao servidor IIS, sendo assim possível escrever no disco C:\, nomeadamente:

1. Abrir a administração IIS.
2. Na pasta onde se encontra os sites no IIS, localizar o *site BCSurvival*.
3. Botão direito do rato → “Permissões”.
4. Clicar no “Adicionar botão” → “Localizações”. Indicar o nome do PC.
5. Na caixa de texto disponível procurar por “NETWORK SERVICE” ou “Serviço de Rede” e adicionar. Nas permissões dar a permissão para escrever.

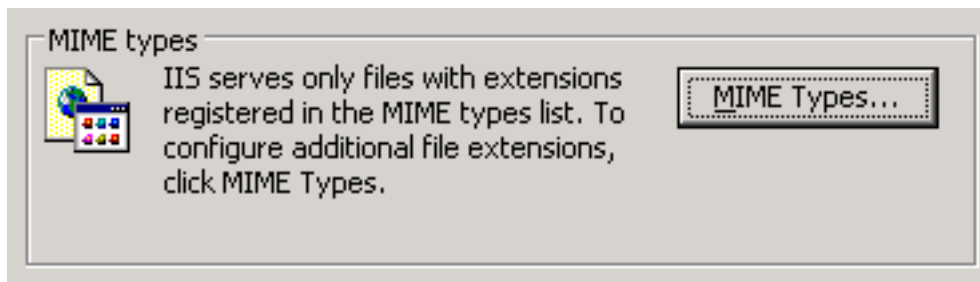
Para além das modificações no servidor IIS é necessário também, no SQL Server 2005, adicionar o “Serviço de Rede” de igual modo como foi feito no IIS, para se ter acesso à base dados “BC” (Breast Cancer).

As *passwords* dos utilizadores registados no *site* estão encriptadas na base de dados e por isso, mesmo que sejam visualizadas por um administrador de base dados, este nunca poderá saber o valor real da *password*.

Para correr o algoritmo em Matlab, que é executado pelas ferramentas de prognóstico do *web site*, é necessário instalar a biblioteca MWArray.dll (é instalada quando se instala o Matlab ou o MCR - MATLAB Component Runtime). Esta biblioteca encarrega-se das classes em .NET associadas ao tratamento de vectores em Matlab.

Para correr as aplicações Silverlight, é obrigatório registar no servidor IIS o tipo MIME do Silverlight, pois de outra forma o servidor não envia a aplicação Silverlight ao cliente (*web browser* do utilizador). A figura 4.15 apresenta um exemplo para adicionar um novo tipo de

MIME, nas propriedades do servidor IIS 5.0. Para tal, em MIME Types adicionar um novo



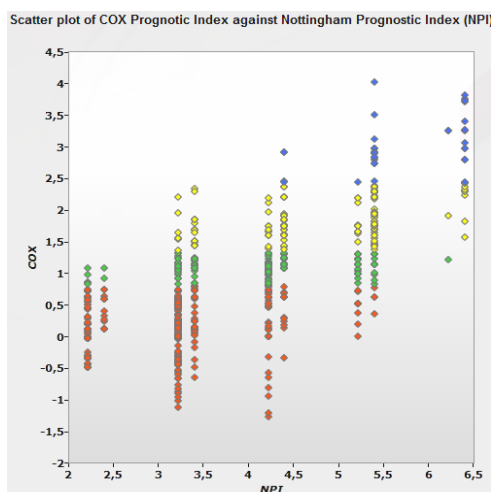
**Figura 4.15:** Adição de um novo tipo de MIME

tipo com extensão .xap e com o nome application/x-silverlight-app.

De forma a obter um domínio gratuito com o nome *BCSurvival*, efectuou-se um registo no *site*: <http://www.pt.vu/step1.php>. O URL do *site* passou a ser o seguinte: <http://bcsurvival.pt.la/>

#### 4.3.4 Sistema de suporte à decisão

A figura 4.16 representa a interface em 2D, desenvolvida em Silverlight, da aplicação referida na secção 4.2.8 a correr no lado do cliente (*browser*) e que é responsável pela localização dos diferentes grupos de risco e índices de prognóstico.



**Figura 4.16:** Aplicação Silverlight

Na navegação por um *site*, existem normalmente os pedidos efectuados pelo *browser* e as respostas por parte do servidor onde está alojado o *site*. O problema é que esta troca de mensagens obriga a uma constante actualização dos dados da página (*refresh*), ou seja, acontece

o efeito de “aparecer e desaparecer” da página, obrigando o *browser* a pedir constantemente ao servidor a aplicação Silverlight.

Para evitar este efeito cria-se uma ponte entre o *browser* e o servidor, através do JavaScript. O Javascript, por sua vez, vai tratar de comunicar com o servidor IIS (mais concretamente com o ASP.NET) através dos *Web Services*, enviando o modelo de prognóstico que se quer usar e determinadas características de um paciente como a idade, tamanho do tumor, etc.

O servidor IIS vai processar o pedido e caso o modelo de prognóstico seja PLANN-ARD, encaminha o pedido com as respectivas variáveis do paciente ao Matlab. O Matlab responde com o índice de risco, sobrevivência e intervalos de confiança associados ao modelo PLANN-ARD, à framework.NET através da dll PLANN.NET criada para esse efeito.

A figura 4.17 representa o diagrama de funcionamento do *web site - BCSurvival*. Entre os factores que contribuíram para a escolha da tecnologia ASP.NET, na construção do *web site*, destacam-se a excelente biblioteca para o suporte de *Web Services*, a facilidade de criar aplicações Silverlight e de integração da framework.NET (C#) com o Matlab.

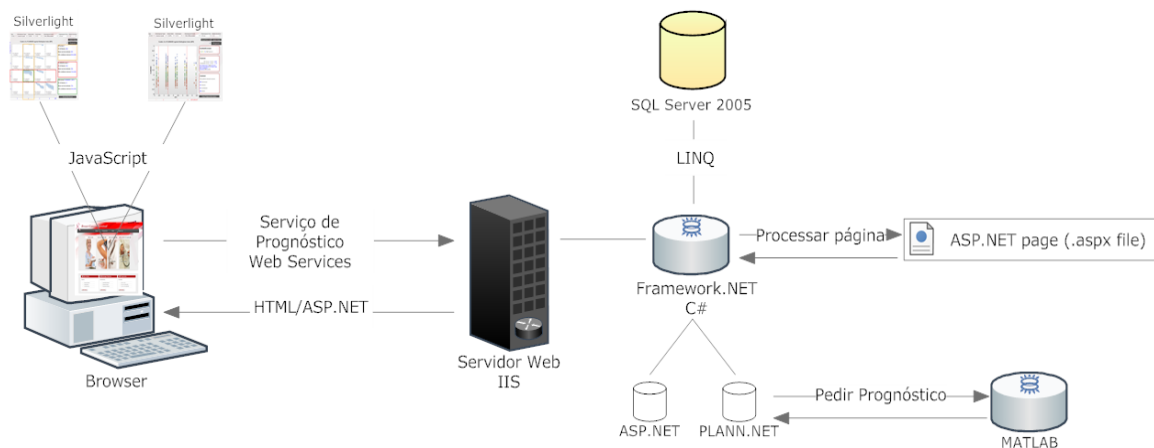


Figura 4.17: Diagrama do *site - BCSurvival*



## WEBSITE - *Breast Cancer Survival*,

### Funcionalidades

#### 5.1 Introdução

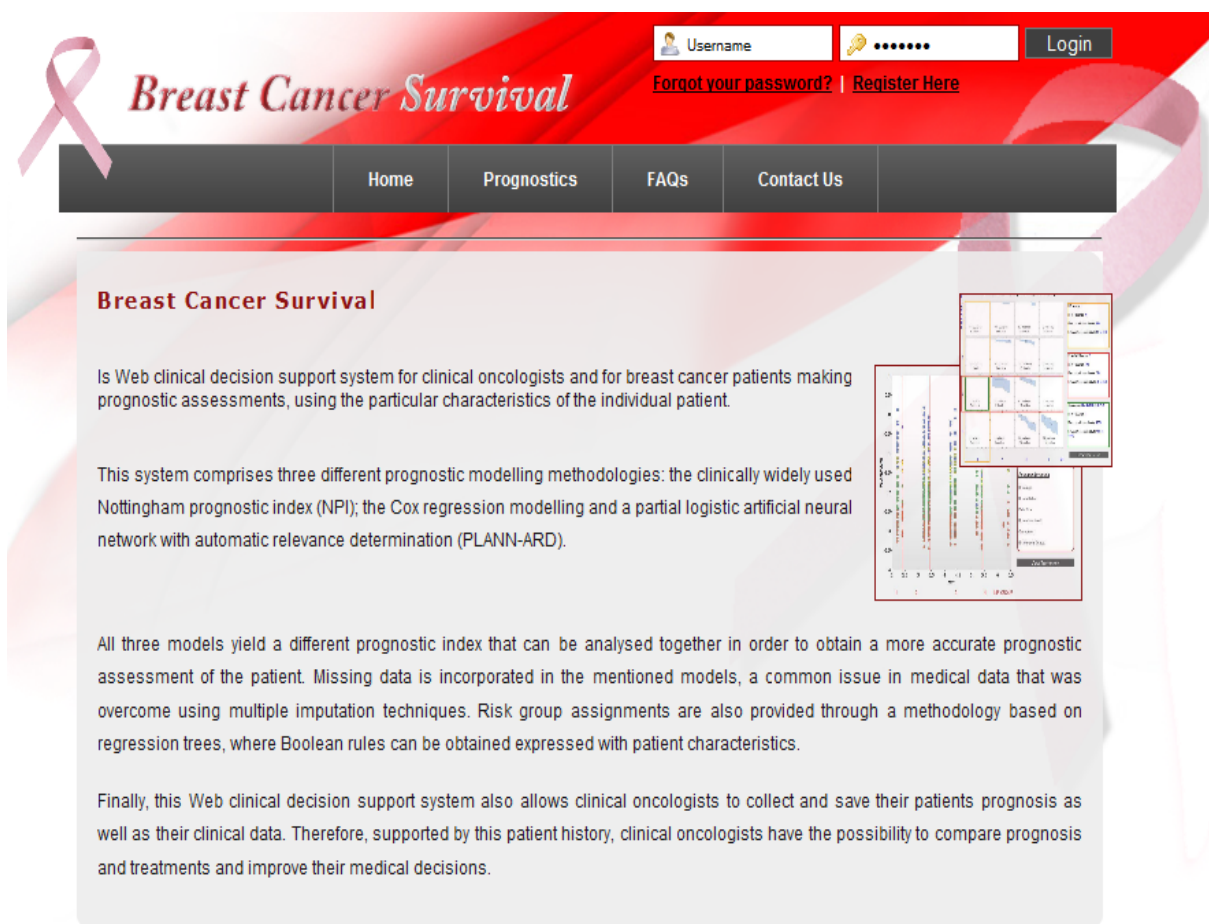
Este capítulo tem como objectivo identificar as funcionalidades do *web site - Breast Cancer Survival*, apresentando as suas características de utilização, nomeadamente as diversas ferramentas de prognóstico do cancro da mama e da gestão de pacientes assim como a disponibilização dos seus prognósticos.

Para melhorar a interactividade com os utilizadores, foi usada a biblioteca de jQuery e os módulos de ASP.NET Ajax anteriormente mencionados no capítulo 4. Ambos permitem maior produtividade, facilidade e rapidez de actualização das páginas web. Em suma, este capítulo pretende detalhar todas as funcionalidades existentes no *web site - BCSurvival*.

## 5.2 Funcionalidades do Site - *BCSurvival*

### 5.2.1 Home Page

O acesso à plataforma que integra o sistema de apoio à decisão é realizado através da internet no *link* <http://bcsurvival.pt.1a/> sem que seja necessário qualquer registo do utilizador. Ao aceder ao URL da aplicação, é apresentada no *browser* a *home page* da plataforma (figura 5.1), onde constam informações relativas ao sistema de apoio à decisão.



**Figura 5.1:** Home page do site - *BCSurvival*

## 5.2.2 Login



Figura 5.2: Area de *Login*.

Ao aceder à *home page* o utilizador além de ter acesso à informação disponível no *web site* tem a possibilidade de aceder a uma área reservada para utilizadores. Para entrar na área reservada é necessário efectuar o login (figura 5.2). Caso o utilizador não tenha ainda efectuado o registo é necessário aceder à página de registo (figura 5.3) e efectuar o preenchimento dos campos que são apresentados, sendo alguns destes campos de carácter obrigatório (*username*, *email*, *password*). No fim do registo, é enviado ao administrador do *site* um *email* com a indicação que foi efectuado um registo.

No caso do utilizador perder o *username* ou a *password* de acesso ao *site*, basta aceder ao link “Forgot your password?” e entrar na página de recuperação (figura 5.4) que oferece um módulo interactivo e apelativo para ambos.

Figura 5.3: Página de registo.

Figura 5.4: Página de recuperação.

### 5.2.3 Menu de Topo



Figura 5.5: Menu de Topo.

O menu de topo (figura 5.5) permite o acesso a algumas funcionalidades:

- “*Home page*”: acesso à página principal do *web site* (figura 5.1).
- “Prognostics”: acesso rápido aos vários modelos de prognóstico. Esta opção está apenas disponível para utilizadores registados.
- “FAQ”: acesso às perguntas mais frequentes, onde os utilizadores podem tirar as suas dúvidas (figura 5.6).
- “Contact Us”: contactar os administradores do *web site* (figura 5.7).

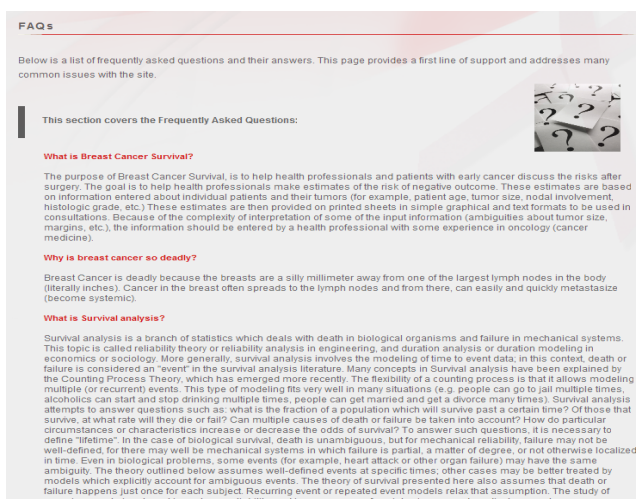


Figura 5.6: Página de FAQ.

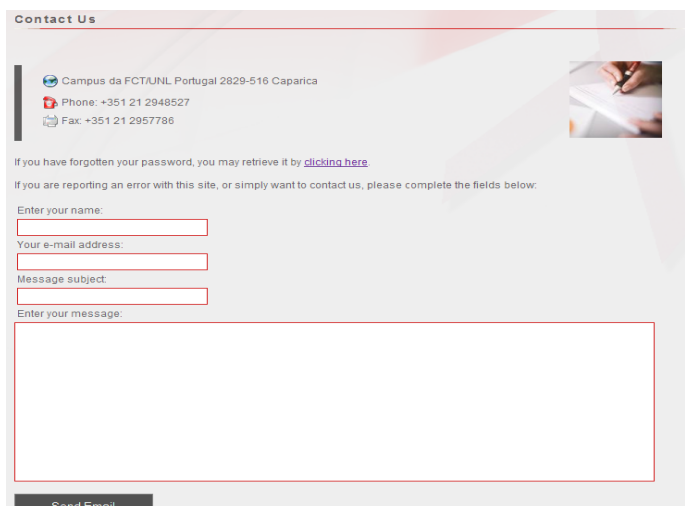


Figura 5.7: Página de Contactos.



## 5.2.4 Menu de Utilizador

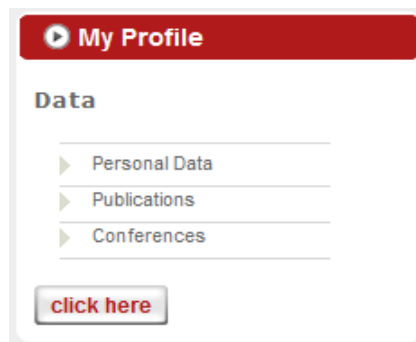


Figura 5.8: Menu de Utilizador.

A figura 5.8 apresenta o menu de utilizador. Este menu, após se carregar no *link* “Personal Data”, oferece ao utilizador autenticado pelo *site* a hipótese de editar e visualizar o seu perfil (figura 5.9). Se os campos de carácter obrigatório (*password*, *email*, *first name* e *last name*) não se encontrarem devidamente preenchidos não será possível prosseguir com a actualização dos dados e aparecerão mensagem(s) de aviso onde o(s) erro(s) foram detectado(s). Para além da capacidade de editar e visualizar as suas informações pessoais o utilizador pode adicionar, editar e consultar o seu curriculum (publicações e conferências), como indicado na figura 5.10.

As publicações e conferências são criadas e editadas através do editor “HTMLEditor”. Os diversos botões da ferramenta são usados para tratamento de conteúdos HTML (criar, editar, etc), sendo possível verificar previamente o texto numa página HTML com esta

Figura 5.9: Tab de Informação pessoal.

Figura 5.10: Tab de publicações.

ferramenta [AJAXToolkit, 2010]. Este controlo é disponibilizado num módulo do AJAX Control Toolkit.

### 5.2.5 Menu de Gestão de Pacientes



Figura 5.11: Menu de Gestão de Pacientes.

A figura 5.11 apresenta o menu de gestão de pacientes. Este menu, após se carregar no *link* “Manage Patients”, permite ao utilizador autenticado pelo *site* a hipótese de adicionar, editar e gerir pacientes, possibilitando também a verificação dos prognósticos dos pacientes (figuras 5.12, 5.13 e 5.14).

A *tab* **Insert/Edit Patient**, indicada na figura 5.12, está dividida em duas secções: a secção “Add Patient” e a secção “Edit Patient”.

#### Add Patient

A cada paciente adicionado está associado um *name*, um *email* e um *comment*. Estes campos são de preenchimento obrigatório à excepção do campo *comment*. O botão “Add Patient” permite ao utilizador adicionar um paciente que, depois de efectuada operação de inserção, irá aparecer na *dropbox* “Select a Patient” juntamente com os restantes pacientes.

#### Edit Patient

Esta secção permite ao utilizador editar os campos *name*, *email* e *Change Patient’s Comment* do paciente seleccionado na lista de pacientes. As informações adicionadas ou

alteradas nesta página web são posteriormente integradas com a base de dados.

Realça-se que a página *web* para adicionar/editar pacientes possui um mecanismo de validação dos dados, ou seja, se os campos não estiverem devidamente preenchidos irá aparecer uma mensagem de aviso.

The screenshot displays a web interface titled "Patients" with three tabs: "Insert/Edit Patient" (active), "Manage Patients", and "Prognosis Patients".

The "Add Patient" section includes:

- An "Add Patient" button with a plus icon and a person icon.
- Input fields for "Name:" and "Email:".
- A "Comment on the Patient:" label above a large text area.
- An "Add Patient" button.

The "Edit Patient" section includes:

- An "Edit Patient" button with a person icon.
- A "Select a Patient:" dropdown menu with "Bilu" selected.
- Input fields for "Change Name:" and "Change Email:".
- A "Change Patient's Comment:" label above a large text area.
- A "Save Changes" button.

**Figura 5.12:** *Tab* para adicionar e editar pacientes.

A *tab* **Manage Patients**, indicada na figura 5.13, apresenta numa lista todos os pacientes que estão associados a um utilizador. A lista de pacientes contém, para além do nome, um comentário associado, a data em que foram inseridos na base de dados do utilizador, uma opção para enviar *emails* ao paciente e uma opção para apagar o paciente. Esta *tab* de gestão de pacientes permite ao utilizador gerir os seus pacientes possibilitando a inserção de comentários, envio de *emails*, etc, oferecendo assim uma visão global de todos os pacientes já adicionados.

A interactividade existente na janela de envio de *emails* e os efeitos de aparecer e desaparecer janelas estão a cargo da biblioteca jQuery, como referido no capítulo anterior e referenciado em [jQuery, 2010].

No início da página existe uma área de pesquisa por paciente ou pacientes. Por exemplo, se se inserir a palavra “Ana”, todos os pacientes associados a este utilizador que tenham Ana no nome irão aparecer na lista final.

The screenshot shows a web interface titled "Patients" with three tabs: "Insert/Edit Patient", "Manage Patients" (selected), and "Prognosis Patients". Below the tabs is a "Manage Patients" section with a "Search Users:" input field and a magnifying glass icon. A navigation bar contains letters A through Z. The main content area is titled "All my Patients" and "Patient: Katia". It features a table with the following data:

| Patient                           | Comment       | Date       |  |  |
|-----------------------------------|---------------|------------|--|--|
| <a href="#">Ana Pena</a>          | Pena do Magno | 23-09-2009 |  |  |
| <a href="#">Tomas</a>             | Aderiu hj     | 22-09-2009 |  |  |
| <a href="#">Catarina Alves</a>    | desculpa      | 15-09-2009 |  |  |
| <a href="#">Ana Catarina Pena</a> | Com gripe A   | 14-09-2009 |  |  |
| <a href="#">Antonio Janota</a>    | Ola           | 14-09-2009 |  |  |
| <a href="#">Bilu</a>              | Sem gripe A   | 14-09-2009 |  |  |
| <a href="#">Katia</a>             | n tem gripe A | 14-09-2009 |  |  |
| <a href="#">Ana Sofia</a>         | EUA           | 14-09-2009 |  |  |
| <a href="#">Jamal</a>             |               | 14-09-2009 |  |  |
| <a href="#">Luis Figo</a>         | Com gripe A+  | 12-09-2009 |  |  |

At the bottom of the interface, there is a "Comment on: Katia" section with a text input field and a "Save Comment" button.

Figura 5.13: Tab para gerir os pacientes.

A *tab* **Prognosis Patients**, indicada na figura 5.14, apresenta um resumo dos diversos prognósticos dos pacientes. Para cada paciente existe um quadro onde constam diversas informações (variáveis e valores de prognóstico). Esta *tab* permite ao utilizador analisar a evolução histórica dos pacientes existentes na base de dados. A lista de pacientes contém, para além do nome, a data em que o prognóstico foi efectuado, sendo que esta é listada por ordem decrescente. Na lista de prognósticos, entre a data e o ícone de remover existe o ícone de acesso rápido para fazer um prognóstico, ou caso o utilizador pretenda, também tem a possibilidade de apagar o prognóstico.

Tal como na *tab* “Manage Patients”, no início da página existe uma forma de pesquisa por paciente ou pacientes.

The screenshot displays the 'Prognosis Patients' interface. At the top, there are navigation tabs: 'Insert/Edit Patient', 'Manage Patients', and 'Prognosis Patients'. Below the tabs is a search bar labeled 'Search Users:'. A navigation menu shows letters A through Z. The main content is divided into two sections: 'Prognosis History' and 'Prognosis Details of: Ana Pena'.

**Prognosis History Table:**

| Patient                        | Date       | Quick Action | Remove |
|--------------------------------|------------|--------------|--------|
| <a href="#">Ana Pena</a>       | 14-10-2009 |              |        |
| <a href="#">Ana Sofia</a>      | 23-09-2009 |              |        |
| <a href="#">Ana Sofia</a>      | 23-09-2009 |              |        |
| <a href="#">JamaI</a>          | 23-09-2009 |              |        |
| <a href="#">Ana Pena</a>       | 23-09-2009 |              |        |
| <a href="#">Ana Pena</a>       | 23-09-2009 |              |        |
| <a href="#">Tomas</a>          | 22-09-2009 |              |        |
| <a href="#">Tomas</a>          | 22-09-2009 |              |        |
| <a href="#">Catarina Alves</a> | 22-09-2009 |              |        |
| <a href="#">Bilu</a>           | 22-09-2009 |              |        |

**Prognosis Details of: Ana Pena**

|                   |                             |
|-------------------|-----------------------------|
| Date              | 14-10-2009 14:07:00         |
| Age               | 20-39                       |
| HistologicalType  | Invasive ductal             |
| NodeRatio         | 0-20%                       |
| Oestrogen         | 0-10                        |
| HistologicalGrade | Moderately differentiated   |
| PathologicalSize  | 2-5 cm                      |
| NodesInvolved     | 1-3                         |
| Prognosis         | NPI: 3, COX: 2, PLANNARD: 0 |
| Comment           |                             |

At the bottom, there is a comment box with the text 'Comment on the patient's prognosis: Ana Pena' and a 'Save Comment' button.

Figura 5.14: Tab para gerir os prognósticos dos pacientes.

### 5.2.6 Menu de Prognósticos

Os diferentes modelos de prognóstico combinados com os métodos de estratificação descritos no capítulo 3, foram integrados na plataforma web de suporte à decisão para clínicos e pacientes, por forma a identificar mais precisamente o grupo de risco de cada paciente e assim, adequar os tratamentos existentes.

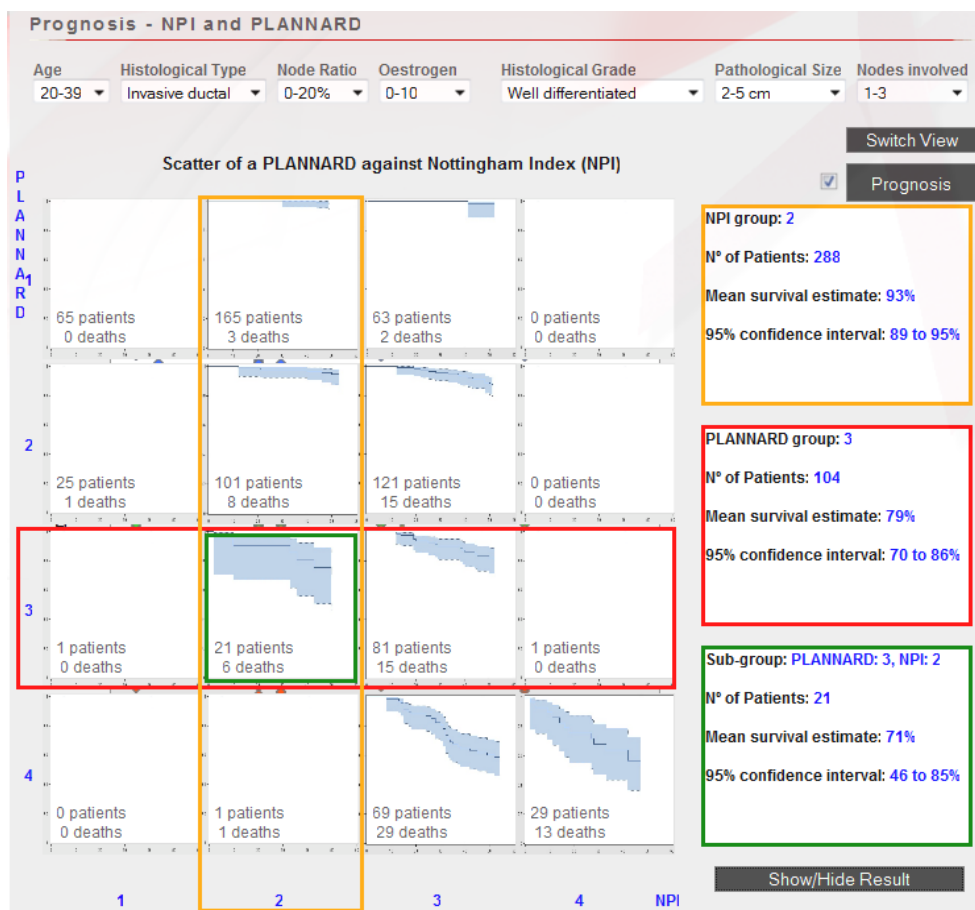


**Figura 5.15:** *Home page* de prognósticos para os utilizadores autenticados.

A figura 5.15 representa a *Home page* de prognósticos para utilizadores autenticados no *web site* com a possibilidade de escolher uma visualização 2D para a localização dos diferentes grupos de risco. Nestas visualizações os diferentes modelos de prognóstico (NPI, COX e PLANN-ARD) são gerados computacionalmente sendo o resultado obtido em cada uma destas metodologias associado a um paciente.

A finalidade principal deste menu é assistir e ajudar os médicos e pacientes a identificar o grupo de risco para determinado paciente de cancro da mama, através do cruzamento das curvas de sobrevivência obtidas nas diferentes metodologias utilizadas. Através desta interface (figura 5.15), pode-se seleccionar a comparação entre os modelos de prognóstico pretendidos, nomeadamente o índice de NPI com dois modelos estatísticos Cox Model (modelo linear) versus PLANN-ARD (modelo não linear).

A figura 5.16 apresenta um exemplo da aplicação onde se compara o índice NPI com o modelo de prognóstico PLANN-ARD.



**Figura 5.16:** Grupo de risco do modelo PLANN-ARD e grupo de risco do NPI.

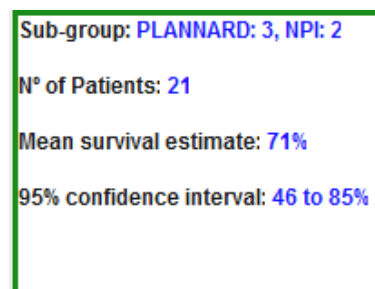
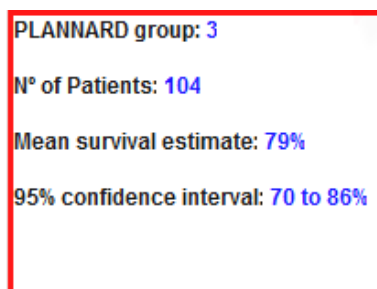
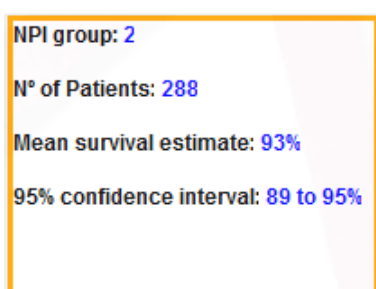
Na figura 5.16 é possível observar as 16 curvas de sobrevivência de *Kaplan-Meier* para os diferentes subgrupos de risco (PLANN-ARD *versus* NPI) além da informação associada ao paciente e ao grupo de risco correspondente. Esta figura representa a coordenação entre os 4 grupos de risco para o NPI, enquanto a abcissa representa os 4 grupos de risco para o PLANN-ARD. As 16 curvas de *Kaplan-Meier* são calculadas da seguinte forma:

1. Para cada um dos 743 pacientes do *Christie Hospital* são obtidos dois índices de risco: um utilizando o modelo de prognóstico PLANN-ARD e outro utilizando o NPI. Após calculados os 2 índices, é identificado para cada modelo o grupo de risco onde cada paciente se encontra. Os grupos encontrados para cada modelo são depois cruzados permitindo identificar assim o sub-grupo de risco.

2. Para cada subgrupo de pacientes constrói-se a curva de sobrevivência de *Kaplan-Meier*, através dos pacientes da base de dados que constituem esse sub-grupo.

Os componentes do lado direito da figura 5.16 permitem ao utilizador ter uma maior precisão da sobrevivência do paciente. Após introduzir as características do paciente (idade, tamanho do tumor, etc) é devolvido pelo sistema: um índice de NPI (3.400) e um grupo de risco (2) associado a esse modelo e um índice de PLANN-ARD (-1.137) com um grupo de risco (3).

Isto indica que:



Para o grupo NPI 2, existem 288 pacientes do *Christie Hospital* com características semelhantes à do paciente introduzido e que estes têm uma estimativa média de sobrevivência de 93% com um intervalo de confiança a 95%, a variar entre 89% e 95%.

Para o grupo PLANN-ARD 3, existem 104 pacientes do *Christie Hospital* com características semelhantes à do paciente introduzido e que estes têm uma estimativa média de sobrevivência de 79% com um intervalo de confiança a 95%, a variar entre 70% e 86%.

Para o sub-grupo PLANN-ARD 3 e NPI 2, existem 21 pacientes do *Christie Hospital* com características semelhantes à do paciente introduzido e que estes têm uma estimativa média de sobrevivência de 71% com um intervalo de confiança a 95%, a variar entre 46% e 85%.

Ao clicar sobre uma imagem da matriz (figura 5.16) é possível visualizá-la mais pormenorizadamente (figura 5.17). As curvas de sobrevivência permitem identificar a heterogeneidade dentro de um sub-grupo de risco de prognóstico.

A figura 5.18 ilustra um gráfico de dispersão que permite comparar os índices de prognóstico obtidos através de diferentes metodologias possibilitando ao utilizador ter uma visão mais geral do índice de risco a que o paciente está sujeito. Cada eixo representa uma metodologia, sendo neste caso específico, o NPI e o PLANN-ARD. Este gráfico permite verificar qual o grupo de risco a que cada paciente pertence com um determinado índice de risco, para ambas as metodologias. Além disso tem também a capacidade de informar o utilizador se os novos



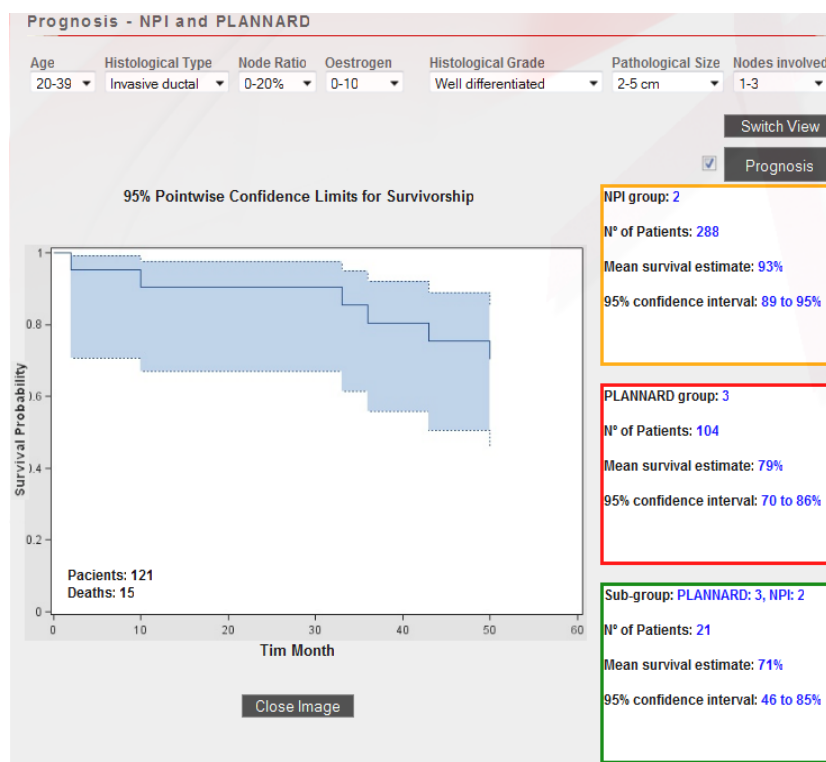
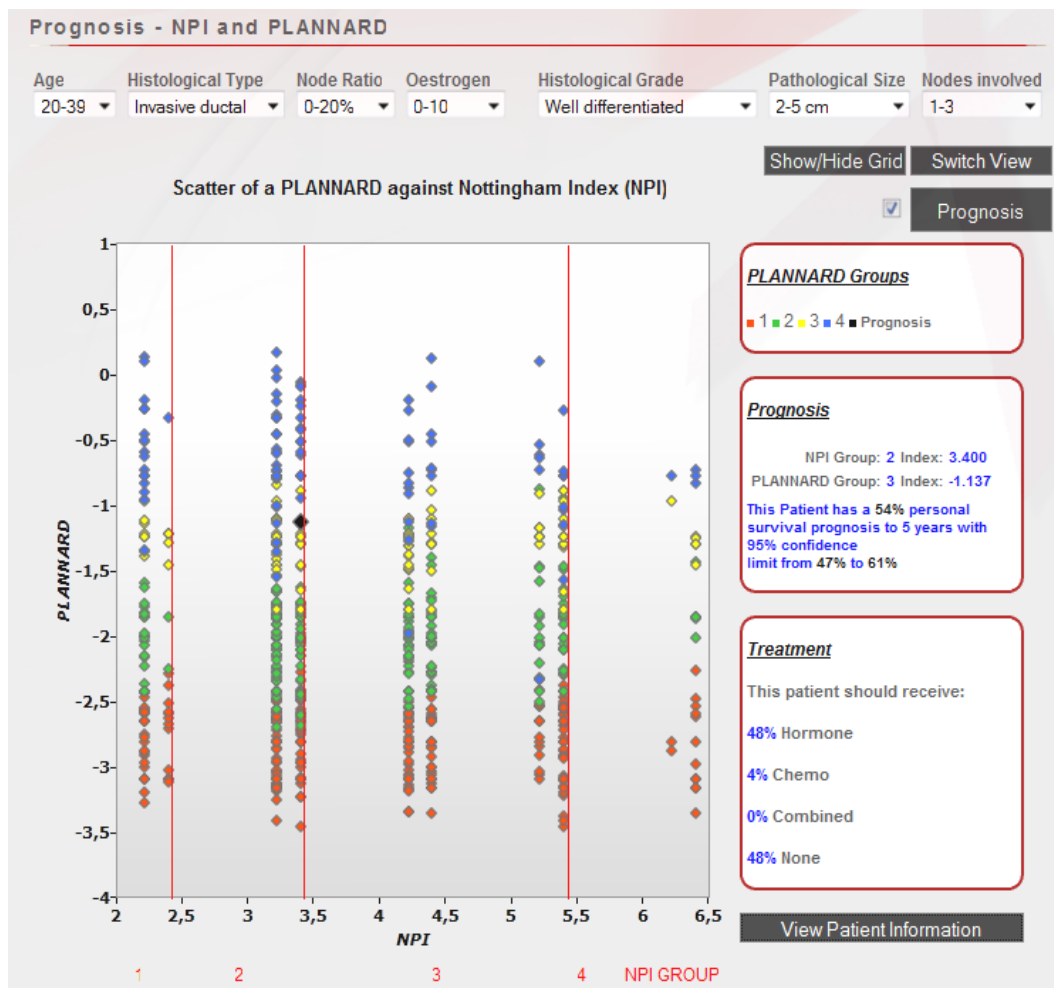


Figura 5.17: Curvas de sobrevivência de *Kaplan Meier*

pacientes para os quais se pretende obter o prognóstico estão afastados do modelo ou na fronteira entre grupos de risco.

As características do paciente (*Age, Histological Type, Node Ratio, Oestrogen, Histological Grade, Pathological Size, Nodes involved*) são introduzidas no sistema e os valores de risco de sobrevivência, os grupos de risco e os tratamentos aconselhados são calculados e devolvidos para o *browser* do utilizador (figura 5.18). As características possíveis de serem introduzidas pelo utilizador são consideradas como as mais predictivas de prognóstico, não sendo permitido a possibilidade de colocar “missing”. Caso houvesse a possibilidade de colocar “missing”, o sistema poderia informar o médico que não existiam na base de dados pacientes com as mesmas características introduzidas, o que poderia induzir, erradamente, o médico na interpretação do resultado.

O resultado devolvido pelo sistema de prognóstico pode ser visualizado num gráfico 2D. A figura 5.18 representa o cruzamento entre o resultado do risco clínico NPI e o índice de prognóstico derivado utilizando o PLANN-ARD. Quando calculado pelo NPI o paciente é



**Figura 5.18:** Interface indicando o índice de risco PLANN-ARD e o índice de risco NPI.

associado ao grupo de risco 2, mas utilizando o PLANN-ARD para o cálculo do prognóstico, o paciente pertence ao grupo de risco 3.

Em relação ao tipo de tratamento, o sistema informa o médico que os pacientes com características idênticas têm a probabilidade de seguir um plano de tratamento de terapêutica hormonal de 48%, quimioterapia de 4%, os dois tratamentos em combinado de 0% e não seguir qualquer tratamento de 48%. Em termos de prognóstico, o sistema notifica o utilizador de que o paciente em avaliação tem um prognóstico individual caracterizado por uma probabilidade de 54% de sobrevivência a 5 anos, com um intervalo de confiança a 95%, ou seja, que pode variar entre 47% e 61%. O cálculo deste prognóstico para um paciente é explicado no capítulo 3 na secção 3.5.3.

### 5.3 Comparação com outros sistemas de identificação de risco de cancro da mama

The Finprog Study - a Nationwide Breastcancer Database

URL: <http://www.finprog.org/>

O *site* **Finprog Study** oferece ferramentas que auxiliam os médicos na avaliação do risco de uma possível reincidência da patologia de cancro da mama. Estas ferramentas permitem fazer previsões individuais de sobrevivência. As características do paciente são introduzidas no sistema e as curvas de sobrevivência ou as curvas de sobrevivência com impacto dos tratamentos são devolvidas para o *browser* do utilizador. O perfil do novo paciente é correlacionado com os existentes na base de dados de cancro da mama finlandesa. Caso haja indivíduos com prognóstico semelhante, para este novo grupo é calculada uma nova curva de sobrevivência, sendo esta posteriormente apresentada no *site* no formato de imagem.

O *site* tem como principais funcionalidades:

- Para cada variável de prognóstico identifica a sua importância no impacto da sobrevivência no cancro da mama.
- O sistema de cálculo de sobrevivência verifica se na base de dados existe algum indivíduo com características idênticas. Se existir, mostra a sobrevivência dessa pessoa e o número total de pacientes com o mesmo perfil. Se não existir, indica que a sobrevivência é de 100% (isto é pouco fiável, pois seria mais correcto afirmar “não há dados disponíveis que permitam inferir sobre a taxa de sobrevivência do paciente”).
- A cada categoria das variáveis é apresentado o valor da sua significância para a sobrevivência.
- Possui vários formatos de apresentação: *Bar chart*, *Pie chart*, etc.

### 5.3. COMPARAÇÃO COM OUTROS SISTEMAS DE IDENTIFICAÇÃO DE RISCO DE CANCRO DA MAMA

---

- Indica o número de pacientes que estão vivos de ano para ano nas curvas de sobrevivência.
- Apresenta resultados do NPI e de outros 9 testes.
- Possui intervalos de confiança, que não se encontram no gráfico, mas numa tabela separada.
- Permite comparar 2 curvas de sobrevivência.

#### **CancerMath.net**

**URL:** <http://www.lifemath.net/cancer/>

O *site* **CancerMath.net** permite calcular o risco e a sobrevivência para um determinado paciente. Consoante as características do paciente introduzidas no sistema, o impacto dos diferentes tratamentos na sobrevivência e risco de morte do paciente são calculados e devolvidos para o *browser* do utilizador.

O *site* tem como principais funcionalidades:

- Apresenta um gráfico de tecnologia Adobe Flash com a evolução das diferentes curvas de sobrevivência para os pacientes com cancro da mama ou cancro mais efeito da terapia, etc.
- Podem ser introduzidas no sistema diferentes terapias.
- Calcula as curvas de sobrevivência para diferentes cancros: Pele, Cólon, Mama, Renal.
- Compara o tempo esperado de vida de um paciente, com ou sem cancro da mama, para a mesma idade.
- Apresenta os anos que um paciente, com determinado prognóstico de cancro da mama, perdeu devido ao cancro.

- Apresenta os efeitos em tempo de vida que um paciente tem, se levar a cabo determinado tratamento.
- Apresenta os resultados em vários formatos: *Bar chart*, *Pie chart*, etc.
- Tem outras ferramentas de interesse para o cancro da mama, por exemplo: cálculo do efeito da terapia na sobrevivência, cálculo da probabilidade de haver *positive nodes* e de *nipple involvement*.

**Risk Assessment Tool:**

URL: <http://www.cancer.gov/bcrisktool/>

O *site* **The Breast Cancer Risk Assessment Tool** fornece uma ferramenta que tem como público alvo mulheres com historial de cancro da mama na família. O risco obtido pela ferramenta não deve ser levado em consideração, caso já tenha sido diagnosticado o cancro da mama na mulher. Ou seja, o objectivo desta ferramenta é estimar o risco de vir a desenvolver cancro da mama. Para isso compara o perfil introduzido no sistema pelo utilizador, com o das mulheres existentes na base de dados.

O *site* tem como principais funcionalidades:

- Para cada variável de prognóstico exhibe uma breve explicação sobre a importância que esta tem no impacto da sobrevivência na mulher.
- A ferramenta de risco é baseada num modelo estatístico conhecido por “*Gail model*”.
- O modelo foi testado em mulheres de raça caucasiana. Em mulheres afro-americanas o risco calculado não foi bem aceite. O modelo ainda precisa de ser testado em mulheres asiáticas.
- Permite fazer o download do programa para dispositivos móveis.

### 5.3. COMPARAÇÃO COM OUTROS SISTEMAS DE IDENTIFICAÇÃO DE RISCO DE CANCRO DA MAMA

---

- A ferramenta apresenta como resultados: a probabilidade de desenvolver cancro da mama no período de 5 anos e a mesma probabilidade num período de vida até aos 90 anos.

#### **American Cancer Society:**

**URL:** <http://www.cancer.org/>

O *site* **American Cancer Society** tem a particularidade de “exigir” ao utilizador o preenchimento de um formulário extensivo sobre os sintomas de cancro da mama que um paciente possa ter. Isto faz todo o sentido, uma vez que os médicos oncologistas do cancro da mama encontram-se envolvidos numa nova fase de gestão. Detalhar ao máximo a patologia permite oferecer ao paciente um maior leque de opções de tratamento da sua doença. Isto é válido tanto para um bom como para um mau prognóstico.

O *site* tem como principais funcionalidades:

- Partilha de experiências sobre o cancro da mama.
- Permite entender melhor o diagnóstico, os resultados dos testes efectuados, e o efeito dos tratamentos na qualidade de vida do paciente.
- O registo para utilizar a ferramenta.
- Permite fazer prognósticos para outros cancros.
- Um glossário para melhor interpretação dos conceitos.
- Artigos e livros de apoio à explicação do cancro da mama, consoante o resultado do questionário que nos propusermos.

#### **Adjuvant! Online**

**URL:** <http://www.adjuvantonline.com/index.jsp>

O site **Adjuvant! Online** tem como públicos alvo os médicos e os pacientes do cancro da mama. Pretende discutir e informar os diferentes riscos e benefícios, na aplicação de diferentes terapias nos pacientes com cancro da mama, após a realização da cirurgia tais como: terapia de quimioterapia, terapia hormonal ou as duas aplicadas simultaneamente.

O site oferece como funcionalidades:

- Ajuda os profissionais de saúde a efectuar estimativas de risco de morte devido ao cancro da mama, com e sem o efeito da aplicação de terapias na sobrevivência dos doentes.
- Permite tirar dúvidas relacionadas com o cancro da mama.
- Explica as estimativas de risco, sobrevivência e os tratamentos.
- Apresenta imagens esquemáticas do tecido mamário, antes e após a cirurgia do cancro da mama.
- Dispõe de ferramentas para outros tipos de cancro.
- Permite o download do programa para um PDA.
- Disponibiliza um gráfico em Java Applet do tipo *Bar chart* com o estado de sobrevivência para um paciente, com os efeitos das diferentes terapias, etc.

A tabela 5.1 sistematiza as funcionalidades dos vários sistemas de identificação de risco de cancro da mama *online* referidos ao longo da secção 5.3, permitindo a sua comparação. A partir da tabela 5.1 é possível verificar que o site proposto para dissertação, o BCSurival expõe como principais pontos fortes, os seguintes:

- Incorpora três modelos de prognóstico num único sistema de visualização.
- Utiliza uma plataforma de visualização 2D em Silverlight para apresentação dos índices de risco de uma população obtida do *Christie Hospital* em Wilmslow, Manchester, no

### 5.3. COMPARAÇÃO COM OUTROS SISTEMAS DE IDENTIFICAÇÃO DE RISCO DE CANCRO DA MAMA

---

Reino Unido no período entre 1990-94. A escolha desta tecnologia, para além de *freeware*, permite a incorporação de bibliotecas gráficas em 2D de forma acessível e de fácil manutenção. Ao contrário dos outros *sites*, o *BCSurvival* é o único que permite a visualização geral do risco dos 743 pacientes obtidos pela base de dados do *Christie Hospital*, assim como, a visualização do risco em 2D para um novo paciente.

- As diferentes curvas de *Kaplan-Meier*, introduzidas para visualização da probabilidade de sobrevivência do paciente, constituem uma melhoria relativamente aos outros *sites*, sendo uma mais-valia na interpretação dos resultados obtidos. Da mesma forma, os intervalos de confiança associados a estas curvas permitem dar uma maior ajuda à tomada de decisão por parte do médico, oferecendo uma maior ou menor incerteza subjacente aos resultados obtidos. Ficando a decisão final a cargo do médico.
- Esta plataforma é a única que permite aos utilizadores reunir e guardar os dados e resultados dos prognósticos dos pacientes.
- Os tratamentos sugeridos pela plataforma têm em conta os tratamentos administrados numa população de 743 pacientes.



Tabela 5.1: Sistemas de identificação de risco de cancro da mama *online*

|                                 | Finprog | CancerMath  | RiskAssessmentTool | AmericanCancerSociety | Adjuvant     | BCSurvival  |
|---------------------------------|---------|-------------|--------------------|-----------------------|--------------|-------------|
| Curvas de <i>Kaplan-Meier</i>   | S       | S           | N                  | N                     | N            | S           |
| Intervalos de Confiança         | S       | N           | N                  | N                     | N            | S           |
| Prognóstico Individual          | N       | S           | S                  | N                     | S            | S           |
| Tratamentos                     | S       | S           | N                  | S                     | S            | S           |
| Tecnologia de Apresentação      | Img     | Adobe Flash | N/D                | N/D                   | Java applets | Silverlight |
| Disponível para PDA             | N       | N           | S                  | N                     | S            | N           |
| Gestão de Pacientes             | N       | N           | N                  | N                     | N            | S           |
| Ferramenta para outros Cancros  | N       | S           | N                  | S                     | S            | N           |
| Informações sobre o Cancro      | N       | N           | S                  | S                     | S            | N           |
| Informação sobre as variáveis   | S       | N           | S                  | S                     | S            | N           |
| Aceitação pelos <i>Browsers</i> | Alta    | Média       | Alta               | Alta                  | Média        | Média       |
| Necessário registo              | N       | N           | N                  | S                     | S            | S           |



## Conclusões

### 6.1 Discussão dos Resultados

Esta dissertação teve o intuito de criar um sistema *online* de apoio à decisão para médicos oncologistas e pacientes de cancro da mama. O sistema desenvolvido incorpora três modelos de prognóstico: NPI, o Regressão de Cox e o PLANN-ARD. Para as três metodologias é identificado o nível de risco a 5 anos, intervalos de confiança e grupos de risco. A plataforma, além de fornecer um sistema de apoio à decisão, permite também aos utilizadores reunir e arquivar os dados e os resultados dos prognósticos. A arquitectura desenvolvida cumpre os objectivos a que se propôs.

A capacidade de prever um acontecimento com um certo grau de sucesso, comparando o resultado obtido com o resultado esperado, faz dos modelos de prognóstico uma mais valia para médicos oncologistas na tomada de decisão do tratamento a aplicar a um paciente. Utilizados apropriadamente, estes modelos ajudam os médicos na decisão clínica, facilitada pela estratificação das classes de risco e na atitude terapêutica. Uma grande preocupação dos médicos está relacionada com a interpretação da natureza da “caixa preta” (*black box*), como é o caso das redes neuronais, em fazer inferências de sobrevivência para um paciente [Lisboa, 2002]. O uso de árvores de decisão, nomeadamente de regressão, no âmbito da área de saúde permite ao médico interpretar os resultados, uma vez que explica, de uma

## 6.1. DISCUSSÃO DOS RESULTADOS

---

forma simples, as regras para cada grupo de risco tendo em conta as variáveis e os índices de prognóstico.

A atribuição de um grupo de prognóstico a um paciente de cancro da mama é de extrema importância, pois esta reflecte-se na estratégia terapêutica a atribuir ao paciente. Por exemplo um paciente com um tumor superior a 5 cm tem poucas hipóteses de sobreviver, como tal, é indispensável um tratamento agressivo.

Os avanços na terapia e nas tecnologias de detecção do cancro da mama, aliado às novas políticas de saúde nos países, como o rastreio de cancro da mama na população e a procura de novos indicadores de prognóstico, permitiu aumentar o poder preditivo do NPI. Novas discussões de ordem ética acerca do uso do NPI têm surgido, como a capacidade de este aconselhar a mulher a engravidar ou não, consoante a sua sobrevivência [Galea et al., 1992], ou o seu uso na triagem clínica, nomeadamente na determinação do estado de gravidade dos doentes, sendo estes tratados por ordem dessa gravidade.

Foram utilizadas as seguintes tecnologias: ASP.NET, Silverlight, AJAX, JavaScript, Matlab entre outras, para a implementação da plataforma *web*, com vista a otimizar a navegação no *site* e no sentido de oferecer todas as funcionalidades disponíveis tendo em conta os pré-requisitos.

Nos últimos anos, têm sido desenvolvidos esforços no sentido de tornar possível a conversão dos seus ficheiros Matlab de modo a que estes possam ser escaláveis para outras plataformas como a framework.NET. Esta evolução possibilita aproveitar melhor esta capacidade de *software* permitindo por exemplo, a criação de soluções *online*.

Procurar, pormenorizar e descrever com uma qualidade superior os prognósticos associados aos grupos de risco, recorrendo para isso a diferentes metodologias foram os objectivos da plataforma construída nesta dissertação. Desta forma evitou-se a substituição das metodologias actualmente em uso.

## 6.2 Trabalho Futuro

Sendo esta plataforma dirigida para a área de saúde, haverá sempre funcionalidades, procedimentos e metodologias médicas que poderão ser integradas para tornar o sistema de apoio à decisão criado mais funcional e composto. Algumas funcionalidades a integrar no sistemas poderiam ser, por exemplo, esquemáticos no formato de imagem, permitindo a visualização antes e após a remoção da mama, um painel informativo com diversas explicações detalhadas sobre o cancro da mama, variáveis de prognóstico, respectivos tratamentos e cirurgia. Com a existência de um administrador do *web site*, seria interessante terminar a parte administrativa do mesmo. O administrador ficaria encarregue da gestão das *mailing lists*, aprovação de novos utilizadores, introdução de notícias sobre o cancro da mama entre outras tarefas. A disponibilização dos resultados de prognóstico para o paciente, no formato PDF seria interessante, caso o paciente quisesse utilizar os resultados num centro de oncologia.

Uma vez que o sistema possui uma base de dados *online* com todos os prognósticos e variáveis dos pacientes, seria possível implementar um sistema de tratamento estatístico dos dados obtidos. A organização estatística teria em conta as variáveis de prognóstico como o tamanho do tumor, idade, entre outras. A apresentação da informação seria realizada através de um quadro histórico da evolução das variáveis de prognóstico, apresentadas num gráfico de barras ao longo do tempo. Por exemplo, num determinado paciente com um tumor de aproximadamente 2 cm seria seguido e determinado, ao longo do tempo, o crescimento do tumor. Esta evolução da doença em determinado instante poderia ser importante para verificar padrões de comportamento do tumor.

Visto que as tecnologias de suporte para *web sites* estão em constante evolução, o próximo passo para um trabalho futuro seria o de migrar a framework 3.5 do sistema de decisão em ASP.NET para a versão 4 e, também, migrar a base de dados SQL 2005 para o SQL 2008, de modo a usufruir da mais recente tecnologia.

Um outro aspecto importante, dando uma outra perspectiva de visualização do índice de risco ao médico, seria a introdução de gráficos 3D. Embora haja esforços por parte da

## 6.2. TRABALHO FUTURO

---

comunidade em torno da criação de modelos 3D em Silverlight e Adobe Flash, estas ainda se encontram num estado embrionário.

O ritmo da evolução tecnológica permite aos médicos flexibilizar cada vez mais a escolha da melhor alternativa para ir de encontro às suas necessidades. A tecnologia móvel está cada vez mais presente com a crescente popularidade do uso de PDA's, iPad e telemóveis. Fazer a implementação do sistema de decisão num dispositivo móvel garantia ao médico a capacidade de fazer avaliações de prognóstico e de tratamento a qualquer momento e em qualquer lugar do mundo independentemente de estar ou não *online*.

# Apêndice A

Regras para o modelo *Cox Proportional Hazard*

**Tabela A.1:** Regras obtidas pelo CART para o modelo *Cox Proportional Hazard*

|         |          |   |
|---------|----------|---|
| Grupo 1 | Regra 1  | Pathsize=1 e Noderatio=1 e Histype=2 e Histgrade=1 ou 2                                 |
|         | Regra 2  | Pathsize=1 e Noderatio=1 e Histype=1 e Histgrade=1                                      |
|         | Regra 3  | Pathsize=1 e Noderatio=1 e Histype=1 e Histgrade=2 e Age =2                             |
|         | Regra 4  | Pathsize=1 e Noderatio=1 e Histype=1 e Histgrade=2 e Age=3 e Oestrogen=2                |
|         | Regra 5  | Pathsize=1 e Noderatio $\neq$ 1 e Histype=2 e Histgrade=1 ou 2 e Age=1 ou 2             |
|         | Regra 6  | Pathsize=1 e Noderatio $\neq$ 1 e Histype=1 e Histgrade=1 e Age=1 ou 2                  |
|         | Regra 7  | Pathsize=2 e Noderatio=1 e Histgrade=1 ou 2 e Histype=2 e Age=2                         |
|         | Regra 8  | Pathsize=2 e Noderatio=1 e Histgrade=1 ou 2 e Histype=2 e Age=1 ou 3 e Oestrogen=2      |
|         | Regra 9  | Pathsize=2 e Noderatio=1 e Histgrade=1 e Histype=1 e Oestrogen=2                        |
| Grupo 2 | Regra 1  | Pathsize=1 e Noderatio=1 e Histype=1 e Histgrade=2 e Age=1 e Oestrogen=2                |
|         | Regra 2  | Pathsize=1 e Noderatio=1 e Histype=1 e Histgrade=2 e Age=1 ou 3 e Oestrogen=1           |
|         | Regra 3  | Pathsize=1 e Noderatio $\neq$ 1 e Histype=2 e Histgrade=1 ou 2 e Age=3                  |
|         | Regra 4  | Pathsize=1 e Noderatio $\neq$ 1 e Histype=1 e Histgrade=1 e Age=3                       |
|         | Regra 5  | Pathsize=1 e Histgrade=3 e Noderatio=1  |
|         | Regra 6  | Pathsize=1 e Histgrade=2 e Noderatio $\neq$ 1 e Histype=1 e Age=1 ou 3                  |
|         | Regra 7  | Pathsize=1 e Histgrade=2 e Noderatio=2 ou 3 e Histype=1 e Age=2                         |
|         | Regra 8  | Pathsize=2 e Noderatio=1 e Histgrade=1 ou 2 e Histype=2 e Age=1 ou 3 e Oestrogen=1      |
|         | Regra 9  | Pathsize=2 e Noderatio=1 e Histgrade=1 e Histype=1 e Oestrogen=1                        |
|         | Regra 10 | Pathsize=2 e Noderatio=1 e Histgrade=2 e Histype=1                                      |
|         | Regra 11 | Pathsize=2 e Noderatio=1 e Histgrade=3 e Age=2  |
|         | Regra 12 | Pathsize=2 e Noderatio=1 e Histgrade=3 e Age=1 ou 3 e Oestrogen=2                       |
|         | Regra 13 | Pathsize=2 e Noderatio=2 ou 3 e Histgrade=1 ou 2 e Histype=2                            |
|         | Regra 14 | Pathsize=2 e Noderatio=2 ou 3 e Histgrade=1 ou 2 e Histype=1 e Age=2                    |
| Grupo 3 | Regra 1  | Pathsize=1 e Histgrade=3 e Noderatio $\neq$ 1   |
|         | Regra 2  | Pathsize=1 e Histgrade=2 e Noderatio=4 e Histype=1 e Age=2                              |
|         | Regra 3  | Pathsize=2 e Noderatio=1 e Histgrade=3 e Age=1 ou 3 e Oestrogen=1                       |
|         | Regra 4  | Pathsize=2 e Noderatio=4 e Histype=2  |
|         | Regra 5  | Pathsize=2 e Noderatio=3 e Histgrade=3  |
|         | Regra 6  | Pathsize=2 e Noderatio=2 ou 3 e Histgrade=1 ou 2 e Histype=1 e Age=1 ou 3 e Oestrogen=2 |
| Grupo 4 | Regra 1  | Pathsize=2 e Noderatio=2 ou 3 e Histgrade=1 ou 2 e Histype=1 e Age=1 ou 3 e Oestrogen=1 |
|         | Regra 2  | Pathsize=2 e Noderatio=2 e Histgrade=3  |
|         | Regra 3  | Pathsize=2 e Noderatio=4 e Histype=1  |





# Apêndice B

Regras para o modelo *PLANN-ARD*

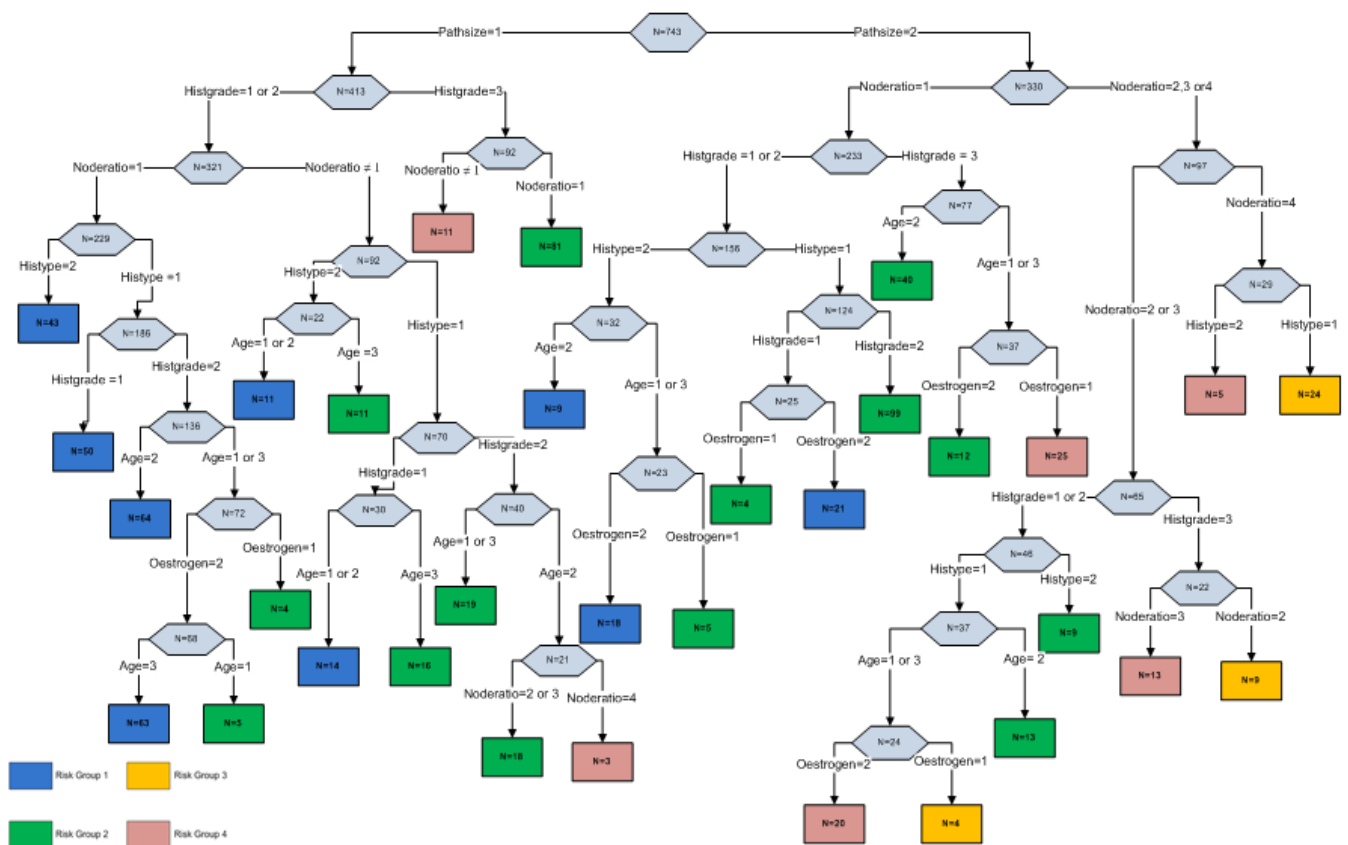
**Tabela B.1:** Regras obtidas pelo CART para o modelo *PLANN-ARD*

|         |          |   |
|---------|----------|---|
| Grupo 1 | Regra 1  | Pathsize=1 e Noderatio=1 e Age=2 e Histgrade=1 ou 2                                     |
|         | Regra 2  | Pathsize=1 e Noderatio=1 e Histype=2 e Histgrade=1 ou 2 e Age=3                         |
|         | Regra 3  | Pathsize=1 e Noderatio=1 e Histype=1 e Histgrade=1 e Age=3 e Oestrogen=2                |
|         | Regra 4  | Pathsize=1 e Noderatio=1 e Histype=1 e Histgrade=2 e Age=3                              |
|         | Regra 5  | Pathsize=1 e Noderatio=2 ou 3 e Histype=2 e Histgrade=1 ou 2 e Age=2                    |
|         | Regra 6  | Pathsize=1 e Noderatio=2 ou 3 e Histype=1 e Histgrade=1 e Age=2                         |
|         | Regra 7  | Pathsize=1 e Noderatio=1 e Histgrade=3 e Oestrogen=2 e Age=2                            |
|         | Regra 8  | Pathsize=2 e Noderatio=1 e Histgrade=1 ou 2 e Histype=2 e Age=2                         |
|         | Regra 9  | Pathsize=2 e Noderatio=1 e Histgrade=1 ou 2 e Histype=2 e Age=1 ou 3 e Oestrogen=2      |
|         | Regra 10 | Pathsize=2 e Noderatio=1 e Histgrade=1 e Oestrogen=2 e Age=2 e Histype=1                |
| Grupo 2 | Regra 1  | Pathsize=1 e Noderatio=1 e Histgrade=1 ou 2 e Age=1                                     |
|         | Regra 2  | Pathsize=1 e Noderatio=1 e Histype=1 e Histgrade=1 e Age=3 e Oestrogen=1                |
|         | Regra 3  | Pathsize=1 e Noderatio=2 ou 3 e Histype=2 e Histgrade=1 ou 2 e Age=1 ou 3               |
|         | Regra 4  | Pathsize=1 e Noderatio=2 ou 3 e Histype=1 e Histgrade=1 e Age=1 ou 3                    |
|         | Regra 5  | Pathsize=1 e Noderatio=2 ou 3 e Histype=1 e Histgrade=2                                 |
|         | Regra 6  | Pathsize=1 e Histgrade=3 e Noderatio=1 e Age=2 e Oestrogen=1                            |
|         | Regra 7  | Pathsize=1 e Histgrade=3 e Noderatio=1 e Age=3  |
|         | Regra 8  | Pathsize=2 e Noderatio=1 e Histgrade=1 ou 2 e Histype=2 e Age=1 ou 3 e Oestrogen=1      |
|         | Regra 9  | Pathsize=2 e Noderatio=1 e Histgrade=1 ou 2 e Histype=1 e Oestrogen=1 e Age=2           |
|         | Regra 10 | Pathsize=2 e Noderatio=1 e Histgrade=2 e Histype=1 e Age=2 e Oestrogen=2                |
|         | Regra 11 | Pathsize=2 e Noderatio=1 e Histgrade=3 e Age=2 e Oestrogen=2                            |
|         | Regra 12 | Pathsize=2 e Noderatio=1 e Histgrade=1 ou 2 e Age=3 e Oestrogen=2 e Histype=1           |
|         | Regra 13 | Pathsize=2 e Noderatio=2 ou 3 e Histgrade=1 ou 2 e Histype=2                            |
| Grupo 3 | Regra 1  | Pathsize=1 e Histgrade=3 e Noderatio=1 e Age=1  |
|         | Regra 2  | Pathsize=2 e Histgrade=1 ou 2 e Noderatio=1 e Histype=1 e Age=1 ou 3 e Oestrogen=1      |
|         | Regra 3  | Pathsize=2 e Histgrade=3 e Noderatio=1 e Age=1 ou 3 e Oestrogen=2                       |
|         | Regra 4  | Pathsize=2 e Histgrade=1 ou 2 e Noderatio=1 e Histype=1 e Age=1 e Oestrogen=2           |
|         | Regra 5  | Pathsize=2 e Noderatio=1 e Histgrade=3 e Age=2 e Oestrogen=1                            |
|         | Regra 6  | Pathsize=2 e Noderatio=2 ou 3 e Histgrade=1 ou 2 e Age=2 e Histype=1                    |
|         | Regra 7  | Pathsize=2 e Noderatio=2 ou 3 e Histgrade=1 ou 2 e Age=3 e Oestrogen=2 e Histype=1      |
| Grupo 4 | Regra 1  | Pathsize=1 e Histgrade=1 ou 2 e Noderatio=4   |
|         | Regra 2  | Pathsize=1 e Noderatio=2 ou 3 ou 4 e Histgrade=3  |
|         | Regra 3  | Pathsize=2 e Noderatio=4  |
|         | Regra 4  | Pathsize=2 e Noderatio=2 ou 3 e Histgrade=3   |
|         | Regra 5  | Pathsize=2 e Noderatio=2 ou 3 e Histgrade=1 ou 2 e Histype=1 e Age=1 ou 3 e Oestrogen=1 |
|         | Regra 6  | Pathsize=2 e Noderatio=2 ou 3 e Histgrade=1 ou 2 e Histype=1 e Age=1 e Oestrogen=2      |
|         | Regra 7  | Pathsize=2 e Noderatio=1 e Histgrade=3 e Age=1 ou 3 e Oestrogen=1                       |



# Apêndice C

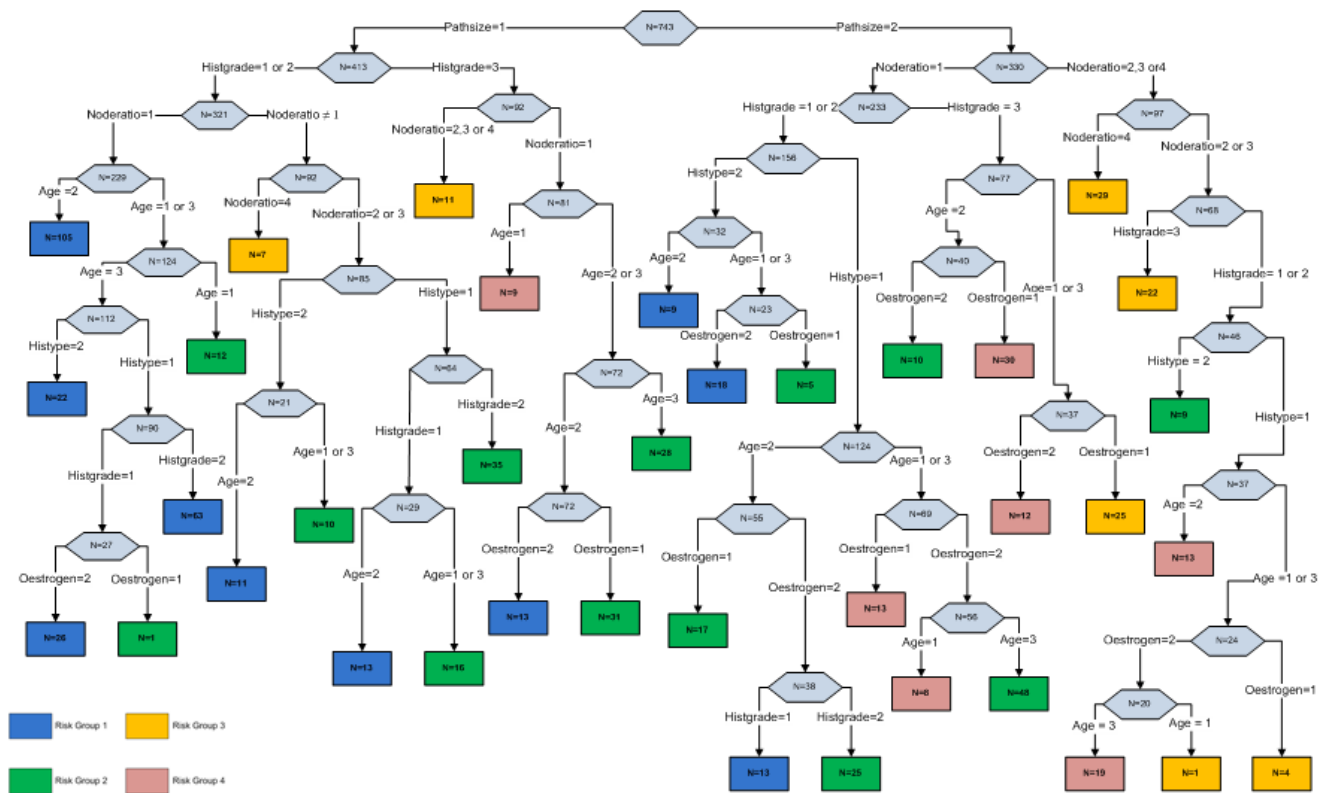
Árvore de decisão para o modelo *Cox Proportional Hazards*





# Apêndice D

Árvore de decisão para o modelo *PLANN-ARD*





# Bibliografia

- [Abreu and Morgado, 2009]Abreu, L. and Morgado, P. (2009). *LINQ com C#*, pages 7–14. Novatec Editora, first edition.
- [Altman and Wyatt, 1995]Altman, D. and Wyatt, J. (1995). Commentary: Prognostic models: clinically useful or quickly forgotten? *BMJ*, 311:1539–1541.
- [Biganzoli et al., 1998]Biganzoli, E., Boracchi, P., Mariani, L., and E.Marubini (1998). Feed forward neural networks for the analysis of censored survival data: a partial logistic regression approach. *Stat Med*, 17:1169–86.
- [Bishop, 1995]Bishop, C. (1995). *Neural networks for pattern recognition*. Oxford: Clarendon Press.
- [Breiman et al., 1984]Breiman, L., Friedman, J., Olsen, A., Stone, and J, C. (1984). Classification and regression trees. *The Wadsworth Brooks*.
- [Christensen, 1987]Christensen, E. (1987). Multivariate survival analysis using cox’s regression model. *Hepatology*, 7(6):1346–987.
- [Clark and Altman, 2003]Clark, T. and Altman, D. (2003). Developing a prognostic model in the presence of missing data an ovarian cancer case study. *Journal of clinical epidemiology*, 56(2):28–37.

- [Clark et al., 2003]Clark, T., Bradburn, M., Love, S., and Altman, D. (2003). Survival analysis part i: basic concepts and first analyses. *Br J Cancer*, 89(2):232–8.
- [Collect, 1994]Collect, D. (1994). *Modelling survival data in medical research*. London: Chapman & Hall.
- [Colosimo and Giolo, 2006]Colosimo, E. A. and Giolo, S. R. (2006). *Análise de sobrevivência aplicada*. ABE - Projeto Fisher.
- [Cox, 1972]Cox, D. R. (1972). Regression models and life-tables. *Journal of the Royal Statistical Society*, 34(2):187–220.
- [Duckett, 2008]Duckett, J. (2008). *Beginning Web Programming with HTML, XHTML, and CSS*. Wiley.
- [Evjen et al., 2008]Evjen, B., Hanselman, S., and Rader, D. (2008). *Professional ASP.NET 3.5 In C# and VB*. Wiley.
- [Fernandes et al., 2009]Fernandes, A., Etchells, T., Jarman, I., Fonseca, J., Elia, B., Chris, B., and Lisboa, P. (2009). Stratification methodologies for neural networks models of survival. *Proceedings of the 10th International Work-Conference on Artificial Neural Networks*, 5517:989–996.
- [Fernandes et al., 2008]Fernandes, A., Jarman, I., Etchells, T., Fonseca, J., Biganzoli, E., Bajdik, C., and Lisboa, P. (2008). Missing data imputation in longitudinal cohort studies - application of plann-ard in breast cancer survival. *Seventh International Conference on Machine Learning and Applications*, ICMLA08:644–649.
- [Fernandes et al., 2010]Fernandes, A. S., Alves, P., Jarman, I. H., Etchells, T. A., Fonseca, J. M., and Lisboa, P. J. (2010). A clinical decision support system for breast cancer patients. *Springer*, pages 122–129.



- [Fonseca et al., 2006]Fonseca, J. M., Mora, A. D., and Barroso, P. (2006). *THE WEB AND THE NEW GENERATION OF MEDICAL INFORMATION*. Elsevier, Studies in Multidisciplinary series Section 5, Chapter 14.
- [Galea et al., 1992]Galea, M. H., Blarney, R. W., Elston, C. E., and Ellis, I. O. (1992). The nottingham prognostic index in primary breast cancer. *Breast Cancer Research and Treatment*, pages 22:207–219.
- [Haybittle et al., 1982]Haybittle, J., Blamey, R., Elston, C., Johnson, J., Doyle, P., Campbell, F., Nicholson, R., and Griffiths, K. (1982). A prognostic index in primary breast cancer. *Brit J Cancer*, 45 3621.
- [Jarman et al., 2008]Jarman, I., Lisboa, P., and Etchells, T. (2008). An integrated framework for risk profiling of breast cancer patients following surgery. *Elsier*, 42(3):165–188.
- [Joe and Tomson, 2002]Joe, M. and Tomson, B. (2002). *Sams Teach Yourself ASP.Net in 24 Hours*. Pearson Education, Inc.
- [Kanjilal and Putrevu, 2009]Kanjilal, J. and Putrevu, S. (2009). *Sams Teach Yourself ASP.NET Ajax in 24 Hours*, pages 3–9. Water Crest, first edition.
- [Kaplan, 1958]Kaplan, E. (1958). Meier p. nonparametric estimation from incomplete observations. *Journal of the American Statistical Association*, 53:457–81.
- [Lisboa, 2002]Lisboa, P. (2002). A review of evidence of health benefit from artificial neural networks in medical intervention. *Neural Networks*, 15:9–37.
- [Lisboa et al., 2003]Lisboa, P., Wong, H., Harris, P., and Swindell, R. (2003). A bayesian neural network approach for modelling censored data with an application to prognosis after surgery for breast cancer. *Elsevier*, 706:1–26.

- [Myers et al., 1966]Myers, M. H., Axtell, L. M., and Zelen, M. (1966). The use of prognostic factors in predicting survival for breast cancer patients. *J. Chron. Dis.*, 19(923).
- [Pereira and Poupa, 2005]Pereira, A. and Poupa, C. (2005). *Linguagens WEB*. Edições Sílabo, lda.
- [Ravdin et al., 1992]Ravdin, P., Clark, F., Hilsenbeck, S., Owens, M., Vendely, P., and Pandian, M. (1992). A demonstration that breast cancer recurrence can be predicted by neural network analysis. *Breast Cancer Res Treatment*, 21:47–53.
- [Rebasa, 2005]Rebasa, P. (2005). Basic concepts in survival analysis. *Cir Esp*, 78(4):222–30.
- [Silva, 2007]Silva, M. S. (2007). *Construindo sites com CSS e (X)HTML*, pages 49–53. Novatec Editora, first edition.
- [Silva, 2008]Silva, M. S. (2008). *jQuery - A Biblioteca do Programador JavaScript*. Novatec Editora.
- [Swift et al., 2009]Swift, J., Barker, C., Wahlin, D., and Patuel, S. A. (2009). *Professional Silverlight 2 for ASP.NET Developers*, pages 3–5. Wiley, first edition.
- [Taktak et al., 2007]Taktak, A., Antolini, L., Aung, M., Boracchi, P., Campbell, I., Damato, B., Ifeachor, E., Lama, N., Lisboa, P., Setzkorn, C., Stalbovskaya, V., and Biganzoli, E. (2007). Double-blind evaluation and benchmarking of survival models in a multi-centre study. *Elsier*, 37:1108–1120.

## Referências Online

[AJAXToolkit, 2010]AJAXToolkit, H. P. (última consulta Março 2010).

. <http://www.asp.net/AJAX/AjaxControlToolkit/Samples/Default.aspx>.

[ASP.NET, 2010]ASP.NET, H. P. (última consulta Março 2010).

. <http://eu.dummies.com/how-to/content/discovering-visual-basic-nets-internet-technology-html>.

[Battles, 2010]Battles, H. P. (última consulta Outubro 2010).

. <http://www.backendbattles.com/backend/jquery>.

[CSSLibrary, 2010]CSSLibrary, H. P. (última consulta Janeiro 2010).

. <http://www.dynamicdrive.com/style/csslibrary/>.

[CSSMenuMaker, 2010]CSSMenuMaker, H. P. (última consulta Janeiro 2010).

. <http://www.cssmenuaker.com/>.

[DCarterDesign, 2010]DCarterDesign, H. P. (última consulta Março 2010).

. <http://www.dcarter.co.uk>.

[HER2, 2010]HER2, H. P. (última consulta Janeiro 2010).

. <http://www.roche.pt/sites-tematicos/her2/index.cfm/livro/quais-sao-os-tratmentos-disponiveis/>.

[IIS, 2010]IIS, H. P. (última consulta Fevereiro 2010).

. <http://www.iis.net/Overview>.

[jQuery, 2010]jQuery, H. P. (última consulta Outubro 2010).

. <http://jquery.com/>.

[LINQ, 2010]LINQ, H. P. (última consulta Março 2010).

. <http://msdn.microsoft.com/en-us/netframework/aa904594.aspx>.

[MatlabCompiler, 2010]MatlabCompiler, H. P. (última consulta Março 2010).

. <http://www.mathworks.com/products/compiler/description1.html>.

[Oncologia, 2010]Oncologia, H. P. (última consulta Janeiro 2010).

. <http://www.pop.eu.com/portal/publico-geral/tipos-de-cancro/cancro-da-mama/o-cancro-da-mama>.

[PortalSaúde, 2010]PortalSaúde, H. P. (última consulta Janeiro 2010).

. <http://www.portaldasaude.pt/portal/conteudos/enciclopedia+da+saude/doencas/>

[RoundedCorner, 2010]RoundedCorner, H. P. (última consulta Janeiro 2010).

. <http://www.roundedcornr.com/>.

[Silverlight, 2010]Silverlight, H. P. (última consulta Janeiro 2010).

. <http://www.microsoft.com/silverlight/>.

[SQLServer, 2010]SQLServer, H. P. (última consulta Março 2010).

. <http://www.microsoft.com/brasil/servidores/sql/2005/prodinfo/overview/what-is-sql-server.aspx>.

[TechTarget, 2010]TechTarget, H. P. (última consulta Março 2010).

. [http://searchsoa.techtarget.com/sDefinition/0,,sid26\\_gci213352,00.html](http://searchsoa.techtarget.com/sDefinition/0,,sid26_gci213352,00.html).

[W3C, 2010]W3C, H. P. (última consulta Março 2010).

. <http://www.w3.org/TR/html401/>.