

• U



C

• Faculdade de Ciências e Tecnologia
Departamento de Física

Universidade de Coimbra

Instituto das Ciências Nucleares Aplicadas à Saúde – ICNAS
Instituto Biomédico de Investigação da Luz e Imagem – IBILI

Tese de Mestrado Integrado em Engenharia Biomédica

Criação de uma Base de Dados Relativa a Elementos Extraídos de Imagens Médicas

Raquel Trindade

Coimbra, Setembro de 2012

• U



C

• Faculdade de Ciências e Tecnologia

Universidade de Coimbra

Departamento de Física

Instituto das Ciências Nucleares Aplicadas à Saúde - ICNAS
Instituto Biomédico de Investigação da Luz e Imagem - IBILI

Tese de Mestrado Integrado em Engenharia Biomédica

Criação de uma Base de Dados Relativa a Elementos Extraídos de Imagens Médicas

Raquel Trindade

Dissertação para obtenção do grau de Mestre em Engenharia Biomédica

Coimbra, Setembro de 2012

Orientadores:

Eng.º Nuno Chichorro Ferreira

Eng.º Francisco José Caramelo



Esta cópia da tese é fornecida na condição de que quem a consulta reconhece que os direitos de autor são pertença do autor da tese e que nenhuma citação ou informação obtida a partir dela pode ser publicada sem a referência apropriada.

This copy of the thesis has been supplied on condition that anyone who consults it is understood to recognize that its copyright rests with its author and that no quotation from the thesis and no information derived from it may be published without proper acknowledgement.

Este documento não se encontra escrito segundo o novo Acordo Ortográfico da Língua Portuguesa.

O estilo de inserção de referências escolhido para a sua escrita foi o definido pelo IEEE, estando a Bibliografia organizada mediante a ordem de referência ao longo do documento.

*“As we acquire more knowledge, things do not become
more comprehensible, but more mysterious.”*

Albert Schweitzer

Agradecimentos

Agradecimentos

Devo agradecer em primeiro lugar aos mentores deste projecto, o Eng.º Nuno Ferreira e o Eng.º Francisco Caramelo que, em primeiro lugar, desenvolveram esta ideia, e depois me apoiaram sempre na concretização das directrizes definidas para a sua implementação e me incentivaram na criação de mais e mais funcionalidades, aperfeiçoando assim a ferramenta desenvolvida. Foram essenciais no que respeita ao processo de detecção das necessidades da instituição onde o trabalho foi concebido, perspicazes na formulação do que deveria ser levado a cabo para satisfazê-las, e nesse sentido sugeriram muitos pormenores importantes para a gestão da Base de Dados implementada.

Este projecto foi desenvolvido em cooperação com o Instituto de Ciências Nucleares Aplicadas à Saúde (ICNAS) e o Instituto Biomédico de Investigação da Luz e Imagem (IBILI), associados à Faculdade de Ciências e Tecnologia da Universidade de Coimbra (FCTUC) e à Faculdade de Medicina da Universidade de Coimbra (FMUC), respectivamente. Quero agradecer a estas instituições as instalações e ferramentas de trabalho cedidas para o desenvolvimento desta tese de mestrado, bem como o suporte financeiro na impressão e publicação dos trabalhos realizados no decorrer deste ano de acolhimento.

Um agradecimento especial à FCTUC pelo apoio financeiro prestado à publicação do artigo “Development of a Medical Image Database for Data Mining and Group Analysis Applications” no âmbito da Conferência HCist, que será apresentado em Outubro no decorrer da Conferência e posteriormente publicado no Jornal “SciVerse Science Direct”, com a colaboração da ELSEVIER® e da Proccedia Technology®.

Aos meus pais e irmãos por todo o apoio que sempre me deram ao longo da minha formação académica, mas principalmente na minha formação pessoal com uma educação baseada no exemplo de moral e humanidade que todos eles são. Obrigada por todo o carinho e paciência.

Aos meus avós que são as melhores pessoas deste mundo, e alegraram sempre os meus fins-de-semana com todo o seu afecto.

À Andreia que me mostrou que a distância não significa nada quando a amizade é verdadeira, e que foi o meu pilar nos momentos mais difíceis.

Às minhas companheiras nesta maratona de escrita, e minhas amigas do coração: Mónica, Daniela, Liliana, Cláudia, Vanessa, Cíntia e Carla. Nunca vos esquecerei e a todos os momentos passados em Coimbra.

A todos os que participaram na minha vida Universitária, tornando-a uma experiência inesquecível. Em especial às minhas maravilhosas companheiras de casa Teresa, Daniela e Marlene, que fizeram do meu terceiro ano o melhor de todos.

A alguém com quem partilhei todos os momentos, bons e maus, mas que garantiu sempre a minha felicidade só pelo simples facto de estar presente.

A todos que mesmo não estando nomeados aqui especificamente, estão dentro do meu coração e fizeram parte do meu crescimento pessoal e profissional nestes 5 anos, ou ao longo de toda a minha existência.

Sumário

Sumário

A prática médica envolve milhares de profissionais de saúde, pacientes, procedimentos, protocolos, exames e tratamentos e cada um deles tem de ser registado, descrito e possui e/ou gera informação própria. Num centro de imagiologia médica são produzidos, diariamente, vários *gigabytes* de informação deste tipo, sendo a maior parte constituída por dados de imagem médica. É necessário armazenar, processar e analisar a informação recolhida o mais rápido e facilmente possível, usando eficientemente os dados disponíveis.

Neste trabalho descreve-se o desenvolvimento de uma Base de Dados de estudos de PET/CT, concebida com o objectivo de permitir não só a análise das imagens em si, mas também da meta-informação a ela associada (armazenada nos cabeçalhos dos ficheiros dos exames) e a inclusão de nova informação gerada a partir do processamento da imagem e de algoritmos de classificação adequados, tratamento estatístico e mineração de dados.

Procedeu-se à recolha de todos os exames deste tipo realizados no ICNAS até ao momento presente, analisou-se a informação que se encontrava estruturada nos seus cabeçalhos, extraiu-se esta informação, e modelou-se uma Base de Dados adequada para recebê-la e permitir a sua modificação, expansão e processamento. Este processo foi realizado recorrendo a ferramentas de modelação (Sybase Power Designer 15[®]) e de geração (PostgreSQL[®] e pgAdmin III[®]) de uma plataforma deste tipo, e as restantes rotinas de extracção, carregamento e processamento de dados foram desenvolvidas na linguagem de programação MATLAB[®], para proceder à gestão da informação.

Criou-se também uma interface gráfica em MATLAB[®] para permitir um fácil acesso à Base de Dados e também às ferramentas de processamento de informação para ela especificamente desenvolvidas.

A plataforma no seu todo servirá também para organizar e integrar a informação ao nível da instituição, e para formação e treino dos investigadores na análise e percepção da imagem médica.

Palavras-chave: Base de Dados, Imagem Médica, PET/CT, Meta-informação, Interface Gráfica com o Utilizador, Processamento de Dados, Integração de Informação.

Abstract

Abstract

The medical practice involves thousands of health professionals, patients, procedures, protocols, treatments and examinations and each one of them has to be recorded, described and possess and/or generate its own information. In a medical imaging center, several gigabytes of this kind of information are daily generated, and the medical image data represents the greater part of it. It is necessary to store, process and analyze the collected information as quickly and easily as possible, using the available information effectively.

In this paper we describe the development of a database of PET/CT studies designed with the goal of allowing not only to analyze the images themselves, but also the metadata associated with it (stored in the examination file headers) and the inclusion of new information generated from image processing and suitable classification algorithms, statistical analysis and data mining.

We proceed to the collection of all of this kind of studies performed in ICNAS until now, analysed the information structured in their headers, extracted this information, and modeled up a database capable of receiving it and to permit its modification, expansion and processing. This process was carried out using tools for modeling (Sybase Power Designer 15[®]) and generating (PostgreSQL[®] and pgAdmin III[®]) such a platform, and the remaining routines for data extraction, upload and data processing were developed in MATLAB[®] programming language, to undertake the information management.

A graphical user interface was also created in MATLAB[®] for enabling an easy access to the database and also to the information processing tools specifically developed for it.

The platform as a whole will also serve to organize and to integrate the information at the institution level, and to the education and training of researchers in the analysis and perception of the medical image.

Keywords: *Database, Medical Image, PET/CT, Metadata, Graphical User Interface, Data Processing, Integration of the Information.*

Índice

Índice

Agradecimentos	V
Sumário	VIII
<i>Abstract</i>	X
Índice	XII
Índice de Figuras	XVI
Índice de Tabelas	XXVI
Acrónimos	XXVIII
Capítulo 1 - Introdução	2
1.1. Contexto/Enquadramento	2
1.2. Objectivo.....	7
1.3. Estado da Arte.....	8
1.4. Organização da Tese	12
Capítulo 2 – Gestão do Projecto	15
2.1. Membros do Projecto	15
2.2. Apresentação das Instituições de Acolhimento.....	15
2.3. Planeamento Inicial vs. Tarefas Delineadas para a sua Execução.....	17
2.3.1. Planeamento Inicial	17
2.3.2. Tarefas Delineadas para a sua Execução	18
Capítulo 3 – Fundamentos Teóricos	21
3.1. Imagem Médica.....	21
3.2. Técnica PET/CT.....	24
3.3. Bases de Dados.....	29
3.3.1. Pessoas envolvidas no Processo de Concepção de uma Base de Dados	35
3.3.2. Processo de Desenvolvimento de uma Base de Dados.....	35
3.3.3. Modelo e Diagrama Entidade-Relacionamento.....	37

3.3.4. Modelo Relacional	39
3.3.5. Modelo Entidade-Atributo-Valor	40
3.3.6. Modo de Armazenamento da Informação de uma Base de Dados	41
3.3.7. Bases de Dados de Imagem Médica.....	42
3.4. Interface com o Utilizador.....	44
Capítulo 4 – Equipamento Utilizado e Dados Recolhidos	48
4.1. Descrição dos Tipos de Ficheiros Gerados em cada Estudo	51
4.1.1. Ficheiros PET	53
4.1.2. Ficheiros TAC	54
Capítulo 5 – Estudo das Variáveis Existentes.....	56
Capítulo 6 – Organização da Informação	59
Capítulo 7 – Método de Implementação da Base de Dados.....	64
7.1. Análise de Requisitos	64
7.2. Casos de Uso	65
7.3. Desenvolvimento da Base de Dados	66
7.3.1. Modelo Conceptual	66
7.3.2. Modelo Físico.....	69
7.3.3. Geração da Base de Dados	70
7.3.4. Carregamento da Base de Dados – Problemas e Restrições.....	71
Capítulo 8 – Desenvolvimento da Interface de Comunicação com a Base de Dados	76
8.1. Aceder e Partilhar Informação	76
8.2. Interface para o Processamento e Análise de Dados	79
8.2.1. Processo de Actualização da Base de Dados.....	82
8.2.2. Funcionalidade de Consulta de Dados da Base de Dados	86
8.2.3. Ferramentas que Permitem a Alteração dos Dados da Base de Dados.....	106
8.2.4. Processamento dos Dados da Base de Dados	116
8.2.5. Ferramentas Administrativas da Base de Dados	119
8.2.6. Menu que Permite a Alteração da Informação Pessoal do Utilizador	140

Capítulo 9 – Testes Realizados e Resultados Obtidos	144
Capítulo 10 – Conclusão	148
10.1. Trabalho Futuro – Perspectivas de Expansão e Integração	148
Bibliografia	152
Bibliografia Utilizada na Concepção da Parte Prática deste Trabalho.....	154
ANEXOS	157

Índice de Figuras

Índice de Figuras

Figura 1 - Concentração de aparelhos de aquisição de imagem (Raio-X, Ecografia, TAC, RMN e PET) por distrito em Portugal Continental. (Adaptado de: [2])	4
Figura 2 - Representação da estrutura química de alguns traçadores em PET (os círculos assinalam o isótopo emissor de positrões que é usado para localizar a molécula). Cada substância deste tipo permite estudar funções particulares do corpo humano, de acordo com o caminho da molécula e as reacções bioquímicas onde ela participa. As moléculas marcadas funcionam como sinais fluorescentes, pois na imagem, as zonas mais activas do organismo vão surgir como pontos luminosos. O radiotraçador mais utilizado é o análogo da glucose (flúordesoxiglicose - FDG), marcado com Flúor 18, um emissor de positrões. A glucose é um dos fornecedores de energia celulares, e assim vai concentrar-se nas zonas do corpo com mais gasto energético, que corresponderão às zonas cancerosas, devido ao facto de estas células serem grandes consumidores de energia. (Adaptado de: [22]).....	26
Figura 3 - Reacção nuclear que está na base do processo de PET/CT. (Adaptado de: [22])	27
Figura 4 - Sequência de eventos num procedimento PET/CT. (a) Representação do sinal em PET (detecção de coincidências verdadeiras): dois fótons de 511 keV são emitidos em sentidos opostos a partir da aniquilação de um positrão e são detectados quase simultaneamente, sem interagir no paciente. (b) Os <i>scans</i> de CT e PET são realizados sequencialmente no tempo. No de PET, os múltiplos detectores (cristais de cintilação) em coincidência definem milhares de linhas de resposta (<i>Lines Of Response – LORs</i>), cada uma contando coincidências verdadeiras. (c) As linhas de resposta são agrupadas em grupos de LORs paralelas correspondendo a múltiplas vistas da distribuição de radioisótopo (projeções). As várias projeções de 0° a 180° podem ser observadas como um sinograma. Os dados de projecção são então reconstruídos para produzir as imagens finais. (d) Corte coronal de uma PET/CT de corpo inteiro, realizado com FDG, com dois tumores visíveis nos pulmões. (Adaptado de: [22]).....	27
Figura 5 - Protocolo padrão de realização de um PET/CT segundo Blodgett et al., 2006. Em primeiro lugar realiza-se uma radiografia digital piloto (“topograma”) de todo o paciente (1), que permite definir a região a examinar, seguindo-se a parte CT do exame (2). De seguida realiza-se a PET (3), e depois as correcções de atenuação e de dispersão (4). O resultado final é a imagem de fusão (5) da parte CT com a parte PET. (Imagem adaptada de: [23])	28
Figura 6 - Componentes de um DBMS. (Adaptado de: [26])	33
Figura 7 - Processo de concepção e implementação de uma BD. (Adaptado de: [26])	36
Figura 8 - O processo de <i>design</i> da interface com o utilizador. (Adaptado de: [28]).....	46
Figura 9 - Fotografia do <i>scanner</i> Gemini GXL da Philips® usado para adquirir imagens de PET/CT no ICNAS.	48
Figura 10 - Imagem para ajudar o leitor a perceber a estrutura genérica do ficheiro de um estudo (volume de imagem reconstruído) no formato <i>imago</i>	49

Figura 11 - Parte inicial do mapa de <i>bytes</i> do cabeçalho principal de um ficheiro <i>imagio</i> . (Adaptado de: [31]).....	51
Figura 12 - Diagrama abreviado do Modelo Conceptual Híbrido concebido para a BD. Mostra apenas os nomes das entidades (no interior das caixas, exceptuando-se a designada ‘EAV Tables’) e as chaves primárias e estrangeiras (fora das caixas, a itálico), bem como uma representação apenas de todas as tabelas EAV que serão necessárias.	61
Figura 13 - Diagrama dos casos de uso definidos para a BD em desenvolvimento. (Diagrama desenvolvido na ferramenta Sybase Power Design 15 [®])	65
Figura 14 - Representação parcial da tabela ‘Estatísticas de Contagem’, visualizada por meio da Interface pgAdmin III [®] , utilizada inicialmente para a interacção com a BD após a sua geração, anteriormente à criação das rotinas em MATLAB [®] para o seu preenchimento e a sua manipulação (no entanto esta imagem foi adquirida já depois do preenchimento da BD). Recorreu-se a esta interface também, sempre que se necessitou de esclarecer qualquer tipo de dúvida relativa aos dados armazenados na BD até ao momento.	70
Figura 15 - Esquema representando as funcionalidades da BD disponíveis para cada perfil.	77
Figura 16 - Diagrama ilustrando as funcionalidades disponibilizadas pela Interface desenvolvida.....	78
Figura 17 - Menu de Boas-Vindas da Interface desenvolvida para se aceder à BD criada.....	80
Figura 18 - Mensagem de erro apresentada ao utilizador da BD quando o nome de utilizador por si introduzido não se encontra presente entre os autorizados a aceder à plataforma.....	80
Figura 19 - Mensagem de erro mostrada quando é introduzida a palavra-chave incorrecta para um dado nome de utilizador constante da BD.....	80
Figura 20 - Menu pessoal de um utilizador da BD com o perfil de Administrador.	81
Figura 21 - Menu de selecção do tipo de actualização que o Administrador pretende realizar. .	82
Figura 22 - Informação ao utilizador da necessidade de seleccionar pelo menos um tipo de actualização.....	82
Figura 23 - Mensagem de erro mostrada quando se cancela a operação.....	82
Figura 24 - Mensagem de erro mostrada quando se pretende proceder à actualização da BD e a conexão ao dispositivo de memória onde os ficheiros se encontram não foi previamente estabelecida.	83
Figura 25 - Mensagem de aviso para que o utilizador trabalhe ao nível da linha de comandos. 83	
Figura 26 - Apresentação parcial da linha de comandos do MATLAB [®] onde é pedido ao utilizador para inserir a sua opção.....	83
Figura 27 - Exemplo da selecção de uma série de directorias para carregamento na BD.....	84
Figura 28 - Exemplo de uma das barras de espera implementadas, para que o utilizador tenha a noção do progresso da execução do carregamento de dados.	85

Figura 29 - Aviso para notificar o utilizador do sucesso da operação de carregamento.	85
Figura 30 - Aviso mostrado para que o utilizador reinicie o MATLAB® no final do carregamento de dados.	85
Figura 31 - Barra de espera para notificar o utilizador da progressão da leitura dos exames presentes na memória física.	86
Figura 32 - Mensagem a informar o utilizador do facto de nenhum ficheiro ter sido apagado do 'BDDISK'.	86
Figura 33 - Mensagem a informar o utilizador do fim da operação de remoção dos ficheiros da BD que já haviam sido removidos da memória física do sistema.	86
Figura 34 - Menu com as funcionalidades de consulta de dados disponibilizadas pela plataforma.	87
Figura 35 - Menu inicial que permite definir o tipo de pesquisa que se pretende realizar. Corresponde à opção de pesquisa de todas as variáveis associadas a uma janela.	88
Figura 36 - Aspecto do menu quando seleccionada a escolha de um conjunto específico de variáveis.	89
Figura 37 - Mensagem apresentada quando se selecciona de uma dada tabela, um item que já havia sido seleccionado e introduzido na lista de selecções.	89
Figura 38 - Apresentação da forma de selecção de variáveis para pesquisar.	90
Figura 39 - Mensagem de erro apresentada quando se tenta fazer uma pesquisa de um conjunto de variáveis, sem ter seleccionado nenhuma variável das existentes na BD para o efeito.	90
Figura 40 - Menu que permite a definição de restrições para a pesquisa de um conjunto de variáveis.	92
Figura 41 - Menu que permite a definição de restrições para a pesquisa de um conjunto de variáveis com ilustração de dois pormenores: modo de selecção de itens da tabela para a lista e variáveis que podem ser escolhidas para ordenar os resultados.	93
Figura 42 - Barra de espera implementada para a monitorização do processo de pesquisa.	94
Figura 43 - Selecção da tabela 'Parâmetros de Configuração do <i>Scanner</i> ' para efectuar a pesquisa de todas as variáveis nela existentes.	94
Figura 44 - Menu que permite escolher entre a implementação ou não de restrições à pesquisa de todas as variáveis existentes numa tabela, e caso se pretenda restringir os valores pesquisados, permite a selecção das variáveis presentes nessa tabela às quais se pretende aplicar as restrições.	95
Figura 45 - Menu de aplicação de restrições, mostrando o pormenor das variáveis disponíveis para organização dos resultados.	96
Figura 46 - Mensagem de notificação do fim da pesquisa, bem como do nome do ficheiro e da directoria onde esta foi armazenada.	96

Figura 47 - Resultado da pesquisa efectuada a todas as variáveis da tabela ‘Parâmetros de Configuração do Scanner’ sem aplicar qualquer tipo de restrição à procura.	97
Figura 48 - Menu de definição das variáveis cujos valores se pretende restringir para a realização da pesquisa, especificando o modo de selecção das variáveis a restringir (preenchimento da lista).....	98
Figura 49 - Tipo de menu desenhado para restringir os valores a serem pesquisados para variáveis do tipo <i>float</i> , <i>int</i> e <i>long</i> . Pode-se ver o pormenor dos sinais lógicos permitidos para o primeiro tipo de restrição a aplicar.....	99
Figura 50 - Mensagem de erro apresentada quando o utilizador insere no segundo painel um valor que não esteja entre o intervalo definido pelos valores mínimo e máximo para a variável em questão, ou então que não seja igual a esses.	99
Figura 51 - Tipo de menu desenhado para restringir os valores a serem pesquisados para variáveis do tipo <i>float</i> , <i>int</i> e <i>long</i> , mostrando a activação do terceiro painel mediante a escolha do utilizador de ‘procura entre dois valores’.....	99
Figura 52 - Mensagem de erro apresentada quando o utilizador insere no terceiro painel um valor que não esteja entre o intervalo definido pelos valores mínimo e máximo para a variável em questão, ou então que não seja igual a esses, e caso o valor definido como mínimo não seja menor que o máximo.....	100
Figura 53 - Menu de implementação de restrições para variáveis do tipo <i>short</i> (que podem referir-se a valores lógicos).....	100
Figura 54 - Tipo de mensagem de erro apresentada caso o utilizador interrompa a operação de definição das restrições a aplicar à pesquisa a meio. No exemplo, seria a mensagem mostrada caso o utilizador premisse o botão <i>Cancel</i> no menu de implementação de restrições da variável ‘ <i>crbTstampPeriod</i> ’.....	101
Figura 55 - Menu para restrição dos valores de variáveis do tipo <i>char</i>	102
Figura 56 - Menu para restrição dos valores de variáveis do tipo <i>char</i> , mostrando a ferramenta de pesquisa de uma dada expressão dentro dos valores da variável.....	102
Figura 57 - Menu apresentando o tipo de visualizadores de que a BD disponibiliza até ao momento.....	103
Figura 58 - Janela com a listagem de todos os exames presentes na BD (as dimensões do menu foram alteradas).....	103
Figura 59 - Janela com a listagem das directorias de pastas de armazenamento de ficheiros de imagem presentes na BD.....	104
Figura 60 - Mensagem de erro notificando o utilizador da necessidade de estabelecer o acesso ao ‘ <i>BDDISK</i> ’ antes de tentar visualizar os ficheiros.	104
Figura 61 - Mensagem de erro apresentada caso o ficheiro de TAC associado ao ‘ <i>.img</i> ’ seleccionado não seja encontrado nos dados imagiológicos constantes das entradas da BD....	104

Figura 62 - Interface de visualização dos ficheiros <i>imago</i> implementada em IDL®	105
Figura 63 - Mensagem de erro mostrada ao utilizador quando a operação de selecção de um volume de dados é cancelada (é seleccionado o botão <i>Back</i> no menu apresentado na Fig. 59).	105
Figura 64 - Menu inicial do visualizador DICOM em <i>background</i> e menu de escolha de tarefas disponibilizado após se pressionar o botão direito do rato (a imagem não se encontra na escala normal).	106
Figura 65 - Mensagem de erro que informa o utilizador do facto de não existirem ficheiros no formato DICOM na directoria por si seleccionada.	106
Figura 66 - Menu de selecção do tipo de informação que se pretende alterar (informação relativa aos pacientes, exames, processamento ou variáveis).	107
Figura 67 - Menu apresentado para que o utilizador possa alterar os campos da BD que contêm informação relativa aos pacientes.	107
Figura 68 - Menu que permite a selecção do paciente ao qual se pretende alterar os dados.....	108
Figura 69 - Pormenor da activação do botão <i>OK</i> apenas após a selecção de um dado paciente da lista.	108
Figura 70 - <i>Warning</i> mostrado ao utilizador quando ele tenta alterar os dados de um dado paciente, sem o ter escolhido previamente.....	108
Figura 71 - Inserção do identificador e nome do paciente escolhido no menu principal de alteração de dados de pacientes.....	109
Figura 72 - Menu que apresenta o nome e o valor de todas as variáveis associadas a um dado paciente (inclui os atributos que foram adicionados em tabelas EAV).....	110
Figura 73 - Edição do campo que havia sido seleccionado para modificação no menu apresentado na Fig. 72.....	110
Figura 74 - À direita vemos as mensagens que vão surgindo ao longo da execução do programa de verificação dos valores das variáveis alteradas, no caso de os valores inseridos não se encontrarem concordantes com as restrições definidas para elas. À esquerda podemos ver a mensagem apresentada no final da execução referida, no caso de terem havido variáveis mal alteradas. Ele apresenta uma lista com os identificadores dos pacientes onde esse erro ocorreu (note-se que na selecção de alteração dos dados de um dado paciente, este será sempre o mesmo, mas no caso da alteração de todos os pacientes ou de um determinado grupo destes, isto já não acontece) e os nomes das variáveis que foram mal alteradas.	110
Figura 75 - Selecção da opção de alteração dos dados referentes a todos os pacientes da BD.	111
Figura 76 - Menu de alteração dos dados referentes a todos os pacientes constantes da BD. ..	111
Figura 77 - Disponibilização do painel de selecção de um grupo de pacientes, quando é esta a opção escolhida.	112
Figura 78 - Menu que permite a alteração apenas de um conjunto de pacientes.	113

Figura 79 - Mensagem de sucesso apresentada ao utilizador quando todas as alterações são aplicadas sem qualquer erro. São apresentadas mensagens semelhantes para os outros casos de alteração que serão analisados: exames, processamentos e variáveis.	113
Figura 80 - Menu de selecção do processamento a alterar.	114
Figura 81 - Pormenores da colocação do identificador e descrição do processamento escolhido na caixa de texto, e do facto do botão <i>OK</i> se tornar activo.	114
Figura 82 - Menu de alteração das variáveis associadas ao processamento seleccionado, por meio de edição das células da tabela onde são apresentadas.	115
Figura 83 - Menu que permite a alteração da descrição das variáveis presentes na BD.	115
Figura 84 - Menu de processamento de dados.	116
Figura 85 - Pode-se ver que, com apenas a selecção dos exames a processar, as funcionalidades de guardar e visualizar os resultados, bem como o botão que permite o processamento, permanecem inactivos.	117
Figura 86 - Com selecção apenas das rotinas, e não dos exames, acontece o mesmo que descrito para a figura anterior.	117
Figura 87 - Finalmente, com ambas as listas preenchidas, as funcionalidades referidas nas duas figuras anteriores ficam disponíveis.	118
Figura 88 - Barra de espera implementada para a monitorização da evolução do processamento de dados.	118
Figura 89 - Aviso mostrado ao utilizador a informá-lo da necessidade de ter de escolher pelo menos uma das opções disponíveis.	118
Figura 90 - Mensagem que notifica o utilizador do fim e do sucesso do processamento.	118
Figura 91 - Ferramentas de gestão e administração da BD.	119
Figura 92 - Menu que permite a gestão dos utilizadores existentes na BD.	120
Figura 93 - Mensagem de erro apresentada ao Administrador quando este tenta inserir um nome de utilizador que já consta dos registos da BD.	120
Figura 94 - Mensagem de erro que é apresentada quando o Administrador tenta inserir um caracter 'espaço' num novo nome de utilizador.	120
Figura 95 - Mensagem de erro que é apresentada quando o Administrador tenta inserir um caracter 'espaço' numa nova palavra-chave.	120
Figura 96 - Mensagem de erro mostrada ao Administrador quando este pressiona o botão <i>Add User</i> sem ter preenchido um dos campos (ou ambos) referentes ao nome de utilizador e à palavra-chave.	121
Figura 97 - Exemplificação da adição de um novo utilizador.	121
Figura 98 - Mensagem que informa que o novo utilizador foi adicionado com sucesso à BD.	121

Figura 99 - Menu após a adição do novo utilizador à BD. Vê-se o pormenor da actualização automática dos elementos presentes na tabela onde se encontram todos os utilizadores existentes na BD.	122
Figura 100 - Menu de adição da informação pessoal do utilizador. (a) - À esquerda podemos ver o menu por preencher. (b) – À direita vemos um exemplo de preenchimento do menu. Note-se que o botão <i>Pick Photo</i> possibilita a escolha da imagem que o utilizador irá ter como foto de perfil (através de uma caixa de diálogo mostrando apenas as imagens constantes de uma dada directoria)	122
Figura 101 - Mensagem apresentada ao Administrador dando-lhe conta do sucesso da introdução dos dados pessoais do novo utilizador na BD.	123
Figura 102 - Aspecto do menu após se seleccionar a opção <i>Add Another User</i>	123
Figura 103 - Selecção dos utilizadores a serem removidos.....	124
Figura 104 - Mensagem apresentada ao Administrador a dar conta do sucesso da operação de remoção.	124
Figura 105 - No final do processo, a lista de utilizadores existentes na BD é automaticamente actualizada, tal como acontecia após a adição de um novo utilizador.	124
Figura 106 - Menu que permite a gestão dos pacientes da BD.	126
Figura 107 - Pormenor do botão de remoção só permanecer activo enquanto a lista contém itens.	126
Figura 108 - Aviso que notifica o utilizador do facto da remoção de pacientes implicar também a remoção dos exames a eles associados, e dos processamentos efectuados a estes últimos. O utilizador deve escolher se pretende ou não prosseguir com a operação.	126
Figura 109 - Barra de espera implementada para monitorizar o processo de procura de pacientes redundantes nas entradas da BD.....	127
Figura 110 - Menu que mostra um dado conjunto de dados suspeitos de pertencerem a um mesmo paciente.....	127
Figura 111 - Menu mostrando o pormenor do botão <i>Confirmation</i> só ficar activo aquando do preenchimento da lista de pacientes redundantes.....	127
Figura 112 - Menu que permite a remoção de exames da BD.	128
Figura 113 - Ilustração da forma de selecção dos exames a remover (note-se que o botão para proceder à remoção só fica activo quando a lista de itens a remover se encontra preenchida).	128
Figura 114 - Menu que informa o utilizador do facto de também serem removidos todos os processamentos associados ao exames seleccionados, e que lhe permite escolher continuar ou não com a operação de remoção.	129
Figura 115 - Mensagem apresentada quando se cancela uma operação de remoção.	129
Figura 116 - Menu de gestão dos processamentos efectuados aos dados.	130

Figura 117 - Janela que permite a remoção de todos os processamentos realizados a um dado conjunto de exames.	130
Figura 118 - Menu que permite a adição de novas rotinas e a remoção destas da BD.	131
Figura 119 - Mensagem de erro mostrada ao utilizador quando ele tenta inserir uma rotina cujo nome já existe nos registos da BD.	131
Figura 120 - Mensagem de erro mostrada ao utilizador quando ele tenta inserir uma rotina na BD, sem ter colocado previamente o ficheiro '.m' que lhe corresponde na directoria da interface.	132
Figura 121 - Disponibilização do campo de descrição da rotina, e do campo para inserir o número de variáveis de entrada que esta utiliza.	133
Figura 122 - Disponibilização do botão <i>Add Input Variables</i> , após a inserção de um número válido de variáveis de entrada.	133
Figura 123 - Mensagem de erro mostrada quando o Administrador tenta inserir um número menor ou igual a zero de variáveis.	134
Figura 124 - Menu de descrição das variáveis de entrada.	134
Figura 125 - Descrição da primeira variável de entrada como sendo uma das variáveis integrantes da BD.	134
Figura 126 - Aspecto do menu quando se pretende descrever uma variável de entrada que não faz parte da BD. Mostra-se o pormenor do tipo de dados que se pode atribuir a essa variável.	135
Figura 127 - Mensagem de erro mostrada ao Administrador no caso de ele interromper a operação de descrição de variáveis de entrada.	135
Figura 128 - Disponibilização do campo para inserir o número de variáveis de saída.	136
Figura 129 - Mensagem de erro mostrada ao utilizador, caso ele não insira um número válido de variáveis de saída.	136
Figura 130 - Activação do botão de descrição das variáveis de saída.	137
Figura 131 - Menu que permite a escolha da entidade à qual se associarão os resultados da rotina. Permite a associar esses dados às entidades: 'Processamento', 'Paciente', 'Estudo' e 'Corte'.	137
Figura 132 - Mensagem de erro mostrada ao utilizador quando a operação de escolha de uma entidade para guardar os dados que resultam da rotina é cancelada.	137
Figura 133 - Menu de descrição das variáveis de saída.	138
Figura 134 - Activação do campo onde o <i>callback</i> da função foi gerado, para que possa ser editado.	138
Figura 135 - Mensagem que informa o utilizador do sucesso da inserção da rotina na BD.	139

Figura 136 - Mensagem de erro mostrada quando se detecta algum problema ao nível da sintaxe e do número de variáveis presentes no <i>callback</i>	139
Figura 137 - Activação do botão que permite adicionar uma nova rotina (<i>Add Another Routine</i>).	139
Figura 138 - Menu onde o utilizador comum pode alterar (ou adicionar caso ainda não o tenha feito) os seus dados pessoais que se encontrarão registados na BD.	140
Figura 139 - Mensagem de erro mostrada quando se tenta introduzir uma data de nascimento descabida para um dos utilizadores da BD (note-se o ano de criação da BD: 2011).	141
Figura 140 - Janela que permite a alteração da <i>password</i>	141
Figura 141 - Mensagem de erro apresentada ao utilizador quando a confirmação da nova palavra-chave não foi bem sucedida.	141
Figura 142 - Janela que permite a selecção da foto de perfil do utilizador.	142
Figura 143 - Mensagem que é apresentada ao utilizador após se efectuarem as alterações por si definidas aos seus dados pessoais.	142
Figura 144 - Possível modelo conceptual relacional para a implementação de uma BD não só de imagens de PET/CT mas também de dados de RMN. Diagrama desenvolvido com auxílio da ferramenta Sybase Power Designer 15 [®]	149

Índice de Tabelas

Índice de Tabelas

Tabela 1 - Contagem do número de aparelhos de imagiologia médica existentes em cada região de Portugal Continental onde foi considerado haver um mercado relevante (ou seja, um número significativo de utentes) para a aquisição de imagens médicas. Os números são seccionados de acordo com o facto das tecnologias enumeradas existirem em Instituições de carácter privado, social ou público. Apresenta-se também a representação percentual dos estabelecimentos por região relativamente ao número de estabelecimentos existentes a nível nacional, e o número de estabelecimentos por 100 000 habitantes em cada região. (Adaptado de: [2])	3
Tabela 2 - Número de aparelhos de PET/CT existentes em alguns países Europeus. (Adaptado de: [3], com fonte em Campos JR.: A Imagiologia no Serviço Nacional de Saúde. Direcção-Geral de Saúde. Lisboa Outubro 2007; pág. 64)	5
Tabela 3 - Casos de maior utilização da tecnologia PET/CT. (Adaptado de: [3])	5
Tabela 4 - Estimativa do número de utentes por ano que procuram os serviços hospitalares devido aos factores discutidos na Tabela 3, divididos mediante a região de realização. (Adaptado de: [3])	6
Tabela 5 - Estimativa do número de exames de PET/CT, RMN e PEM realizados no ICNAS até ao momento.	7
Tabela 6 - Breve descrição da estrutura e conteúdo do documento.	12
Tabela 7 - Informações sobre os membros do projecto.....	15
Tabela 8 - Recursos tecnológicos de que o ICNAS dispõe (à esquerda) e exames médicos nele realizados (à direita). (Adaptado de: [14], [16]).....	16
Tabela 9 - Dimensões e tamanhos das imagens biomédicas. (Adaptado de: [18])	22
Tabela 10 - Lista das versões de cabeçalhos principais. (Adaptado de: [31]).....	50
Tabela 11 - Lista das versões dos sub-cabeçalhos. (Adaptado de: [31]).....	51
Tabela 12 - Tipo de ficheiros scan que podem ser gerados durante uma aquisição. (Adaptado de: [32]).....	52
Tabela 13 - Tipos de Ficheiros TAC que podem ser gerados numa aquisição. (Adaptado de: [32]).....	54
Tabela 14 - Descrição das restantes entidades (e atributos que as constituem) que fazem parte da BD.	62

Acrónimos

Acrónimos

BD – Base de Dados

CT – *Computed Tomography*
(Tomografia Computorizada)

DBMS – *Database Management System* (Sistema de Gerenciamento da Base de Dados)

DDL – *Data-Definition Language*
(Linguagem de Definição de Dados)

DICOM – *Digital Imaging and Communications in Medicine* (Imagiologia Digital e Comunicações em Medicina)

DML – *Data-Manipulation Language*
(Linguagem de Manipulação de Dados)

EAV – Entidade-Atributo-Valor

ER – Entidade-Relação ou Entidade-Relacionamento

FCTUC – Faculdade de Ciências e Tecnologia da Universidade de Coimbra

FMUC – Faculdade de Medicina da Universidade de Coimbra

FOV – *Field Of View* (Campo de Visão)

HPP – Hospitais Privados de Portugal

IBILI – Instituto Biomédico de Investigação da Luz e Imagem

ICNAS – Instituto das Ciências Nucleares Aplicadas à Saúde

IPO – Instituto Português de Oncologia

MIDS – *Medical Image Database Systems* (Sistemas de Bases de Dados de Imagem Médica)

MRI – *Magnetic Resonance Imaging* (Imagiologia por Ressonância Magnética)

PEM – *Positron Emission Mammography* (Mamografia por Emissão de Positrões)

PET – *Positron Emission Tomography* (Tomografia por Emissão de Positrões)

RMN – Ressonância Magnética Nuclear

SA – Sociedade Anónima

TAC – Tomografia Axial Computorizada

UC – Universidade de Coimbra

UE – União Europeia

Capítulo 1

Introdução

1. Introdução

Hoje em dia, armazenar e organizar a informação de um modo eficiente é essencial para se obter dela o *feedback* desejado, ou seja, a maneira como se estruturam os dados é crucial para extrair informação útil a partir deles. As Bases de Dados (BDs) são uma forma comum e fácil de o fazer, tornando possível a manipulação, análise, correlação e visualização da informação guardada.

O objectivo deste projecto é a criação de uma Base de Dados (BD) para armazenamento de todo o tipo de informação que pode ser extraído de um ficheiro de Imagem Médica. Particularmente, neste trabalho foram utilizadas as imagens de PET/CT recolhidas no ICNAS até ao momento pelo tomógrafo Gemini GXL da Philips®.

A ideia geral foi parametrizar a informação codificada em *bytes* nos ficheiros que guardam o exame médico, recorrendo a rotinas de extracção já pré-desenvolvidas e fornecidas pelos criadores da tecnologia. Elas permitem a recolha tanto da meta-informação armazenada nos cabeçalhos de cada ficheiro, como da imagem adquirida propriamente dita.

Após a recolha da informação, esta foi estudada de forma a averiguar a melhor maneira de organizá-la convenientemente para a sua integração numa BD. Teve-se também atenção a que depois fosse possível visualizar os dados, de modo a relacioná-los e extrair conclusões, recorrendo a uma interface gráfica que permitisse tal processamento.

1.1. Contexto/Enquadramento

Nas décadas de 1980 e 1990 houve um grande avanço na imagiologia biomédica, com o aparecimento de várias novas modalidades de aquisição de imagem médica, e desenvolvimento das já existentes. Este proporcionou uma autêntica revolução no modo de aquisição de imagens, meios flexíveis e de fácil manuseamento para ver secções anatómicas e estados fisiológicos e, em alguns casos, a redução da dose de radiação a ser injectada nos pacientes e do trauma causado pelo procedimento do exame. Destacam-se como principais acontecimentos neste período o aparecimento de novos *scanners* e aparelhos de ultra-sons, o desenvolvimento da Ressonância Magnética Nuclear (RMN), da Tomografia Computacional (*Computed Tomography* - CT), da Tomografia por Emissão de Positrões (*Positron Emission Tomography* - PET), da Tomografia Computorizada por Emissão de um Positrão Único (*Single Positron Emission Computed Tomography* – SPECT), a Angiografia de Subtracção Digital (*Digital*

Subtraction Angiography – DSA) e a Imagiologia de Fonte Magnética (Magnetic Source Imaging – MSI) [1].

Estes tipos de exames mudaram drasticamente as práticas médicas e são eles que geram o maior volume de dados imagiológicos ao nível hospitalar. A sua dinamização tem levantado questões relacionadas com a integração dos vários sistemas de aquisição de imagem, a necessidade de armazenamento dos dados recolhidos de modo a que todos os departamentos possam depois aceder-lhes, e que possam também quiçá ser transmitidos para outras Instituições de Cuidados de Saúde para possibilitar a análise de outros profissionais de saúde. Há então uma preocupação crescente (dos administradores hospitalares e dos desenvolvedores dos sistemas de informação) em integrar desde o início os vários sistemas de informação médica para melhorar a qualidade dos cuidados de saúde prestados, e de forma a reduzir os custos que advêm de ter este tipo de uniformização em estágios tardios.

As referidas modalidades constituem cerca de 30% dos exames radiológicos nos Estados Unidos da América (EUA) e os outros 70% dos exames são realizados recorrendo à radiologia tradicional (raios-X) e à radiografia digital luminescente.

Em Portugal este sector ainda não se encontra tão desenvolvido. Segundo o *Estudo sobre a Concorrência no Sector da Imagiologia*, em 2009, a distribuição de instituições com aparelhos de imagiologia em Portugal continental era a descrita na tabela 1, e representada na Fig. 1. As técnicas consideradas neste estudo foram: o raio-X, a ecografia, a TAC, a RMN e a PET [2].

Tabela 1 - Contagem do número de aparelhos de imagiologia médica existentes em cada região de Portugal Continental onde foi considerado haver um mercado relevante (ou seja, um número significativo de utentes) para a aquisição de imagens médicas. Os números são seccionados de acordo com o facto das tecnologias enumeradas existirem em Instituições de carácter privado, social ou pública. Apresenta-se também a representação percentual dos estabelecimentos por região relativamente ao número de estabelecimentos existentes a nível nacional, e o número de estabelecimentos por 100 000 habitantes em cada região. (Adaptado de: [2])

Zonas com um Mercado Relevante para a Imagiologia Médica	Privada	Social	Pública	N.º Total de Estabelecimentos	% do Total de Estabelecimentos a nível Nacional	Estabelecimentos/ 100 000 Habitantes
Bragança	1	0	1	2	0 %	3,49
Chaves	3	0	1	4	1 %	4,92
Viana do Castelo	8	2	2	12	2 %	4,76
Braga	11	3	2	16	3 %	3,93
Mirandela	2	0	2	4	1 %	4,20
Guimarães	23	2	2	27	5 %	5,17
Felgueiras	14	3	1	18	3 %	3,22
Vila Real	6	0	3	9	2 %	4,48
Porto	72	5	7	84	15 %	6,58
São João da Madeira	13	0	2	15	3 %	5,25
Viseu	6	0	2	8	1 %	2,75
Aveiro	12	2	4	18	3 %	4,54
Guarda	2	0	1	3	1 %	2,68
Seia	1	0	1	2	0 %	4,12
Covilhã	0	1	1	2	0 %	2,17
Coimbra	19	3	4	26	5 %	7,75
Lousã	1	2	0	3	1 %	2,18
Castelo Branco	4	0	1	5	1 %	6,64

Focando-nos na modalidade de PET/CT, a utilizada neste trabalho, em 2007 existiam quatro equipamentos de PET (distribuídos entre PET unicamente e PET/CT) em território nacional (ver tabela 2) [3].

Tabela 2 - Número de aparelhos de PET/CT existentes em alguns países Europeus. (Adaptado de: [3], com fonte em Campos JR.: A Imagiologia no Serviço Nacional de Saúde. Direcção-Geral de Saúde. Lisboa Outubro 2007; pág. 64)

	França	Alemanha	Itália	Reino Unido/Irlanda	Espanha	Portugal
N.º de Aparelhos PET/CT	42	-	59	22	42	4
Unidade/Milhão de Habitantes	0.7	0.0	1.0	0.3	0.9	0.4

O primeiro estudo PET foi realizado no Porto, em Outubro de 2002, nas instalações dos Hospitais Privados de Portugal (HPP-Medicina Molecular, SA). Em 2008, estavam instaladas 7 unidades PET: um tomógrafo PET no Instituto Português de Oncologia de Lisboa Francisco Gentil (IPOLFG), e seis sistemas PET/CT, distribuídos pelas seguintes instituições: IPO do Porto Francisco Gentil (IPOPFG), HPP-Medicina Molecular (SA), Hospitais da Universidade de Coimbra, Quadrantes (Clínica Médica e de Diagnóstico – Grupo Joaquim Chaves), Hospital CUF Descobertas e Hospital da Luz em Lisboa [4]. No corrente ano, 2012, tem-se conhecimento da existência de pelo menos mais dois equipamentos deste tipo: um no IPO de Braga e outro na Fundação Champalimaud.

Giraldes, 2010, avaliou as necessidades nacionais de equipamentos de PET/CT mediante as estimativas populacionais de 2007, e seleccionou os seguintes casos como os mais prováveis utentes para esta tecnologia: a prevenção da doença cardíaca nos Diabéticos Tipo II (nos adultos de 50 ou mais anos), o *Screening* coloproctológico (não usado frequentemente em Portugal, mas que se estima que deva ser realizado a pelo menos 10% dos adultos a partir dos 50 anos), o tratamento de cancro (deverá realizar-se, apenas, nos seguintes tipos de cancro: cancro da mama, cancro coloproctológico, cancro do esófago, cancro da cabeça e pescoço, cancro do pulmão, linfoma, melanoma, nódulos pulmonares solitários, cancro da tiroide e cancro cervical; estes constituem cerca de 46,7 % da totalidade da mortalidade por cancro), o acompanhamento de uma cirurgia a um cancro do pulmão ou ao coração, em cardiologia (por exemplo, na Doença Arterial Coronária), e rastreamento da causa de uma dor no peito numa Urgência Hospitalar. Uma síntese detalhada destes casos, bem como a quantidade de utentes nestas circunstâncias encontram-se nas tabelas 3 e 4, respectivamente.

Tabela 3 - Casos de maior utilização da tecnologia PET/CT. (Adaptado de: [3])

Utilizações da PET/CT			
Diabetes	Cancro	Pulmão	Cardiologia
<i>Prevenção</i>			
Ambulatório	Ambulatório	Ambulatório	Urgência Hospitalar
Despistagem da doença	<i>Screening</i> coloproctológico em	Não se deve aplicar no	Dor no peito com ECG

cardíaca.	adultos com 50 ou mais anos.	<i>screening</i> do cancro do pulmão.	normal. Pode proceder-se à detecção de uma doença cardíaca.
<i>Tratamento</i>			
	Ambulatório		Ambulatório
	Tratamento do cancro: da mama, coloretal, esófago, cabeça e pescoço, pulmão, linfoma, melanoma, nódulos pulmonares solitários, tireoide e cervical.		Doença Arterial Coronária.
	Internamento		Internamento
	Fase pré-operatória de oncologia pulmonar.		Doentes cardíacos na fase pré-operatória.

Tabela 4 - Estimativa do número de utentes por ano que procuram os serviços hospitalares devido aos factores discutidos na Tabela 3, divididos mediante a região de realização. (Adaptado de: [3])

Estimativa Exames/Ano	Prevenção		Tratamento			Urgência	
	Diabetes	Screening Coloretal	Cardiologia	Intervenções Cirúrgicas Córdio-Torácicas	Cirurgia Cancro do Pulmão	Morbilidade por Cancro	Dor no Peito
Norte	5206	18077	1277	9094	981	9052	13352
Centro	3773	13104	379	3253	564	6987	4051
Lisboa	5355	18594	1175	22163	1130	14499	12789
Alentejo	686	2383	166	-	202	1932	1808
Algarve	675	2345	156	-	150	1544	1696
Total	15695	54503	3153	34510	3027	34014	33696

Os resultados apurados pelo 4.º Inquérito Nacional de Saúde (INS¹) realizado entre Fevereiro de 2005 e o mesmo mês de 2006 pelo Instituto Nacional de Estatística (INE) em parceria com o Instituto Nacional de Saúde Dr. Ricardo Jorge (INSA) mostram que de entre as 10 549 424 pessoas que constituíam a população residente, 686 315 sofriam de Diabetes, 2 085 013 de tensão arterial alta, 197 665 de um tumor/cancro maligno, 171 638 já sofreram um Acidente Vascular Cerebral (AVC), 398 918 sofrem de obesidade, 131 544 já tiveram um enfarte do miocárdio, 2 566 677 de uma outra qualquer doença crónica [5]. Ignorando o montante respeitante a outras doenças crónicas (que podem ou não estar relacionadas com os factores discutidos anteriormente para a realização de uma PET/CT), podemos aferir pelos outros valores apresentados, que respeitam a factores que podem levar à necessidade de realização de uma PET/CT mediante os grupos identificados anteriormente, que aproximadamente 35% da população nacional pode eventualmente necessitar de realizar exames de PET/CT de forma continuada ou esporádica.

Vemos então o quão urgente é a implementação de métodos que permitam armazenar e aceder à informação constante destes exames, de forma continuada, para se poder seguir o histórico dos exames realizados por determinado paciente, por exemplo.

¹ INS - “É um instrumento de medida e de observação em saúde, que recolhe dados de base populacional,

A situação ao nível da Instituição onde este projecto foi desenvolvido é traduzida pelas estimativas apresentadas na tabela 5, mostrando-se assim a importância da instauração do tipo de ferramentas referidas anteriormente no seu seio.

Tabela 5 - Estimativa do número de exames de PET/CT, RMN e PEM realizados no ICNAS até ao momento.

Ano	PET/CT	RMN*	PEM**
2009	171	1300	-
2010	775	1300	-
2011	1252	1300	-
2012	235	1300	7

*Apenas se conseguiu aceder a uma estimativa do número de exames realizados por ano, nos quatro anos de funcionamento da instituição (e não ao valor exacto de exames realizados em cada ano).

**Ainda em fase experimental. Realizaram-se a aquisição de 14 mamas (7 mamas esquerdas e 7 mamas direitas) e 14 axilas (correspondendo também metade ao lado esquerdo e a outra metade ao lado direito).

1.2. Objetivo

O objectivo deste projecto consiste na criação de uma BD de imagem médica, não apenas com as imagens em si, mas também com a meta-informação guardada nos cabeçalhos dos ficheiros dos exames, e nova informação obtida a partir da aplicação de vários métodos de processamento de imagem, análise de dados e algoritmos de classificação aos dados disponíveis.

O desafio proposto foi a recolha da informação dos exames da PET/CT adquiridos no centro de imagiologia ICNAS da UC até ao momento presente, e organizá-la da melhor maneira possível para a construção de uma plataforma computacional que permitisse a aplicação de métodos de análise estatística, processamento de imagem (segmentação, visualização e análise de exames, rastreio de tumores, etc.), mineração de dados (*data mining*), extracção de correlações, e outros tipos de manipulação de dados que possam ser implementados. A BD deve permitir adição de nova informação calculada a partir destes algoritmos de uma forma fácil e intuitiva, enquanto preserva a privacidade dos pacientes e restringe o acesso aos dados somente a pessoal autorizado.

O ICNAS, para além da vertente clínica de realização de exames, tem a investigação como uma importante componente da sua área de actuação, distinguindo-se assim da maioria dos centros que utilizam a técnica PET/CT. Assim, ao nível da Instituição, pretende-se que esta plataforma permita a monitorização de casos e rotinas do centro (para despiste de eventuais problemas e optimização de protocolos), o desenvolvimento da pesquisa (análise de grupos,

classificação...) e a educação/formação para a percepção da imagem médica pelos profissionais que com ela colaboram.

1.3. Estado da Arte

As Bases de Dados de imagem médica criadas até ao momento geralmente só armazenam a imagem em si e não exploram as potencialidades da meta-informação a ela associada. Genericamente, a forma que o utilizador tem de lhes aceder é através de visualizadores de imagem muito simples, que permitem somente a realização de tarefas básicas de processamento, tais como: a visualização (de uma ou várias imagens ao mesmo tempo), geralmente com métodos de amplificação e redução da imagem, possibilidade de escolha da paleta de cores, extracção de coordenadas, delimitação de zonas de interesse, redimensionamento, *zoom in/out*, mudanças de luminosidade/contraste e algumas outras tarefas simples.

Foi realizada uma pesquisa de Bases de Dados existentes *online*, e concluiu-se que geralmente são apenas fornecidas as imagens para *download* (ou somente um *link* que permite a sua visualização), ou então é possível visualizá-las com possibilidade de alteração apenas de algumas das características já mencionadas. Relativamente à meta-informação, aquela que se pretende explorar neste projecto, é fornecida muito pouca, e geralmente apresentada em pequenas tabelas em 'Excel' ou CSV. Não há, de uma maneira geral, uma estruturação da informação textual relativa ao estudo que deu origem a essa imagem. De seguida apresenta-se o *site* encontrado com informação mais relevante, uma vez que aglomera uma série de Bases de Dados públicas de imagem médica existentes na *web* para fins educacionais, de desenvolvimento de *software* associado ao processamento de imagem e de realização de pesquisas [6].

Este *website* foi desenvolvido pelo Grupo de Visão Computacional e Análise de Imagem da Universidade de Cornell. Começa por apresentar duas Bases de Dados de raios-X digitais realizados ao tórax: a primeira, disponibilizada pela *Japanese Society of Radiological Technology*, apresenta ficheiros de imagem com e sem nódulos pulmonares, no formato '.img', disponíveis para *download*; já a segunda, para além de permitir o descarregamento das imagens, tem um visualizador dos resultados da segmentação pulmonar aplicada a todas as radiografias presentes na BD, variando características associadas à distância de contorno mínima, ao tamanho e tipo de compressão aplicada, a ajustes de luminosidade e contrastes usados no algoritmo, a contornos realizados por diferentes observadores humanos especializados, à

modelação de forma e aparência e à sobreposição de dois objectos. Esta última é disponibilizada pelo *Image Sciences Institute*.

De seguida são apresentadas cinco BDs com imagens de CT. A primeira pertence ao ‘Relatório de Tamanho de Nódulos’ do *Lung Image Database Consortium* (LIDC) e nela constam imagens de nódulos pulmonares, com o acréscimo de informações relativas ao seu volume, diâmetro, posição (coordenadas horizontais e verticais) e cortes (ou secções, do inglês *slices*) onde se encontram, num ficheiro à parte em formato Excel ou CSV. A segunda é a *Reference Image Database to Evaluate Response* (RIDER) que está a ser implementada com vista à terapia do cancro do pulmão. As imagens referidas nas duas BDs anteriores, bem como outras imagens, estão também disponíveis na terceira BD apresentada, que pertence ao *National Cancer Institute* onde podemos encontrar o *The Cancer Imaging Archiving* (TCIA), uma BD de imagens no formato DICOM que estão disponíveis para pré-visualização e *download* após registo no *site*. Note-se que as pesquisas nesta plataforma são redireccionadas para o *National Biomedical Imaging Archive*, que disponibiliza um motor de busca de imagens organizado segundo a doença em causa, a região adquirida e a modalidade de imagem utilizada (CT, MRI, etc.). Apurou-se posteriormente que esta BD também inclui imagens de PET. A quarta BD apresentada contém também imagens de pulmão em CT, e designa-se *Public Lung Image Database*; é gerida pela *Vision and Image Analysis Group* (VIA) e pela *International Early Lung Cancer Action Program* (I-ELCAP), e tem um total de 50 imagens de pulmão inteiro adquiridas com baixa dosagem, identificadas pela data de aquisição mais um identificador numérico, com referência ao seu tamanho, tipo (geralmente, DICOM(CT)) e resolução; disponibiliza um visualizador Java para as imagens, que permite fazer *zoom in/out*, escolher a paleta de cores, navegar pela imagem e mudar o janelamento que lhe é aplicado; é também apresentada uma lista com os dados DICOM da imagem. A última BD apresentada nesta subsecção pertence também aos mesmo gestores da anterior e designa-se *Public Lung Database to Assess Drug Response*; disponibiliza também ferramentas parecidas com a anterior para visualização e manipulação da imagem, com a adição da análise tridimensional da imagem e a possibilidade do utilizador realizar anotações nas imagens, no entanto não disponibiliza os dados DICOM.

No referido *site*, segue-se uma BD com dados simulados de MRI, chamada *Simulated Brain Database* (SBD). Pretende-se ter um repositório de dados 3D de MRI, com simulação de situações normais e de esclerose múltipla, aplicando diferentes parâmetros de aquisição. Permite a visualização e também o *download* das imagens, mostrando detalhes relativos à dimensão e posição, e permitindo somente a ampliação das imagens e três opções de coloração.

Depois são listadas duas Bases de Dados respeitantes a mamografias, designadamente: *Digital Database for Screening Mammography* (DDSM) e *Mammographic Image Analysis Society* (MIAS). A primeira tem 2620 caso disponíveis em 43 volumes de dados. Os ficheiros

de imagem para *download* encontram-se no formato LJPEG (*Lossless JPEG*), no entanto é possível visualizar as imagens no *site*, bem como uma lista de informações a elas associadas (nome do ficheiro, data do estudo, idade do paciente, tipo de filme, densidade, data da digitalização) e anotações feitas por quem as analisou. A segunda BD disponibiliza ficheiros nos formatos PBM (*portable bitmap* – binário), PGM (*portable greymap* – escala de cinzentos), PPM (*portable pixmap* – cor) e PNM (*portable anymap* – corresponde a PBM, PGM ou PPM), e fornece informação quanto ao tecido de *background* (gordura, gordura-glandular, denso-glandular), ao tipo de anomalia presente (calcificação, massa, assimetria, normal, etc.), à gravidade da anomalia (benigna ou maligna) e às coordenadas e raio da anomalia.

Refere também duas Bases de Dados de casos de retinopatia diabética. A primeira, DIARETDB1 (*Standard Diabetic Retinopathy Database*), apresenta inúmeras imagens da retina no formato ‘.png’, e disponibiliza uma ferramenta em MATLAB[®] que pretende classificar a retinopatia com base em análise de densidades por processamento de imagem. A segunda, DRIVE (*Digital Retinal Images for Vessel Extraction*), pertence ao *Image Sciences Institute*, e fornece uma enorme variedade de imagens dos vasos da retina para descarregamento, no formato ‘.gif’.

Ao nível da já referida NBIA, é apresentada uma BD de imagens de Colonoscopia Virtual, e ao nível do TCIA apresentam-se imagens de Colonografia por CT. Estas últimas têm associada meta-informação em ficheiros Excel, onde são descritas as características dos pólipos encontrados em cada imagem. Ambas se encontram no formato DICOM, com pré-visualização disponível no *website* e possibilidade de descarregamento.

Por fim, é apresentada uma BD para registo 2D-3D existente no *Image Sciences Institute*, que tem associada uma descrição das operações realizadas para o registo das imagens.

Para além deste *site* de compilação de BDs de Imagem Médica encontramos outros que disponibilizam uma pequena BD própria. Por exemplo, em ‘*imageprocessingplace.com*’ encontramos uma série de conjuntos de imagens, organizados por tema e funcionalidade para que o utilizador possa extraí-los da plataforma na *Internet*, e utilizá-los no seu próprio trabalho de investigação ou processamento de imagem [7]. Encontram-se aqui imagens de livros de referência em processamento de imagem, associadas a potentes ferramentas no processamento de imagem digital (como é o caso da *toolbox* do MATLAB[®]), imagens mundialmente conhecidas que são utilizadas como referência para o estudo do tratamento de imagem (‘Lena’, ‘Cameraman’...), imagens de formas (usando o *standard* de descrição de conteúdo multimédia dos dados MPEG7), imagens de microscopia (adquiridas com microscópios de vários tipos), imagens médicas (de radiologia), *links* para várias outras Bases de Dados de imagens ou *datasets* com elas relacionadas, imagens espaciais e planetárias, imagens espaciais e meteorológicas da Terra, Bases de Dados de faces e impressões digitais, imagens de sequências de vídeo, imagens de e para segmentação, silhuetas de aeronaves, sombras de folhas, Bases de

Dados científicas públicas (que servem para estudos comparativos), Bases de Dados de imagens usadas para detecção do movimento de pessoas e veículos, Bases de Dados com anotações de pesquisa do ‘Image Sciences Institute’ (contém imagens da retina, radiografias do tórax, imagens de técnicas experimentais de co-registo, imagens do fígado e *scans* cerebrais de MRI), imagens de profundidade (adquiridas com 4 câmaras com diferentes profundidades, e usando diferentes níveis de intensidade).

Na BD de imagem do Instituto de Processamento de Imagem e Sinal (SIPI – *Signal and Image Processing Institute*) da Universidade da Califórnia do Sul, pode-se encontrar uma colecção de imagens digitalizadas desde 1977, com o objectivo de apoiar a pesquisa ao nível do processamento e análise de imagem, e também da visão computacional [8]. Estão disponíveis 4 volumes de imagens: texturas, imagens aéreas, imagens diversas e sequências de imagens representando movimento. Estas imagens possuem vários tamanhos (256x256 pixéis, 512x512 pixéis, ou 1024x1024 pixéis), mas todas têm 8 *bits* por pixel no caso de serem monocromáticas, e 24 *bits* por pixel no caso de serem a cores. As imagens encontram-se armazenadas no formato TIFF (*Tagged Image File Format*) e não é fornecido qualquer tipo de *software* de visualização para lhes aceder, apenas a possibilidade de descarregar a imagem.

Depois existem inúmeros locais de armazenamento de fotos dos mais diversos tipos, criados por pessoas a nível individual (muitas vezes anónimas), que partilham as suas aquisições com o resto da comunidade cibernauta, para partilhar o seu trabalho pessoal, pesquisas, e receber algum *feedback* de outros empreendedores nessa área. A título de exemplo temos o *site* pessoal de Fabien Peticolas, que disponibiliza uma Base de Dados de fotos diversificadas para processamento de imagem e sinal e avaliação digital das marcas de água [9].

Foram também encontrados livros que constituem uma autêntica BD com a apresentação das imagens dos exames médicos realizados nos processos médicos que discutem. A título de exemplo, Conti e Cham, 2005, catalogaram uma série de casos clínicos estudados com exames PET/CT e descritos por outros autores. Em cada um dos casos é apresentada um resumo do historial clínico do paciente, o que se descobriu com a execução do exame, a discussão do caso e algumas das imagens recolhidas. Constituindo assim uma BD de conhecimentos empíricos sobre os carcinomas incidentes, e a sua identificação através da imagiologia médica, neste caso com recurso em específico à PET/CT [10].

Relativamente a trabalhos mais parecidos com o que está a ser levado a cabo nesta tese, foram encontradas apenas duas publicações descrevendo pesquisas similares.

A primeira foi desenvolvida em Boston no Departamento de Medicina Nuclear/PET do Instituto de Cancro Dana-Farber, e consiste numa BD de dados clínicos que geralmente não se encontram nos sistemas PACS e HIS/RIS, que permite pesquisar/aceder a estudos antigos dos pacientes mediante um sistema otimizado de inserção e recuperação de dados. Jeffrey Yap e os seus colegas desenvolveram-na com mais de 100 variáveis (contendo entre elas o tempo de

aquisição PET, a dose de FDG, o tempo de captação, os dados de CT e o nível de glicose no sangue) e com o objectivo de assegurar que cada novo exame realizado a um determinado paciente é executado de acordo com as condições dos estudos anteriores [11].

No segundo caso, trata-se de uma BD relacional desenvolvida para efectuar pesquisas relativas a informações demográficas e a informações de exames pertencentes a pacientes que realizaram estudos de PET/CT. Esta plataforma foi desenvolvida no Medical College da Georgia, mais concretamente no Centro Ambulatório de Terapia por Radiação. Durante 3 anos de recolha de dados (desde Julho de 2003), procedeu-se à inserção de 2000 pacientes na BD. A implementação da BD implicou o desenvolvimento de fichas para recolher a informação do paciente e do estudo. Estas registavam o historial médico do paciente até ao momento e, após o seu preenchimento, eram lidas de forma automática para campos correspondentes na BD. Foi também implementada uma ferramenta na BD que devolve relatórios para suprirem as necessidades clínicas e administrativas das equipas de pesquisa. Eles incluem a contagem dos procedimentos de PET/CT realizados por mês, o volume de pacientes a nível administrativo, os dados demográficos dos pacientes, e o volume de pacientes que foram indicados para a realização de exames pelos médicos. Estes dados podem ser filtrados por indicações do exame, tipo de exame, sexo, raça, idade e factores de risco do paciente. Esta BD revelou-se útil para o fornecimento de relatórios financeiros e para monitorizar a produtividade e desenvolver orçamentos futuros. As ferramentas implementadas fornecem tratamentos estatísticos úteis para publicações e apresentações de pesquisas [12].

1.4. Organização da Tese

Este documento encontra-se dividido em duas partes. Uma primeira parte onde são dadas informações acerca da autoridade do seu criador sobre ele (direitos de autor), da forma como foi escrito (notas do autor, acrónimos), da estrutura que segue (índices) e onde é apresentado um breve resumo do projecto levado a cabo (sumário/*abstract*). A segunda parte consiste no conteúdo do trabalho propriamente dito, e é composta por 10 capítulos, cujo nome e uma breve descrição são apresentados na tabela 6.

Tabela 6 - Breve descrição da estrutura e conteúdo do documento.

Nome do Capítulo	Breve Descrição do seu Conteúdo
Introdução	Apresenta-se o problema em causa, descrevendo o contexto em que se insere. Define-se o objectivo deste trabalho e apresenta-se a forma como o documento se encontra organizado.
Gestão do Projecto	Breve descrição das pessoas e entidades associadas a

	este projecto e do planeamento inicial das tarefas a levar a cabo para concretizá-lo.
Fundamentos Teóricos	Definição dos conceitos teóricos chave: imagem médica, PET/CT, BD e Interface Gráfica.
Equipamento Utilizado e Dados Recolhidos	Descrição da tecnologia e ferramentas utilizadas, e dos dados possíveis de serem extraídos dos ficheiros gerados aquando da realização de um exame.
Estudo das Variáveis Existentes	Estudo aprofundado da meta-informação existente nos cabeçalhos dos ficheiros <i>imago</i> , correspondendo uma variável a um dos campos existentes no “mapa de bytes” que descreve os cabeçalhos (cabeçalho principal e sub-cabeçalhos) de cada ficheiro.
Organização da Informação	Apresentação da estrutura que se definiu para hierarquizar e estabelecer associações entre a informação recolhida.
Método de Implementação da Base de Dados	Explicação dos passos realizados desde a formulação dos requisitos da BD, passando pelos casos de uso e pela construção dos diagramas conceptual e físico, até à geração e preenchimento da BD.
Desenvolvimento da Interface de Comunicação com a Base de Dados	Exposição da Interface desenvolvida para estabelecer a comunicação entre o utilizador e os dados armazenados na BD, permitindo-lhe visualizar, manipular, pesquisar, processar e gerir a informação.
Testes Realizados e Resultados Obtidos	Síntese das funcionalidades da plataforma final (resultados), e de alguns testes realizados para avaliar a qualidade da mesma (principalmente do seu desempenho a nível temporal comparativamente a uma análise manual dos dados).
Conclusão	Principais conclusões tiradas da realização deste trabalho e apresentação de melhorias que podem vir a ser implementadas.

Capítulo 2

Gestão do Projecto

2. Gestão do Projecto

Entende-se por gestão um conjunto de tarefas que procuram atingir os objectivos pré-determinados para um dado empreendedorismo, usando de forma eficaz todos os recursos disponibilizados pela instituição onde se encontra a ser realizado. [13]

Nesta secção vão ser apresentadas informações acerca das pessoas e das entidades ligadas a este projecto, os objectivos iniciais e o planeamento desenvolvido para executá-lo.

2.1. Membros do Projeto

Este projecto foi realizado por uma aluna do 5.º ano do Mestrado Integrado em Engenharia Biomédica sob a orientação dos Eng.ºs Francisco José Caramelo e Nuno Chichorro Ferreira. Na tabela 7 são apresentadas informações detalhadas sobre cada um dos membros deste projecto.

Tabela 7 - Informações sobre os membros do projecto.

Nome	Função	Grau Académico	Contacto
Francisco José Caramelo	Orientador	Doutorado em Ciências da Saúde pela UC. Exerce funções de Professor Auxiliar na FMUC. Investigador nos centros IBILI e ICNAS.	fcaramelo@ibili.uc.pt
Nuno Chichorro Ferreira	Orientador	Doutorado em Ciências Biomédicas pela UC. Exerce funções de Professor Auxiliar no IBILI (FMUC). Investigador nos centros IBILI e ICNAS.	nuno@ibili.uc.pt
Raquel Trindade	Mestrando	Licenciada em Ciências da Engenharia Biomédica. Aluna do Mestrado Integrado em Engenharia Biomédica com especialização na área de Informática Clínica e Bioinformática.	raquelmarisa3@hotmail.com

2.2. Apresentação das Instituições de Acolhimento

O ICNAS é um centro multidisciplinar da UC, que tem como principais objectivos o desenvolvimento da investigação científica com recurso a novas técnicas e tecnologias de Medicina Nuclear aplicadas à saúde, a divulgação dos avanços alcançados no decorrer da sua

actividade, a prestação de serviços especializados mediante a aplicação de radiações na área da Biomedicina, o alcance de uma maior interdisciplinaridade através da exploração de novas associações entre as áreas científicas relacionadas com a sua actividade, e, por fim, a promoção a nível nacional e internacional da comunicação entre as diversas instituições que realizam investigação, educação e prestação de serviços na área dos cuidados de saúde. [14]

Para concretizar estes propósitos, esta entidade dispõe de uma unidade completa de Medicina Nuclear: uma unidade clínica equipada com 5 Câmaras Gama, uma câmara de PET/CT, e uma unidade de produção constituída por um Laboratório de Radiofarmácia e um Ciclotrão (acelerador de partículas utilizado para a produção de radionuclídeos artificiais que podem ser utilizados em PET [4]). Este equipamento encontra-se instalado num edifício moderno e especificamente concebido para o seu tipo de actividade (destacando-se as suas características associadas à protecção radiológica), localizado no Pólo III – Pólo das Ciências da Saúde – da UC.

Estas instalações estão em funcionamento desde 2009, e o centro é dirigido pelo Doutor Miguel Castelo-Branco (Professor Auxiliar na FMUC, que dirige também o IBILI [15]). Existe também um Conselho Científico constituído por 15 elementos que trata da definição das políticas de investigação, actividades, regimento e Estatutos a serem implementadas no ICNAS.

Na tabela 8 podemos ver sintetizados os recursos tecnológicos [16] de que esta instituição dispõe, bem como os tipos de exames clínicos nela realizados [14].

Tabela 8 - Recursos tecnológicos de que o ICNAS dispõe (à esquerda) e exames médicos nele realizados (à direita). (Adaptado de: [14], [16])

Recursos	Exames
<ul style="list-style-type: none"> - Ciclotrão IBA Cyclone 18/9 - Radioquímica - <i>Scanners</i>: <ul style="list-style-type: none"> • ClearPEM (PET) • Philips Gemini GXL (PET/CT) • SIEMENS Trio Tim 3T (MRI) • 2 GE Millenium MG (SPECT) • GE Millenium (SPECT) • GE Starcam 600 XR/T (SPECT) • GE Starcam 400 AC/T (SPECT) • Hologic QDR 4500 Elite (Densitometria Óssea) - Fantasmas - Fontes - <i>Software</i> - Várias <i>Hot Cells</i> (existem pelo menos 4 câmaras blindadas de contenção de radiação nuclear para proteger os profissionais dos isótopos radioactivos) - Módulos de Síntese de Radiofármacos 	<ul style="list-style-type: none"> - Aparelho cardiovascular: <ul style="list-style-type: none"> • Angiografia de Radionuclídeos (ARN) • Estudo de perfusão do miocárdio (Tálio, Sestamibi, Tetrafosmina) - Sistema músculo-esquelético: <ul style="list-style-type: none"> • Cintigrafia do esqueleto • Cintigrafia com TL-201 • Cintigrafia com MIBI (Tc-99m) - Glândulas endócrinas: <ul style="list-style-type: none"> • Cintigrafia da tiroide com Tc-99m • Cintigrafia corporal com DMSA (Tc – 99m) • Estudos de distribuição do MIBG (Iodo 123/131) • Cintigrafia das paratiroides com TI-201/MIBI (Tc-99m) • Cintigrafia corporal com TL-201 • Cintigrafia com análogos da somatostanina • Estudo de distribuição do iodo-colesterol - Aparelho digestivo: <ul style="list-style-type: none"> • Cintigrafia Hepatobiliar e hepatoesplénica • Cintigrafia hepática com glóbulos vermelhos marcados • Pesquisa de hemorragia digestiva • Estudo funcional das glândulas salivares • Trânsito esofágico e Esvaziamento gástrico • Pesquisa de refluxo gastroesofágico • Cintigrafia intestinal com leucócitos marcados • Divertículo de Meckel

	<ul style="list-style-type: none"> - Aparelho urinário <ul style="list-style-type: none"> • Cintigrafia Renal (DTPA; DMSA; MAG3) • Renografia • Cistografia directa - Avaliação de infecção e inflamação / Estudos hematológicos: <ul style="list-style-type: none"> • Estudo de distribuição dos leucócitos marcados • Cintigrafia da medula óssea • Estudo da distribuição de Ga-67 - Aparelho respiratório: <ul style="list-style-type: none"> • Cintigrafia pulmonar de perfusão e ventilação - Tomografia de Emissão de Positrões (PET/CT)
--	---

O IBILI (Instituto Biomédico de Investigação de Luz e Imagem) é também um centro que promove a investigação, educação e formação ao nível das ciências da saúde, a um nível multidisciplinar, principalmente integrando as áreas da Medicina, Biologia e Engenharia. É reconhecido internacionalmente, tendo sido classificado como ‘Excelente’ pela Fundação para a Ciência e Tecnologia (FCT) nos anos de 1999, 2004 e 2008 [17].

As suas instalações situam-se ao lado das do ICNAS, e incluem óptimas clínicas e laboratórios. A sua equipa é constituída por pessoal altamente qualificado, na medida em que esta Instituição tenta promover e apoiar a aquisição dos mais elevados níveis de educação profissional, especialmente no ensino de pós-graduação.

2.3. Planeamento Inicial vs. Tarefas Delineadas para a sua Execução

2.3.1. Planeamento Inicial

Inicialmente foram definidos os seguintes pontos como metas a atingir ao longo da realização deste trabalho:

- Descrição dos cabeçalhos dos ficheiros de imagem (com especial atenção à existência de campos vazios na estrutura destes).
- Explorar a melhor forma de identificar os registos guardados na BD: qual a chave conjugada ideal de forma a não violar a privacidade dos doentes.
- Módulos: desenvolver programas que sejam passíveis de ser integrados noutros já previamente estabelecidos, e que tenham capacidade de receber outras plataformas modulares a serem desenvolvidas.
- Interface: pretende-se a apresentação de uma janela organizada de forma a poder apresentar a informação de dezenas de exames.

- Averiguar a melhor forma de manter a confidencialidade dos dados, garantindo a privacidade dos doentes.
- Implementar restrições de forma a prevenir a introdução de um mesmo paciente diversas vezes na BD, devido a diferenças na introdução do seu nome (por exemplo, ao nível das maiúsculas e minúsculas definidas: “João Alberto” e “João alberto”), ou de qualquer outro conjunto de dados que o identifiquem. Fazer programas que procurem campos comuns a todos os pacientes (tais como: o nome, a idade, o sexo e a residência) que tenham valores iguais para pacientes diferentes, e fazer uma indicação de um possível erro/redundância naquele registo “duplicado”.
- Tentar reduzir os erros existentes nos dados: ao nível da inserção de dados e da maior uniformização possível dos dados já introduzidos pelo utilizador da aplicação.
- Mudança nos ‘Protocolos de Aquisição’ e ‘Registo de Dados’ (apesar das dificuldades devido a toda a componente burocrática implicada).
- ‘Cabeçalho Padrão’: estudar as variáveis comuns aos diversos tipos de cabeçalhos e exames, e também as não comuns mas deveras importantes, de forma a seleccionar aquelas que devem estar presentes na constituição da BD.
- Definição de uma forma de adicionar dados à BD – descrição desta.
- Elaborar ferramentas de pesquisa de forma a procurar, por exemplo, por características do paciente e elaborar estatísticas a partir destas: averiguar a probabilidade/possibilidade de o paciente vir a sofrer de determinada doença, avaliar o risco de agravamento da sua condição de saúde actual, fazer histórico da evolução do estado de saúde do paciente (representações gráficas).
- Aplicar filtros na passagem dos dados da BD ligada às máquinas de aquisição dos exames para a BD a ser construída.

2.3.2. Tarefas Delineadas para a sua Execução

Foram planeadas as seguintes acções a fim de concretizar os desafios propostos:

- Recolha dos dados presentes nos ficheiros *imago* (meta-informação presente nos cabeçalhos de cada ficheiro) usando as ferramentas fornecidas pela Philips Medical Systems® para o efeito (designadas ‘Luminary Tools Package’). Tratam-se de rotinas desenvolvidas em IDL® e MATLAB® para a extracção da informação presente no mapa de *bytes* de cada ficheiro.
- Proceder à organização da informação recolhida de acordo com temas (sugeridos no manual de descrição de apoio ao uso das ‘Luminary Tools’), utilidade (p.e., armazenamento das directorias dos ficheiros para que seja mais simples aceder-lhes, bem como o armazenamento de dados processados na BD), necessidade (registo da informação pessoal associada aos utilizadores da BD), etc..

- Selecção apenas da informação relevante (ou seja, de variáveis com importância para o estudo/investigação).

- Desenvolvimento de um modelo conceptual, usando as ferramentas do Sybase Power Designer 15[®]: desenvolver um modelo relacional para inserir as variáveis comuns a todos os cabeçalhos dos exames, e adicionar tabelas EAV para o armazenamento de novas variáveis que possam surgir ao longo do tempo. Isto resultou no desenvolvimento de um Modelo Conceptual Híbrido para a nossa BD.

- Gerar a BD em si recorrendo à linguagem PostgreSQL[®], ao nível da interface de visualização de dados pgAdmin III[®].

- Criação de rotinas em MATLAB[®] para extrair a informação dos ficheiros dos exames (de extensão ‘.img’ e ‘.syn’), organizá-la, corrigir campos eventualmente corrompidos e inserir estes dados (e outros resultantes do processamento destes) na BD gerada – preenchimento da BD.

- Desenvolver uma interface gráfica em MATLAB[®] para permitir a visualização, manipulação, correlação e pesquisa da informação contida na BD.

- Realizar testes para a validação do trabalho desenvolvido.

Capítulo 3

Fundamentos Teóricos

3. Fundamentos Teóricos

Para a realização deste trabalho foi necessário o estudo aprofundado de uma série de conceitos teóricos relacionados com o problema em estudo. Como o pretendido seria a criação de uma BD de imagem médica, mais concretamente de exames de PET/CT (posteriormente expansível para outras modalidades de imagem existentes no ICNAS), analisaram-se as definições de imagem médica e BD e os princípios envolvidos na técnica. Pesquisou-se também qual a melhor forma para desenvolver uma interface com o utilizador, que satisfizesse os requisitos de funcionalidades e facilidade de manipulação.

3.1. Imagem Médica

Começamos então por definir o que é uma imagem médica: trata-se de uma representação do todo ou apenas de uma parte do corpo humano, recolhida *in vivo*, que apresenta informação morfológica (características anatómicas, como por exemplo ossos partidos ou tumores) e/ou funcional (acerca de um estado/processo fisiológico, representando modificações/desregulações na actividade de certos órgãos) normalmente com o intuito de realização de um diagnóstico, uma terapia, um planeamento de cirurgia ou uma pesquisa ao nível clínico (fins educacionais). As imagens recolhidas localizam-se maioritariamente ao nível do crânio (mais especificamente do cérebro), peito (coração e caixa torácica), abdómen (estômago e intestino) e osso. As imagens médicas são quase sempre mostradas em tons de cinzento (escala monocromática) em vez de cor, exceptuando-se as imagens de ultra-som Doppler e as imagens pseudo-cor de medicina nuclear [18], [19].

O nascimento da imagem médica pode ser associado à descoberta dos raios-X por Wilhelm Conrad Roentgen em 1895 e subsequente realização da primeira radiografia alguma vez feita à mão esquerda da sua esposa, mostrando o seu esqueleto e a sua aliança [18].

A grande maioria das imagens médicas são produzidas ao nível dos departamentos de Radiologia, no entanto uma pequena parte do volume total de dados imagiológicos consiste em fotografias anatómicas, microfotografias histológicas, imagens endoscópicas e de OCT, entre outras. As imagens radiológicas capturam propriedades físicas do corpo, tais como a opacidade (nos exames de raios-X, nomeadamente, radiografias e *scans* CT), a reflectância (nos ultrasons), a concentração de água ou de outro químico (nas imagens de RMN) e a distribuição de radiotraçadores dentro de um órgão (como acontece na PET), em condições e com uma qualidade que as técnicas tradicionais usando filme radiográfico não conseguem alcançar. Isto

decorre do facto dos sensores actuais disporem de visualizadores de alta qualidade e elevada resolução que disponibilizam melhores características (melhor sensibilidade e o facto de estarem integrados em redes de elevada velocidade de transferência de dados, que os guiam até sistemas de armazenamento sem custos adicionais) comparativamente aos filmes e papéis (utilizados pelos clínicos para gerir os dados de registo dos pacientes) produzidos pelos sistemas tradicionais de radiologia [19]. Estes últimos tendem a excluir informação complementar proveniente de outros sistemas de informação clínica, e o registo manual acarreta geralmente custos adicionais e perda de produtividade devido a registos perdidos, ficheiros mal colocados, estudos duplicados e à necessidade de um grande espaço físico para armazenamento. Assim sendo, constata-se que cada vez mais se pretende a implementação de um sistema totalmente digital, sem recurso a suportes físicos (p.e., filmes e papel), de modo a que a informação digital seja integrada e partilhada pelos visualizadores e Bases de Dados de toda a rede hospitalar. Neste sentido, existem diferentes tipos de digitalizadores de filmes (*scanners* de laser, câmaras de estado sólido, câmaras de vídeo, entre outros) que podem ser usados para converter filmes de raios-X (imagens analógicas) em formatos digitais (imagens digitais) para futuro processamento [1].

As imagens médicas são representadas digitalmente em diversos formatos dependendo da modalidade, da zona anatómica analisada e da técnica de recolha de imagem utilizada. O tamanho da imagem e o número de imagens recolhidas no exame de um paciente varia entre modalidades. Uma imagem médica bidimensional (2D), tendo M pixéis de altura, N de largura e 2^k níveis de cinzento, terá $M \times N \times k$ bits. Na tabela 9 vemos o número médio de megabytes (MB) por exame gerado pelas diferentes técnicas de imagiologia médica, onde uma imagem de 12 bits é representada por 2 bytes na memória [18].

Tabela 9 - Dimensões e tamanhos das imagens biomédicas. (Adaptado de: [18])

Modalidade	Dimensões Típicas da Imagem	Escala de Cinzentos (<i>bits</i>)	Tamanho Médio/Exame
Medicina Nuclear	128×128	8 ou 16	2 MB
MRI	256×256	12	8-20 MB
Ultra-som	512×512	8	5-8 MB
Ultra-som <i>Doppler</i>	512×512	24	15-24 MB
Angiografia de Subtracção Digital (<i>Digital Subtraction Angiography - DAS</i>)	512×512	8	4-10 MB
CT	512×512	12	20 MB
CT em Espiral ou Helical	512×512	12	40-150 MB
Microscopia Digital Electrónica (<i>Digital Electronic Microscopy - DEM</i>)	512×512	8	Varia
Microscopia Digital Colorida (<i>Digital Color Microscopy - DCM</i>)	512×512	24	Varia
Caracterização Cardíaca	512×512 ou 1024×1024	8	500-1000 MB
Filmes Raio-X Digitalizados	2048×2048	12	8
Radiografia Computorizada	2048×2048	12	8-32 MB
Mamografia Digitalizada	4096×4096	12	64 MB (um par)

Para que se possa proceder à gestão e partilha de imagens digitais são necessários uma série de passos: aquisição dos dados, armazenamento dos mesmos, gestão do sistema de armazenamento (da BD), compressão dos dados (para poupar espaço de armazenamento e facilitar a transmissão), transmissão destes (geralmente envolve a conversão dos dados imagiológicos do formato definido pelo fabricante da tecnologia que os adquiriu, para a representação DICOM, que será abordada mais adiante nesta secção) e, por fim, a preparação da imagem para que haja condições óptimas de visualização desta na estação de exibição (para apresentar devidamente a imagem no ecrã ela deve ter o tamanho formatado correcto, bom ajuste de luminosidade e contraste, orientação correcta e nenhum fundo capaz de distrair o observador) [1], [19].

Os maiores desafios enfrentados pelas Bases de Dados de imagem médica são a integração com o sistema de fluxo de informação do Hospital e a interoperabilidade entre si.

Para tentar resolver esta questão de adaptação do sistema hospitalar ao armazenamento e comunicação da imagem digital de um modo rápido, fácil e seguro, desenvolveu-se o conceito de Sistema de Arquivamento e Comunicação de Imagens (*Picture Archiving and Communication System* – PACS). Foi criado originalmente nos anos de 1980, mas o seu maior desenvolvimento (bem como período de implementação) ocorreu na década de 1990. As funções do PACS incluem: aquisição da imagem e de dados, padrões de comunicação de dados médicos, pré-processamento de imagens médicas, gestão de armazenamento hierárquico, recuperação e visualização de imagens. Todos estes elementos estão interconectados por redes de banda larga e servidos por um poderoso controlador que garante o fluxo confiável e atempado de dados multimédia desde os computadores de aquisição até às estações de visualização de destino. Um PACS de larga escala inclui: *scanners* de aquisição de imagem médica, um sistema de informação radiológica (*Radiology Information System* - RIS), computadores de aquisição de dados, sistemas de armazenamento de ficheiros, plataformas de trabalho para radiologistas (para visualização de exames e realização de relatórios médicos) e, por vezes, sistemas de informação hospitalares (*Hospital Information System* - HIS) interconectados por redes computacionais. Esta integração de larga escala providencia comunicação correcta e atempada entre profissionais e departamentos de forma conveniente e eficiente, suporta instalações dispersas e teleradiologia (transmissão de imagens radiológicas entre diferentes locais, geralmente distantes, para diagnóstico ou tratamento), diminui desperdícios (evitando repetição e perda de exames), reduz os custos com filmes e pessoal, e facilita a reunião de dados para pesquisas clínicas. Da instalação de um sistema deste porte incorrem custos muito elevados, pelo que as instituições hospitalares optam por instalar apenas sistemas confinados a uma única modalidade digital, designados de mini-PACS [1].

A aquisição e partilha de informação radiológica requerem a integração de vários tipos de informação: vocal, textual, imagens e gravações de vídeo. Tipicamente os radiologistas

primeiro registam as descobertas de um estudo num sistema de voz digital, depois estes registos vocais são associados às imagens adquiridas (ou gravações de vídeo) e aos registos clínicos textuais dos pacientes e, por fim, são arquivos dentro de uma BD. O avanço das tecnologias computacionais e *standards* de comunicação permitiram a integração destes conjuntos de dados autónomos e heterogéneos.

A abordagem corrente para interligar quaisquer dois dispositivos médicos envolve um formato padrão para a troca de dados e um protocolo de comunicação para a transferência de mensagens. Por exemplo, o padrão reconhecido a nível industrial para a comunicação textual é o HL7 (*Health Level 7*): conjunto de normas que tentam uniformizar a forma de transmissão de dados, para que se estabeleça comunicação entre aplicações computacionais que diferem no seu *hardware* (provenientes de diferentes fornecedores), modo de operação e representação de dados; define o formato de dados e protocolo para a troca de ‘dados textuais chave’ para a comunicação entre sistemas de informação clínica, tais como os sistemas de informação radiológica, hospitalar (registos administrativos dos pacientes no Hospital), laboratorial (análises clínicas) e farmacêutica (medicação prescrita). No entanto, o método padrão para a comunicação de dados imagiológicos dentro da rede hospitalar é o formato DICOM (*Digital Imaging and Communications in Medicine*). Este *standard* serve já como estrutura base para a comunicação em imagiologia médica. É uma ferramenta que funciona ao nível da aplicação no modelo das 7 camadas (a camada 7, que é a camada mais superior). Foi desenvolvido pelo Colégio Americano de Radiologia (*American College of Radiology - ACR*) e pela Associação Nacional Eléctrica de Fabricantes (*National Electrical Manufacturers Association - NEMA*). O DICOM fornece formatos padronizados para imagens, um modelo comum de informação, definições do serviço da aplicação e protocolos para comunicação. O objectivo destas características é tornar o equipamento de imagiologia médica interoperável.

No caso das tecnologias utilizadas para a realização deste projecto, o *standard* DICOM é utilizado geralmente para a visualização só da parte CT dos exames. Para a recolha da meta-informação e visualização conjugada das imagens de PET e de CT usa-se o formato *imago*.

3.2. Técnica PET/CT

As imagens usadas na realização deste projecto são da modalidade híbrida PET/CT utilizada em Medicina Nuclear. Na União Europeia (UE), a Medicina Nuclear foi reconhecida como especialidade médica autónoma pela Directiva Comunitária de 1989, e que em 2000 existia em 24 dos 26 países europeus (incluindo países não abrangidos pela UE) [4].

A Medicina Nuclear utiliza propriedades dos núcleos da matéria para investigar perturbações no metabolismo, na fisiologia e na anatomia humanos de modo a diagnosticar uma possível patologia, e para tratar doenças já diagnosticadas com fontes radioactivas não seladas. As técnicas próprias da Medicina Nuclear são fáceis de executar, estão associadas a muito baixa morbidade e, virtualmente, a nenhuma mortalidade. As doses de radiação absorvidas pelos doentes são, quase sempre, semelhantes ou inferiores às dos métodos radiológicos convencionais, contrariamente ao muitas vezes suposto [4].

Mas em que consiste um exame de PET/CT? Quais as suas bases físicas?

A Tomografia por Emissão de Positrões com Tomografia Computorizada (PET/CT) é uma técnica híbrida de Medicina Nuclear usada para diagnóstico, estadiamento, monitorização da terapia e avaliação da recorrência em Oncologia, assim como em muitas outras aplicações em Neurologia, Cardiologia, Farmacologia e pesquisa/investigação. Uma única aquisição deste tipo pode gerar muitas centenas de imagens funcionais PET (que, em muitos casos, fornecem um diagnóstico precoce e mais preciso, uma vez que, em situação de doença, as alterações funcionais e moleculares precedem as alterações morfológicas) associadas a referências anatómicas provenientes da CT (também várias centenas de imagens) [20], [21].

Trata-se de um exame inócuo, de simples execução, minimamente invasivo, indolor, cujo resultado pode ser determinante na orientação do diagnóstico bem como na programação terapêutica, particularmente no doente oncológico. Ao nível oncológico este tipo de exame está indicado para: distinguir entre lesões benignas e malignas; determinar o grau de malignidade; estabelecer o estágio da doença; averiguar a existência de recorrência ou doença residual; investigar a localização da recidiva perante a subida de marcadores tumorais; avaliar a resposta à terapêutica – imagens de antes, durante e após o tratamento; identificar a localização de um tumor primário desconhecido, para biópsia ou para fins terapêuticos [14].

A execução da técnica PET/CT inicia-se com a selecção do traçador apropriado (um fármaco marcado com um radionuclídeo emissor de positrões – Fig. 2) que é administrado no organismo do paciente (mas em concentrações moleculares tão reduzidas que não alterem os processos fisiológicos normais). Normalmente, depois de um período de captação, onde as moléculas se distribuem pelo corpo, o paciente é posicionado então dentro do *scanner* (usualmente com uma disposição cilíndrica dos detectores) e a aquisição de imagens (tridimensionais) começa. Este processo pode demorar desde 10 a mais de 30 minutos, dependendo do paciente, da aplicação clínica (e o protocolo de aquisição correspondente) e das características da máquina onde o exame está a ser realizado.

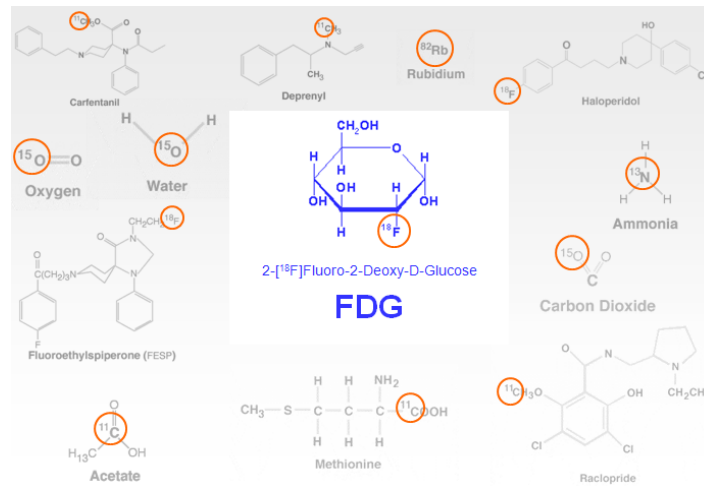


Figura 2 - Representação da estrutura química de alguns traçadores em PET (os círculos assinalam o isótopo emissor de positrões que é usado para localizar a molécula). Cada substância deste tipo permite estudar funções particulares do corpo humano, de acordo com o caminho da molécula e as reacções bioquímicas onde ela participa. As moléculas marcadas funcionam como sinais fluorescentes, pois na imagem, as zonas mais activas do organismo vão surgir como pontos luminosos. O radiotraçador mais utilizado é o análogo da glucose (flúordesoxiglicose - FDG), marcado com Flúor 18, um emissor de positrões. A glucose é um dos fornecedores de energia celulares, e assim vai concentrar-se nas zonas do corpo com mais gasto energético, que corresponderão às zonas cancerosas, devido ao facto de estas células serem grandes consumidores de energia. (Adaptado de: [22])

Durante este período de tempo, fisicamente falando, ocorre o seguinte:

1. Os emissores de positrões são isótopos deficientes em neutrões que alcançam a sua estabilidade através da transmutação de um protão do seu núcleo na combinação neutrão+positrão+neutrino (Fig. 3).
2. Cada positrão emitido irá perder a sua energia através de interacções com os tecidos circundantes, até se aniquilar com um electrão.
3. Em cada aniquilação são emitidos 2 fotões gama, em direcções quasi-colineares, em sentidos opostos que podem ser detectados em coincidência (Fig. 4-a) por dois detectores.
4. Muitos pares de fotões coincidentes (Fig. 4-b) são detectados e geralmente guardados em matrizes (sinogramas), onde cada elemento representa uma linha de projecção da distribuição de actividade no paciente num ângulo e posição radial específicos (Fig. 4-c) [20].

Para finalizar, é aplicado um algoritmo de reconstrução de imagem aos dados para recuperar a distribuição de actividade na forma de séries de imagens, mapeando assim indirectamente o processo funcional (Fig. 4-d) que criou a distribuição do emissor de positrões [20]. O estudo CT que é realizado no mesmo sistema complementa a informação PET, ajudando na localização das lesões e na correcção para a atenuação da radiação.

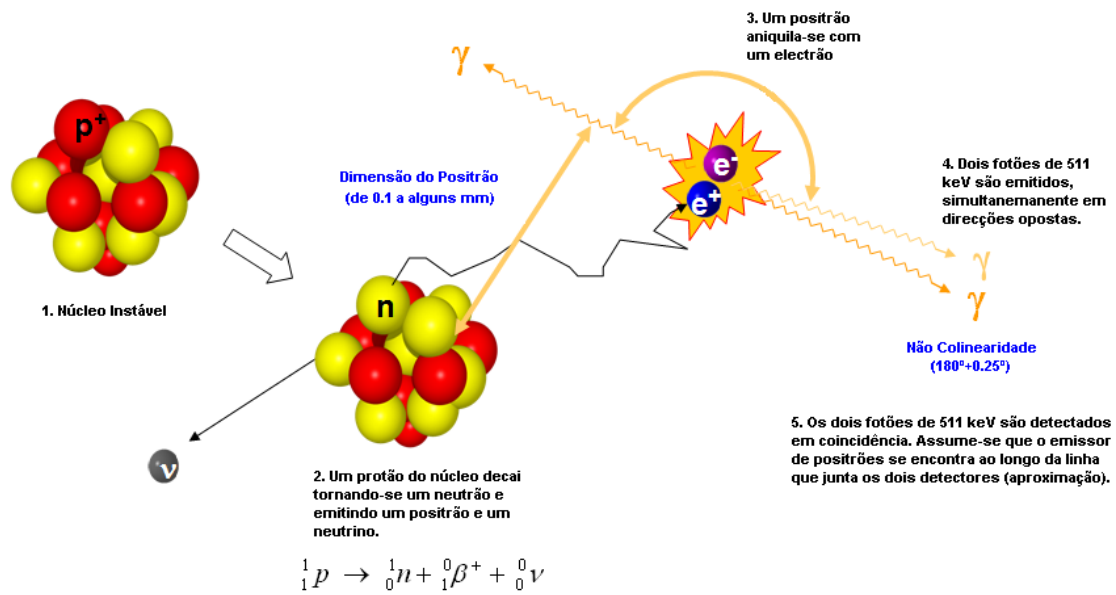


Figura 3 - Reacção nuclear que está na base do processo de PET/CT. (Adaptado de: [22])

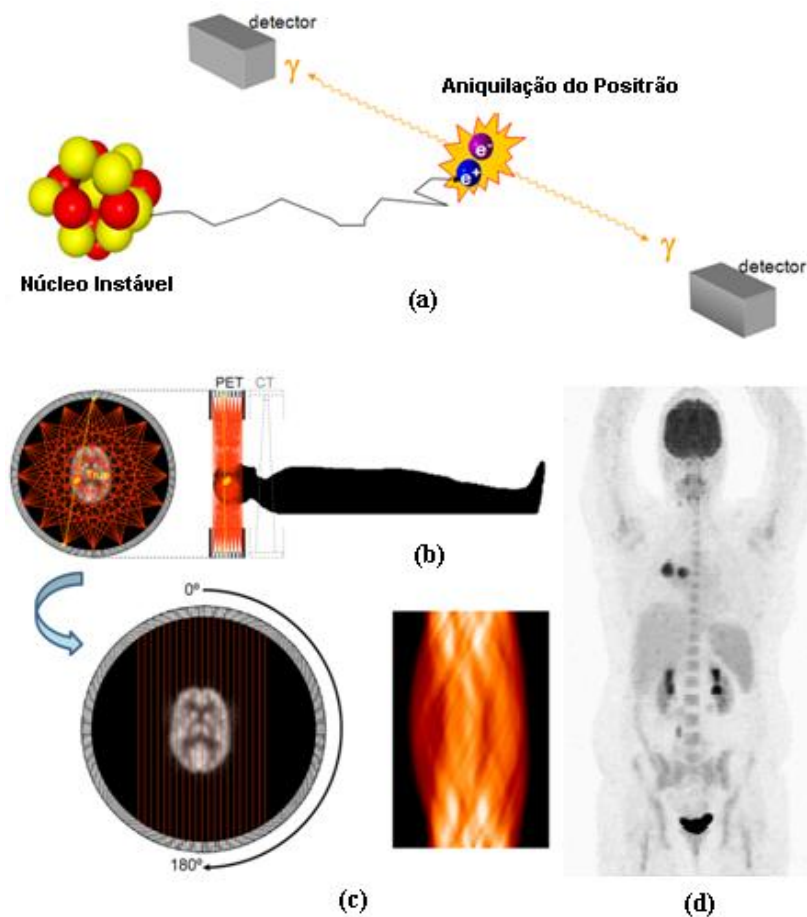


Figura 4 - Sequência de eventos num procedimento PET/CT. (a) Representação do sinal em PET (detecção de coincidências verdadeiras): dois fótons de 511 keV são emitidos em sentidos opostos a partir da aniquilação de um pósitron e são detectados quase simultaneamente, sem interagir no paciente. (b) Os scans de CT e PET são realizados sequencialmente no tempo. No de PET, os múltiplos detectores (cristais de cintilação) em coincidência definem milhares de linhas de resposta (*Lines Of Response* – LORs), cada uma contando coincidências verdadeiras. (c) As linhas de resposta são agrupadas em grupos de LORs paralelas

correspondendo a múltiplas vistas da distribuição de radioisótopo (projeções). As várias projeções de 0° a 180° podem ser observadas como um sinograma. Os dados de projecção são então reconstruídos para produzir as imagens finais. (d) Corte coronal de uma PET/CT de corpo inteiro, realizado com FDG, com dois tumores visíveis nos pulmões. (Adaptado de: [22])

Cada exame radiológico envolve um procedimento bem definido, e deve proceder-se à utilização de mais que uma modalidade de imagiologia para apoiar o processo de tomada de decisão clínica [18].

Existem vários protocolos que podem ser usados para a recolha de imagens dos pacientes em aparelhos de PET/CT, no entanto o procedimento padrão (Fig. 5) de execução de um exame deste tipo envolve: administração oral ou intravenosa do contraste usado em CT, aproximadamente uma hora antes do *scan* (no ICNAS não é usado; injecta-se, sim, o paciente com uma pequena concentração de glicose radioactiva, FDG, numa veia periférica; a quantidade administrada é determinada previamente de acordo com o peso do doente; há um período de espera, de cerca de 1 hora, para a distribuição e captação do radionuclídeo); posicionamento do paciente no *scanner* e imobilização deste, se necessário (por exemplo, com a utilização de dispositivos para reduzir o movimento do pescoço); depois realiza-se a parte CT do exame, com a aquisição de uma radiografia digital piloto (“topograma”) para a visualização de todo o corpo do paciente, e que permite depois a selecção da área de interesse; passa-se então à aquisição PET (no ICNAS, a aquisição das imagens propriamente dita demora entre 25 a 35 minutos); por fim, executam-se as correcções de atenuação e de dispersão, usando os coeficientes de atenuação retirados da porção CT da aquisição. Obtêm-se assim imagens co-registadas para futura interpretação por um profissional de saúde [14], [23].

Saliente-se que o protocolo deve ser sempre ajustado para a aquisição da zona do corpo que se pretende avaliar, e principalmente ao nível da escolha do traçador usado para monitorizar o processo, do tempo e do local de administração.

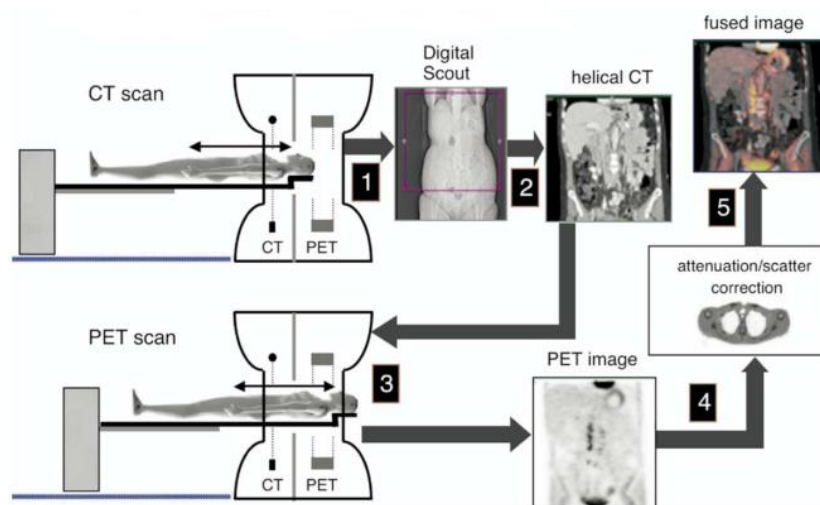


Figura 5 - Protocolo padrão de realização de um PET/CT segundo Blodgett et al., 2006. Em primeiro lugar realiza-se uma radiografia digital piloto (“topograma”) de todo o paciente (1), que permite definir a região a

examinar, seguindo-se a parte CT do exame (2). De seguida realiza-se a PET (3), e depois as correcções de atenuação e de dispersão (4). O resultado final é a imagem de fusão (5) da parte CT com a parte PET. (Imagem adaptada de: [23])

No procedimento normal de realização de um exame de PET/CT no ICNAS, o tempo total de permanência na instituição é de aproximadamente 2 a 4 horas. Antes do exame o paciente deverá fazer um período de jejum de pelo menos 4 a 6 horas, e beber entre 0,5 a 1 litro de água; deve também dar conhecimento das seguintes situações: o seu peso (a fim de se determinar a dose a administrar), se possui diabetes ou intolerância à glicose (por causa da administração de FDG), se tomou corticóides e/ou factores hematopoiéticos (que podem afectar a fixação do radiofármaco), se está a realizar quimioterapia e quando foi a sua última sessão (o exame só deve fazer-se 3 semanas após o último ciclo) e se está a realizar radioterapia e quando foi o último tratamento (deve aguardar-se um período de 2 meses desde esse até à realização de um exame PET) [14].

3.3. Bases de Dados

Uma BD serve para analisar resultados de uma experiência, guardar e analisar dados resultantes de processamento por máquinas (na maior parte das vezes computadores), descobrir aspectos estatísticos em dados empíricos, gerir uma actividade, entre outras tarefas. Chama-se *data warehouse* à infraestrutura que contém os dados para processamento, análise e mineração de dados (*data mining* - consiste em aplicar algoritmos de extracção de informação útil a partir dos dados; por exemplo: encontrar padrões comuns nos dados) [24].

Hoje em dia, as Bases de Dados são essenciais para o registo e gestão da maioria das actividades humanas na sociedade moderna. As BDs empregadas nas transacções comerciais (negócios e empresas) são, sem dúvida, as mais utilizadas no dia-a-dia, e, ao nível da investigação científica, as áreas que mais se servem das suas potencialidades são: a astronomia (para o armazenamento de informação relativa aos inúmeros corpos celestes estudados), a genética (investigação do genoma humano), a bioquímica (exploração das propriedades medicinais das proteínas), a medicina (com os grandes volumes de dados vocais, audiovisuais, textuais e imagiológicos gerados todos os dias), a engenharia (no desenvolvimento de novas ferramentas e tecnologias, por exemplo) e a educação de um modo geral (armazenamento da informação para futura transmissão de conhecimentos) [25], [26].

Inicialmente as BDs eram constituídas por informação maioritariamente textual e numérica; estes sistemas são agora designados ‘Aplicações Tradicionais de BDs’. O grande impacto das BDs associa-se ao crescente e comum uso dos computadores, devido ao facto de

desempenharem um papel crítico em quase todas as áreas em que estes são utilizados. A evolução tecnológica nos últimos anos levou ao desenvolvimento de novas aplicações para sistemas de BDs, que permitem o arquivamento de dados multimédia (imagens, vídeos, sons) e geográficos (*Geographic Information Systems* – GIS – que guardam e analisam mapas, dados meteorológicos e imagens de satélite), o processamento analítico *online* (OLAP – sistemas usados para extrair e analisar informação útil proveniente de BDs muito grandes de apoio à decisão), monitorização em tempo real (de processos de fabrico, por exemplo) e técnicas de pesquisa que melhoram a qualidade da informação devolvida ao utilizador (em sistemas globalmente usados como a *World Wide Web*) [25].

Uma BD pode definir-se como uma colecção de dados relacionados (factos conhecidos que podem ser gravados, e que têm um significado implícito) que tem as seguintes características [25]:

- Representa alguns aspectos do mundo real, muitas vezes designados ‘Universo do Discurso’. Mudanças nesses aspectos traduzem-se em mudanças na BD.

- É uma colecção coerente de dados com algum significado inerente. Um conjunto de dados aleatórios nunca pode ser designado como uma BD.

- É desenhada, construída e preenchida com dados para um propósito específico. Tem um grupo de utilizadores entendidos e algumas aplicações pré-concebidas nas quais esses utilizadores estão interessados.

Uma BD tem assim uma fonte da qual os dados derivam, um grau de interacção com eventos do mundo real e uma audiência que está interessada no seu conteúdo. Pode ser de qualquer tamanho e de complexidade variada. A sua enorme quantidade de dados tem de ser organizada e gerida de forma a que os utilizadores possam pesquisá-los, recuperá-los e actualizá-los quantas vezes forem necessárias. Pode ser gerada e mantida manualmente ou pode ser computadorizada. Uma BD computadorizada pode ser criada e mantida tanto por um grupo de aplicações ou programas desenvolvidos especialmente para o efeito, ou por um sistema de gerenciamento de BDs.

O grande potencial das BDs surge do sistema de gerenciamento de uma BD (*Database Management System* – DBMS) [26]. Este é um conjunto de programas que possibilita o armazenamento e gestão de grandes quantidades de dados de uma forma eficiente e permite que estes persistam por longos períodos de tempo, de uma forma segura. Este sistema está entre os mais complexos tipos de *software* existentes. Eis as funcionalidades que ele disponibiliza ao utilizador:

1. Armazenamento persistente de grandes volumes de dados (vários *gigabytes*, ou mais), mantendo-os seguros (protegidos de acidentes ou acessos não autorizados).

2. Uma interface para que o utilizador ou uma aplicação computacional possam aceder e modificar os dados usando uma linguagem específica, designada ‘linguagem de *queries*’ (uma

query é um termo associado à linguagem da BD que pretende designar ‘uma questão acerca dos dados’) ou ‘linguagem de manipulação de dados’ (*data-manipulation language* - DML).

3. Permite aos utilizadores criar novas BDs e definir o seu esquema (estrutura lógica de dados), usando uma linguagem específica para o efeito, que se designa ‘linguagem de definição de dados’ (*data-definition language* - DDL).

4. Um ‘sistema de gestão de transacções’ que controla o acesso simultâneo por vários processos distintos (é normal agrupar uma ou mais operações da BD numa transacção, que é a unidade de trabalho de uma BD), não permitindo que as acções de um utilizador afectem outros utilizadores (impedindo assim que acessos simultâneos corrompessem os dados acidentalmente). As transacções bem implementadas devem dispor das chamadas propriedades ACID:

- Atomicidade: a execução ‘tudo-ou-nada’ de transacções; obriga a que uma acção seja executada até ao fim, ou então que o sistema volte ao estado em que se encontrava antes da acção ser iniciada.

- Consistência: todas as BDs têm restrições de consistência (p. e., um determinado valor ter de ser sempre positivo), ou obrigações nas relações entre os elementos de dados.

- Isolamento: dá a ilusão ao utilizador de que só uma transacção está a ser executada num dado momento, ou seja, há um isolamento aparente relativamente a outras transacções.

- Durabilidade: o efeito de uma transacção na BD nunca pode ser perdido, uma vez completada a transacção; há a habilidade para recuperar de falhas e erros no sistema.

O primeiro DBMS comercial surgiu no fim da década de 60. Este sistema evoluiu a partir de sistemas de ficheiros que guardavam a informação por um longo período de tempo, e também permitiam o armazenamento de grandes quantidades de informação. No entanto não garantiam que os dados não pudessem ser perdidos (era necessário realizar cópias de segurança), ou que pudessem ser acedidos de maneira eficiente (quando a sua localização num ficheiro em particular não fosse conhecida), não suportavam uma linguagem de pesquisa dos dados, o seu esquema limitava-se à criação de estruturas de directorias para os ficheiros (organização estruturada em pastas), e não preveniam a corrupção de dados causada pelo acesso de vários utilizadores em simultâneo (as mudanças realizadas por um utilizador não iriam aparecer em tempo real a um outro utilizador a aceder em simultâneo).

Nas primeiras aplicações comerciais importantes de DBMSs já eram aplicadas muitas *queries* e modificações aos dados. Algumas dessas aplicações eram: os sistemas de reservas de linhas aéreas, os sistemas bancários e os sistemas de gestão das empresas (dos dados e salários dos funcionários, das vendas e compras efectuadas, etc.).

De seguida apresentam-se as capacidades que um bom DBMS deve disponibilizar actualmente [25]:

- Controlo de redundâncias: a BD deve ter um esquema que guarde cada item de dados lógicos num único sítio na sua estrutura. Isto não permite inconsistências e preserva espaço de armazenamento. Pode também ser bom para melhorar a performance de algumas *queries*.

- Restrição de acesso a pessoas não autorizadas: alguns utilizadores podem não estar autorizados a aceder a toda a informação da BD. Devem-se então criar contas e especificar as restrições de acesso para cada conta. Um DBMS tem de permitir e controlar o acesso de vários utilizadores ao mesmo tempo ao repositório de dados comum, de modo a que a concorrência entre utilizadores não incorra em erros, ou seja, as actualizações realizadas ao mesmo tempo sejam guardadas de forma correcta.

- Fornecer armazenamento persistente para objectos do programa e estruturas de dados: a estrutura de dados do sistema tem de ser compatível com uma ou mais linguagens de programação que permitam a sua manipulação.

- Representar relações complexas entre os dados: tem que ter a capacidade de representar várias relações complexas entre os dados, e também de recuperar e actualizar dados relacionados de forma fácil e eficiente.

- Reforçar restrições de integridade: definir e reforçar restrições de integridade dos dados, por exemplo, com a especificação do tipo de dados para cada elemento.

- Fornecer um sistema de *backup* e recuperação de dados: disponibilizar ferramentas para recuperação de falhas de *hardware* e *software* do sistema computacional. Por exemplo, no caso de o sistema falhar a meio de uma operação de actualização da BD, deve ser restaurado para o estado em que se encontrava antes do início do processo, ou então gravar um resumo com o estado em que a execução parou, e reiniciá-la a partir desse ponto quando o sistema voltar a funcionar.

- Potenciar e reforçar padrões: uniformização entre diferentes utilizadores dentro de uma organização. Promover a comunicação e cooperação entre os vários departamentos e projectos de uma instituição através da definição de nomes e formatos de dados, estruturas de relatórios e terminologias comuns entre eles.

- Reduzir o tempo de desenvolvimento da BD: conceber BDs de raiz pode ocupar muito mais tempo do que simplesmente criar o ficheiro de construção de uma BD num DBMS. É também muito mais simples criar novos programas para a BD usando este sistema. Visto isto, estima-se que o tempo de desenvolvimento usando um DBMS é entre 1/6 a 1/4 do tempo que levaria um sistema tradicional.

- Flexibilidade: permite mudar a estrutura da BD conforme a mudança de requisitos para esta. Permite adicionar novos elementos à estrutura (por exemplo, adicionar um novo grupo de utilizadores, que necessite de informação que ainda não esteja armazenada na BD) pela simples execução de um ficheiro de criação adicional, isto sem afectar os elementos já existentes e os programas já implementados na aplicação.

- Disponibilidade de informação sempre actualizada: visto que o DBMS disponibiliza imediatamente as alterações feitas por um dado utilizador aos dados, para que todos os outros possam vê-las.

- Economia de escala: pois a consolidação de dados e aplicações permite reduzir a quantidade de sobreposições entre actividades de processamento de dados em diferentes projectos e departamentos dentro de uma instituição. Isto vai reduzir também, a longo prazo, os custos de operação e manutenção

Podemos ver um esquema de um DBMS completo na Fig. 6. As caixas simples representam os componentes do sistema e as caixas duplas representam as estruturas de dados na memória. As linhas sólidas indicam o controlo e fluxo de dados e as tracejadas indicam somente fluxo de dados. Os tipos de informação que os vários componentes podem precisar são: dados (o conteúdo da BD em si), meta-informação (o esquema da BD que descreve a sua estrutura e as restrições que lhe foram aplicadas), dados estatísticos (informação recolhida e armazenada pelo DBMS acerca das propriedades dos dados, tais como: os tamanhos, os valores e as várias relações dos componentes da BD) e índices (estruturas de dados que permitem o acesso eficiente à informação) [26].

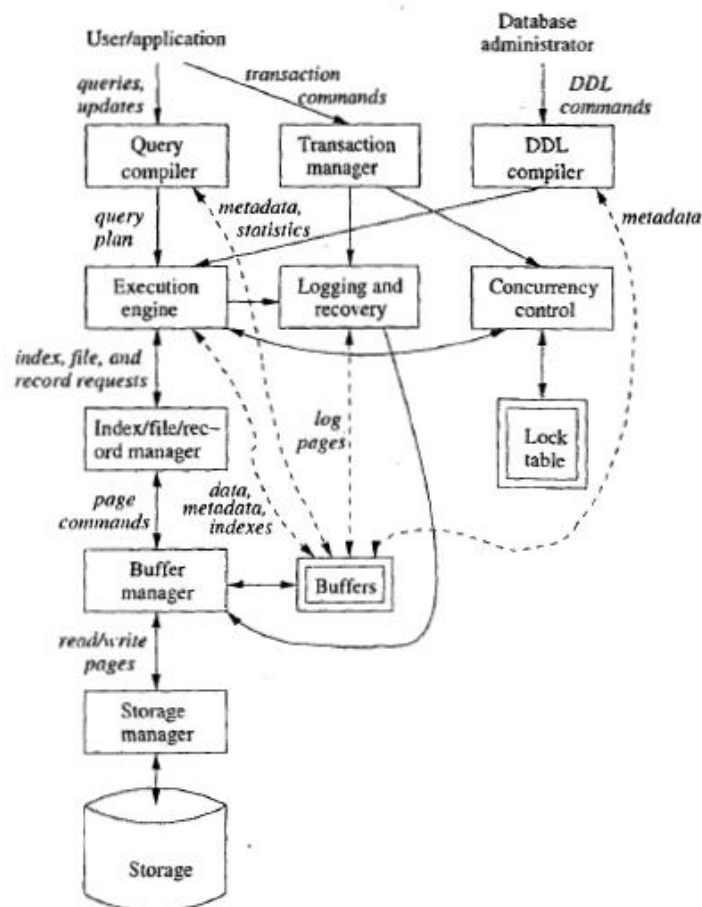


Figura 6 - Componentes de um DBMS. (Adaptado de: [26])

O administrador pode executar comandos que alterem o esquema da BD. Estes comandos DDL de alteração de esquema são analisados por um processador DDL e passados ao motor/engenharia de execução, que depois entra no gestor de índices/ficheiros/registos para alterar o esquema de informação para a BD.

A maioria das interacções com a BD segue o caminho descrito no lado esquerdo do esquema apresentado na Fig. 6. Um utilizador ou uma aplicação computacional inicia uma determinada acção que não afecta o esquema da BD, mas pode afectar o seu conteúdo (se for um comando de modificação) ou extrair dados da BD (se for uma *query*).

Diferentes DBMSs suportam diferentes linguagens de pesquisa (DML), mas há uma que está semi-implementada como padrão: a linguagem SQL (*Structured Query Language*). Uma pesquisa é um pedido de informação sobre os dados da BD, tem-se como uma pergunta estruturada. Uma linguagem de pesquisa é um conjunto de regras para a construção das pesquisas definidas anteriormente [24].

O processo de desenvolvimento de uma BD é geralmente demorado e complexo, e as pessoas tendem a confundir os seus passos principais: definição, construção e manipulação dos dados. De seguida será clarificado o trabalho levado a cabo em cada uma destas etapas. *Definir* uma BD envolve especificar tipos de dados, estruturas e restrições para que os dados sejam guardados de acordo com o problema em questão. *Construir* uma BD é o processo de guardar os dados em si no meio de armazenamento disponibilizado e controlado pelo DBMS. *Manipular* uma BD consiste em utilizar *queries* para que esta responda às nossas necessidades devolvendo dados específicos, actualizando a informação e gerando relatórios a partir dos dados [25].

Uma BD prima ao nível da abstracção pelo facto de esconder detalhes relacionados com o armazenamento dos dados que nada interessam à maioria dos utilizadores. Esta abstracção é fornecida pela utilização de um modelo de dados, uma colecção de conceitos que são usados para descrever a estrutura da BD: tipos de dados, relações, restrições que devem ser aplicadas aos dados e operações básicas que especificam as actualizações e recuperações de dados a serem executadas pela BD. Os mais relevantes são os modelos conceptuais, os modelos físicos e os esquemas de BDs. Os modelos conceptuais usam entidades, atributos e relações da seguinte forma: uma entidade representa um objecto ou conceito do mundo real, um atributo representa uma propriedade que descreve uma entidade, e uma relação entre duas ou mais entidades representa uma interacção, influência ou domínio entre elas. Os modelos físicos descrevem o modo como os dados são armazenados no computador mediante a representação de informação referente a formatos de registo, ordem de registo e formas de acesso (uma forma de acesso é uma estrutura que aumenta a eficiência de uma pesquisa realizada a um conjunto específico de registos de uma BD). O esquema da BD é uma descrição desta, que é especificada durante a concepção da BD e não muda frequentemente. O diagrama associado a este esquema mostra

unicamente alguns dos aspectos a ele associados, tais como os nomes dos tipos de registos e itens de dados, e alguns tipos de restrições.

3.3.1. Pessoas envolvidas no Processo de Concepção de uma Base de Dados

Uma BD envolve várias pessoas desde a sua concepção, passando pela sua gestão, até à sua utilização:

- Um administrador (*Database Administrator* – DBA) que é responsável por autorizar o acesso à BD, por coordenar e monitorizar o seu uso, e por adquirir *software* e *hardware* para esta conforme seja necessário. Ele responde por problemas de falhas de segurança e mau tempo de resposta do sistema.

- Os *designers* da BD devem identificar a informação que deve ser armazenada na BD e escolher as estruturas apropriadas para representar os dados. São eles que devem investigar junto dos futuros utilizadores os requisitos da BD, antes de esta ser implementada e preenchida. A BD final deve ser capaz de satisfazer os requisitos de todos os grupos de utilizadores.

- Os utilizadores finais que são as pessoas cujos empregos implicam o acesso à BD para aplicação de *queries*, actualizações e geração de relatórios. A BD existe para o seu uso.

- Os Analistas do Sistema determinam os requisitos dos utilizadores finais, e desenvolvem especificações para as transacções que vão de encontro às funcionalidades por eles especificadas.

- Os Programadores da Aplicação implementam as especificações definidas pelos analistas como programas. Testam, documentam, depuram, e fazem a manutenção destas transacções.

Numa arquitectura básica de cliente-servidor, uma aplicação de uma BD distribui-se entre dois tipos de módulos: o módulo do cliente, que funciona na plataforma de trabalho do utilizador ou num computador pessoal, e usa programas e interfaces (GUIs – *Graphical User Interfaces*) para aceder à BD remotamente; e o módulo do servidor, que armazena os dados e disponibiliza o acesso a estes, a realização de pesquisas, entre outras funções.

3.3.2. Processo de Desenvolvimento de uma Base de Dados

O processo de desenvolvimento de uma aplicação de uma BD envolve o desenho, implementação e teste desta. De seguida vão explicar-se sucintamente os passos envolvidos nesse processo.

O primeiro passo é a definição dos requisitos que a BD deve satisfazer, e a análise destes. Estes requisitos devem ser especificados da forma mais detalhada e completa possível. Em paralelo com este processo devem definir-se as operações (ou transacções) que vão ser aplicadas à BD. Neste passo é comum usarem-se digramas de fluxo de dados, diagramas de sequências e diagramas de cenários (um exemplo destes digramas, é o diagrama de casos de usos para a especificação dos requisitos do sistema).

O segundo passo é a criação de um esquema conceptual, usando um modelo de dados conceptual de alto nível. Este esquema é uma descrição concisa dos requisitos de dados que os utilizadores fizeram, dos tipos de entidades, das relações e das restrições; estes conceitos não incluem ainda detalhes de implementação. Convém averiguar-se nesta fase se os requisitos estabelecidos não geram conflitos entre si (e modificar o esquema conceptual para eliminá-los) e introduzir requisitos funcionais que não possam ter sido especificados no esquema inicial.

O passo seguinte é a implementação da BD, usando um DBMS. O modelo conceptual é transformado num modelo de implementação de dados, designando-se este passo por ‘concepção lógica’ ou ‘mapeamento do modelo de dados’ (modelo lógico).

Por fim, chega-se à fase de geração do modelo físico, durante a qual as estruturas de armazenamento, formas/caminhos de acesso e organização de ficheiros são especificadas. Em paralelo são concebidos programas para o estabelecimento das transacções da BD e das especificações a elas associadas.

O processo de concepção de uma BD começa com a análise de qual a informação que ela deve guardar e de quais as relações que se podem estabelecer entre as componentes dessa informação. Os desenvolvedores de BDs começam por desenvolver um esquema usando o modelo ‘Entidade-Relação’ (modelo ER), ou um outro modelo baseado em objectos; no entanto, a abordagem tradicional e também a mais utilizada é este modelo. Este modelo é restrito, no sentido em que só se podem representar certas estruturas. No entanto, pela sua simplicidade e utilidade, é o modelo utilizado pela maioria dos DBMSs comerciais hoje em dia [26].



Figura 7 - Processo de concepção e implementação de uma BD. (Adaptado de: [26])

Na Fig. 7 vê-se o modo como o modelo ER é utilizado na concepção de uma BD. Começa-se agrupando ideias acerca da informação que nós queremos estruturar e modelar, e depois convertemo-las num modelo ER. O modelo abstracto ER é depois convertido na

linguagem de especificação de dados de um DBMS. Como a maioria destes sistemas comerciais usa o modelo relacional, assume-se que há a conversão de um modelo ER para um esquema relacional. Converte-se assim um modelo abstracto num modelo concreto relacional, o ‘esquema relacional da BD’.

3.3.3. Modelo e Diagramas Entidade-Relacionamento

Os diagramas Entidade-Relacionamento (diagrama ER) permitem descrever a estrutura de dados adequada para a aplicação de uma BD. Traduz-se numa representação gráfica constituída por conjuntos de entidades, atributos e relações. Um modelo ER geralmente não é considerado adequado para servir de base à linguagem de manipulação dos dados. No entanto, é, muitas vezes, mais fácil desenhar BDs usando a notação ER. A concepção de um diagrama ER é um processo que vai do geral para o particular, e os seus passos principais são: identificar todas as entidades e todos os relacionamentos importantes para a situação a tratar; construir o diagrama ER em si; gerar o conjunto de tabelas a partir do diagrama ER (geralmente feito de forma automática por ferramentas de *design* concebidas para o efeito).

A unidade base de um modelo ER é a entidade, que representa uma ‘coisa’ do mundo real, que tem uma existência independente. Pode ser um objecto com existência física ou um conceito abstracto. Cada entidade colecciona atributos inter-relacionados de alguma forma, associados a um determinado tema. Estes são propriedades particulares que a descrevem. Num diagrama ER, todas as entidades têm de ter uma chave (este conceito será definido na subsecção 3.3.4) e sublinham-se os atributos que pertencem à chave primária da entidade. Cada entrada de uma entidade tem um valor para cada um dos atributos da entidade. Os valores dos atributos que descrevem cada entidade constituem a maior parte dos dados armazenados numa BD. Existem vários tipos de atributos no modelo ER: simples (ou atómicos, não são divisíveis) e compostos (podem ser divididos em parte menores, que representam mais atributos básicos com significados independentes); de valor único (um único valor para uma dada entidade) e de vários valores (podem estar limitados – necessidade de terem entre um número limite menor e um número limite maior - no número de valores permitidos para cada entidade individual); armazenados (de um modo geral todos são armazenados na BD) e derivados (quando a partir de um dado atributo podemos derivar ou inferir o valor de outro – por exemplo, determinar a idade a partir da data de nascimento); de valor nulo (quando uma entidade não tem um valor aplicável para um dado atributo, ou quando não se sabe o valor de um atributo numa entidade em particular) [25].

As relações são conexões entre duas ou mais entidades. Os relacionamentos são caracterizados por: participação (indica se o relacionamento é obrigatório ou não) e

cardinalidade (indica o grau do relacionamento - quantas entidades estão envolvidas). As relações binárias, entre duas entidades, são o tipo de relações mais comum. Podem ter diferentes multiplicidades: podem ser *one-to-one*, *many-to-one*, *one-to-many* ou *many-to-many*. Numa relação *one-to-one* uma entrada da primeira entidade pode estar associada unicamente a uma entrada da segunda entidade envolvida na relação. Numa relação *one-to-many* ou *many-to-one* cada entrada da entidade colocada do lado *many* pode estar associada no máximo a uma entrada da entidade do lado do *one*, por sua vez, uma entrada do lado *one* pode estar associada a várias entradas da entidade do lado *many*. As relações *many-to-many* permitem que cada entrada de ambos os lados esteja associada a várias entradas do lado oposto [26].

Existem também as heranças: um tipo de relação em que se associam diversas entidades de uma forma hierárquica, sendo que as entidades filhas (sub-entidades) herdam os atributos e as relações das entidades mãe (super-entidades). Uma sub-entidade representa o mesmo elemento do mundo real que a super-entidade a que pertence, ela tem de possuir os mesmos atributos que esta última mais os seus atributos específicos. Chama-se especialização ao processo de definir um conjunto de sub-entidades a partir de uma super-entidade. O conjunto de sub-entidades que forma uma especialização é definido com base em algumas características distintivas dessas entidades na super-entidade. Generalização é o processo inverso do descrito anteriormente, no sentido em que se suprimem as diferenças entre várias entidades, identificando as suas características comuns, e generalizando-as dentro de uma única super-entidade, onde as entidades originais funcionam como as suas sub-entidades. São aplicadas algumas regras de inserção e eliminação na especialização e generalização. Por exemplo, apagar uma entidade que seja uma super-entidade, implica apagar todas as sub-entidades que lhe pertencem [25].

Os princípios que devem ser tidos em conta para o *design* de um bom modelo ER são [26]:

- Fidelidade: o modelo deve ser fiel às especificações da aplicação. As entidades e os seus atributos, e as relações implementadas, devem retratar a realidade.

- Evitar Redundâncias: devemos abordar um determinado tema uma única vez. Não devemos repetir nomes de entidades, atributos ou relações, nem implementar a mesma acção (representar o mesmo facto) mas com outros nomes duas vezes.

- Simplicidade no número de elementos: evitar introduzir mais elementos no diagrama do que aqueles que sejam absolutamente necessários, para evitar dizer a mesma coisa duas vezes, ou de uma forma indirecta ou excessivamente complexa.

- Escolher as relações certas: as entidades podem ser conectadas de diversas formas, mas devem-se escolher quais as que estabelecem o mínimo de relações possíveis para satisfazer as necessidades da plataforma de maneira correcta e evitando redundâncias.

- Escolher o tipo certo de elementos: não colocar, por exemplo, uma entidade a representar aquilo que poderia ser descrito por um atributo dentro de uma outra entidade.

3.3.4. Modelo Relacional

O modelo relacional de dados representa uma BD como um conjunto de tabelas (bidimensionais, que são chamadas de ‘relações’), onde cada tabela é armazenada como um ficheiro em separado. Foi introduzido em 1970, e em 1990 os sistemas de BDs relacionais já eram a norma [25], [26].

Nas BDs relacionais, os dados são armazenados em tabelas, e estas relacionam-se entre si. O conjunto de tabelas de uma BD relacional mais os outros objectos geridos pelo DBMS (vistas, índices, procedimentos, *triggers*, sequências, etc.) é designado ‘esquema de dados relacional’ ou simplesmente ‘esquema relacional’ [24].

No Modelo de Dados Relacional todos os dados (assim como os relacionamentos existentes entre os dados) são representados por tabelas relacionadas entre si. Cada tabela tem um nome único, pelo qual é referenciada. Cada linha da tabela (registo) representa uma entidade única ou um relacionamento entre entidades. Cada tabela pode conter muitos registos (ou nenhum no caso de uma tabela vazia). Cada coluna (atributo) tem um nome e refere-se a um dado aspecto do objecto representado na tabela.

Cada atributo possui um conjunto de valores, e o conjunto de todos os valores possíveis para um dado atributo designa-se domínio do atributo. Dependendo do domínio, assim é definido o tipo de dados de um atributo (números, texto, datas...).

Uma chave primária (ou identificador primário) é o atributo, ou conjunto de atributos, que permitem identificar de forma inequívoca qualquer registo de uma tabela. Ela é unívoca (o atributo, ou atributos, que a constituem têm um valor unívoco para qualquer registo da tabela) e não nula (não pode haver registos a tabela que tenham o atributo, ou atributos, da chave primária nulo(s), isto é vazio(s)).

Uma chave estrangeira é um atributo, ou conjunto de atributos, de uma tabela que aparece como chave primária numa outra tabela, permitindo estabelecer o relacionamento entre os registos dessas duas tabelas.

Entende-se como integridade de uma BD o facto de esta conter somente dados válidos, ou seja, os dados nela armazenados estão de acordo com a realidade. Às regras que definem se os dados são válidos ou não chamam-se restrições de integridade.

Existem vários tipos de integridade: a integridade de domínio são as restrições aplicadas aos valores dos atributos de uma dada tabela para se obrigar o valor de um atributo a constar de um conjunto específico de valores ou a encontrar-se dentro de um determinado intervalo; a

integridade de entidade consiste na declaração de uma chave primária para a tabela em questão, e na consequência imediata de não poderem haver dois registos nessa tabela com a mesma chave primária; integridade referencial quando se relaciona duas tabelas por meio de uma chave estrangeira, o valor da variável que é chave estrangeira ou é nulo na tabela onde é chave estrangeira, ou terá de existir na tabela onde é chave primária; há também regras de integridade mais complexas, que não podem ser definidas na estrutura das tabelas, e que são aplicadas pela execução de programas de verificação sempre que os dados são alterados – inseridos, apagados ou actualizados – que são designadas regras de negócio [24–26].

Para a manipulação e utilização de dados relacionais é usada uma linguagem, geralmente também a linguagem SQL, que permite: a definição de dados (criação e alteração de tabelas, vistas, etc.), a manutenção da integridade dos dados (definindo as regras para tal), a pesquisa de dados nas tabelas, a manipulação de dados (inserir, alterar e apagar dados), a gestão das pessoas autorizadas a aceder à BD e a preservação da segurança dos dados (definição de privilégios, perfis, etc.) e o controlo de transacções (iniciar e terminar transacções) [24]. Na prática, o sistema deve analisar a *query* e encontrar uma maneira eficiente de lhe responder, mesmo que as relações envolvidas possam ser muito complexas [26].

3.3.5. Modelo Entidade-Atributo-Valor

Outro modelo conceptual de implementação de uma BD é o modelo Entidade-Atributo-Valor (EAV). Este modelo consiste apenas numa tabela com três atributos: na primeira coluna deve colocar-se o identificador da entidade, na segunda o nome do atributo que se quer adicionar a essa entidade, e na terceira o valor que lhe corresponde. A grande vantagem deste modelo é ter os atributos organizados por linhas (e não por colunas como acontece no modelo relacional), e portanto pode-se sempre acrescentar um novo atributo à BD simplesmente adicionando uma nova entrada à tabela (e não alterar a estrutura base da BD, como acontece no modelo relacional em que é necessário adicionar uma nova coluna para o efeito). É assim um modelo altamente flexível e expansível a qualquer momento. Por este motivo, é ideal para o armazenamento de informação médica, que se encontra em constante expansão como já foi discutido. No entanto, tem associado um pequeno inconveniente: que tipo de dados associar à coluna dos valores? Pois podem querer-se adicionar atributos que sejam portadores de diferentes tipos de dados, no entanto, estando todos os valores na terceira coluna da tabela, essa coluna tem de ter em todas as suas entradas um mesmo tipo de dados (no modelo relacional isto não acontece, porque a cada atributo corresponde uma coluna diferente; este é um dos factores que o torna ideal para a estruturação de dados). Para contornar este problema costumam-se associar referências para outras tabelas na terceira coluna, e nessas tabelas definem-se os tipos de dados

permitidos para cada atributo do modelo EAV. Outra solução é a adição de várias tabelas EAV, uma para cada tipo de dados que se pretenda introduzir [27].

Ao associar tabelas EAV a um modelo conceptual relacional constitui-se um modelo híbrido, que tenta aproveitar as vantagens dos dois modelos anteriores, e colmatar as falhas de um com as potencialidades do outro.

3.3.6. Modo de Armazenamento da Informação de uma Base de Dados

Os dados de uma BD (tipicamente em grandes quantidades, que têm que persistir durante longos períodos de tempo, e que são acedidos e processados repetidamente durante esses períodos), bem como o DBMS usado, normalmente permanecem num sistema de armazenamento secundário, que nos sistemas computacionais actuais geralmente corresponde a um disco magnético (o disco rígido do computador). O acesso a este é controlado, em primeira mão, pelo sistema operativo, que procede à gestão das entradas e saídas de informação do disco. No entanto, para que se possa realizar qualquer tipo de operação útil nos dados, estes têm que se encontrar na memória principal (*main memory*). A função do gerenciador de armazenamento (*storage manager*) é controlar a colocação de dados no disco e o seu movimento entre o disco e a memória principal. As BDs são armazenadas permanentemente num disco magnético secundário pelas seguintes razões: geralmente elas são demasiado grandes para caberem inteiramente na memória principal de um computador; as circunstâncias que causam a perda permanente de dados guardados aparecem menos frequentemente em discos de armazenamento secundário que nos primários, dado que os primeiros são de memória não volátil, enquanto que a memória principal é muitas vezes um armazenamento volátil; o custo de armazenamento por unidade geralmente é menor para unidades secundárias do que para armazenamento primário [25], [26].

As BDs são armazenadas fisicamente como ficheiros de registos (cada registo consiste num conjunto de valores de dados ou itens relacionados, onde cada valor é formado por um ou mais *bytes* e corresponde a um campo específico no registo; geralmente descrevem as entidades e os seus atributos) [25].

O tipo de dados de um campo geralmente é um dos tipos de dados padrão usados em programação: numéricos (inteiro, inteiro longo ou vírgula flutuante), caracteres (de comprimento fixo ou variável), Boolean (representando um valor ‘verdadeiro’ ou ‘falso’, por meio da associação dos valores 0 e 1 a estas descrições) e, por vezes, valores de data e tempo especialmente codificados. O número de *bytes* necessários para armazenar cada tipo de dados é fixo para um dado sistema computacional. Normalmente, um inteiro requer 4 *bytes*, um inteiro longo 8 *bytes*, um número real 4 *bytes*, um Boolean 1 *byte*, uma data 10 *bytes* (usando o formato

YYYY-MM-DD, 'ano-mês-dia'), uma *string* de comprimento fixo de k caracteres k bytes, e as *strings* de comprimento variável podem requerer tantos bytes quantos os caracteres existentes em cada campo de valores.

Nas BDs mais recentes, surgiu a necessidade de armazenar itens que consistem em objectos não estruturados de grandes dimensões (representando imagens, vídeos, gravações sonoras ou pedaços de texto), estes são chamados de BLOBs (*Binary Large Objects* – objectos binários de grandes dimensões) e são gravados separadamente do seu registo, num conjunto de blocos de memória no disco (mas é incluído um ponteiro para estes blocos no seu registo).

3.3.7. Bases de Dados de Imagem Médica

As funcionalidades mais importantes dos sistemas de bases de dados de imagem médica (*Medical Image Database Systems* - MIDS) são adquirir, armazenar e permitir a visualização das imagens digitais para que os radiologistas possam realizar um diagnóstico primário [18]. Estes sistemas são responsáveis por gerenciar as imagens desde as suas modalidades de aquisição até às estações (aos terminais) onde vão ser visualizados. Um MIDS armazena dados de imagem médica e a informação textual a eles associados, com o objectivo de suportar sistemas de apoio à decisão em ambiente de cuidados de saúde. Os dados imagiológicos são multimodais, heterogéneos e estão em constante mudança ao longo do tempo. Os pacientes podem ter imagens de diferentes partes do corpo, com diferentes modalidades usadas para as adquirir, e também repetição de exames do mesmo tipo em intervalos de tempo regulares para rastrear a progressão da doença.

Identificam-se as seguintes causas para a complexidade associada ao desenvolvimento e implementação de uma BD de imagem médica [18]:

- Grandes Conjuntos de Dados: as imagens adquiridas num exame podem variar de 2 MB em modalidades de Medicina Nuclear até cerca de 32 MB em mamogramas e radiografias digitais. No caso do ICNAS, que sendo um centro de investigação armazena não só os dados de imagem mas também os dados brutos e os sinogramas para futuro tratamento e análise, o número de bytes ocupados é substancialmente maior. Um Hospital tipicamente gera cerca de um *terabyte* de dados de imagiologia digital num ano.

- Multimodalidade: as modalidades de imagiologia médica diferenciam-se pelo tipo de informação biomédica (anatômica, bioquímica, fisiológica, geométrica e/ou espacial) que podem fornecer acerca do órgão em estudo *in vivo* (cérebro, coração, peito, fígado...). A selecção da modalidade para realizar diagnóstico depende do tipo de doença, e o radiologista tem como função sintetizar a informação resultante da imagem recolhida de forma a auxiliar a tomada de decisão (por si, ou pelo médico que pediu o exame). As características e outras

informações contidas nas imagens multimodais são diversas e inter-relacionadas de maneira complexa, o que torna difícil a sua interpretação e correlação. Mesmo dentro da mesma modalidade e para a mesma estrutura anatómica, dois conjuntos de imagens médicas podem variar muito: na espessura do corte, na orientação do conjunto de dados, na extensão da zona em análise, e na representação dos dados.

- Heterogeneidade dos dados: as imagens médicas são heterogêneas em vários aspectos: na maneira como são recolhidas, formatadas, distribuídas e visualizadas. As imagens são adquiridas em *scanners* de diferentes modalidades e em diferentes posições, representadas em formatos de dados internos que podem variar segundo a modalidade e o fabricante, e diferir na aparência, orientação, tamanho, resolução espacial, e no número de *bits* por pixel. Já a escolha de cor para a visualização das imagens posteriormente, pode servir tanto como factor de homogeneização, como de heterogeneização, na medida em que é uma escolha do utilizador (pois, de um modo geral, as imagens inicialmente são apresentadas em escala de cinzentos).

- Contextos estruturais e funcionais: numa imagem médica, a informação estrutural contribui para o conhecimento do estado da doença à medida que ela afecta a morfologia do corpo. Exemplificando, a localização de um tumor relativamente às estruturas anatómicas adjacentes (contexto espacial) tem implicações profundas no planeamento da terapia, e a monitorização do crescimento ou redução de um tumor (contexto geométrico) é um importante indicador da evolução do paciente ao longo da terapia. No entanto, o que distingue as imagens médicas de outros tipos de imagens digitais é a representação de informação funcional (bioquímica e fisiologia) acerca das partes do corpo, em adição aos seus conteúdos anatómicos e estruturas.

- Imprecisão: por causa da resolução espacial e do contraste limitados e pela presença de ruído, as imagens médicas só podem fornecer aos médicos uma representação aproximada e muitas vezes imprecisa das estruturas anatómicas e funcionalidades fisiológicas. Este fenómeno aplica-se a conjuntos de dados inteiros e a características dentro de uma mesma imagem.

- Dimensão temporal: rastrear o estado da doença e monitorizar a evolução do paciente durante o tratamento ao longo do tempo são dois procedimentos fundamentais para o diagnóstico e a decisão terapêutica, e para avaliação dos resultados a longo prazo. A habilidade para definir e detectar relações temporais em grupos de imagens de um paciente, adquiridas em momentos diferentes, juntamente com o historial médico desse paciente, é um componente essencial numa base de dados de imagem médica. À medida que a base de dados aumenta para incorporar uma grande quantidade de dados similares de um mesmo paciente, as ferramentas desenvolvidas para monitorização temporal intra-sujeito podem ser expandidas para a comparação inter-sujeito para melhorar os resultados de prognóstico e diagnóstico.

- Suporte e infraestruturas: muitas vezes existem barreiras tecnológicas e administrativas que tornam a aglomeração da informação textual e imagiológica para análise do

corpo inteiro de um dado paciente um verdadeiro desafio. A maioria das BDs e sistemas de informação são entidades *stand-alone* sob o controlo de diferentes secções e departamentos médicos. Estas entidades administrativas são relutantes à construção de pontes para outros sistemas de informação por recearem a perda de controlo sobre os seus próprios dados. É necessária uma infraestrutura integrada de comunicação que estabeleça a interface entre as diversas fontes de imagens, informação de pacientes e materiais de referência médica, de uma forma fácil e clara para o utilizador clínico.

- Segurança: com a integração de informação multimédia, surge a necessidade de salvaguardar a integridade e privacidade para os registos e imagens médicas associados a cada paciente.

- Co-Registo: o co-registo de imagens médicas é muitas vezes necessário para alinhar geometricamente 2 ou mais conjuntos de imagens 3D, podendo assim os vóxeis que representam a mesma estrutura anatómica serem sobrepostos ou comparados directamente.

3.4. Interface com o Utilizador

Com uma interface gráfica pretende-se o estabelecimento de um meio de comunicação efectivo entre o ser humano e o computador. A concepção de uma interface com o utilizador começa com a identificação dos utilizadores, das tarefas que irão desempenhar e dos requisitos associados ao ambiente em que estão inseridos. Após a identificação das tarefas, são criados diferentes cenários de utilização. Estes são posteriormente analisados para se definir o conjunto de objectos e acções (necessários à implementação das funcionalidades) e as suas representações no ecrã (estas, por sua vez, constituem a base do *layout* do ecrã: o *design* gráfico resultante da disposição dos ícones, pedaços de texto descritivo, zonas de inserção de dados e selecção de itens, formato e características da janela, menus disponibilizados, etc.) [28].

De acordo com Mandel, 1997, os princípios básicos para a criação de uma interface com o utilizador assentam em 3 regras de ouro [29]:

- Colocar o utilizador no controlo de tudo: o utilizador precisa de um sistema que “reaja” às suas necessidades e o ajude a concretizar as tarefas. Ele pretende ter o controlo da máquina, e não que a máquina o controle a ele, condicionando/limitando as suas acções. O construtor da interface deve implementar restrições que simplifiquem o modo de interacção: definindo modos de contacto que não forcem o utilizador a realizar acções desnecessárias ou indesejadas; fornecer diferentes modos de interacção, para que o utilizador possa utilizar o da sua preferência; permitir ao utilizador a interrupção da tarefa que está a desempenhar, e fazer com que as acções sejam reversíveis; desenvolver macro mecanismos que permitam a realização

de uma sequência habitual de acções de uma só vez; esconder pormenores técnicos internos do utilizador normal (não um gestor da aplicação); permitir uma manipulação intuitiva dos objectos que aparecem no ecrã (arrastar objectos, redimensioná-los, etc.)

- Reduzir a carga de memória necessária para o utilizador desempenhar as tarefas: quanto mais o utilizador tiver que memorizar, mais a interacção com o sistema será vulnerável à ocorrência de erros.

- Fazer a interface de forma consistente: a informação que aparece no ecrã deve estar organizada da mesma forma, deve manter-se o mesmo *design* ao longo de todos os menus.

Após a análise de tarefas que a plataforma deve executar e a definição dos objectos e acções em que se irão traduzir, a primeira abordagem para a concepção da interface deve incluir: o estabelecimento das finalidades e intenções de cada tarefa a ser executada (inclui a avaliação da utilidade da tarefa, da sua eficiência, da facilidade com que cada tarefa pode ser aprendida e do grau de satisfação do utilizador com o seu modo de funcionamento); o mapeamento de cada objectivo estabelecido anteriormente numa sequência de acções específicas; a especificação da sequência de acções de todas as tarefas e sub-tarefas num cenário de utilização, tal e qual como será apresentado ao utilizador na interface; a indicação do estado do sistema (o modo como a interface irá apresentar-se graficamente aquando da execução dos cenários de utilização); a definição de mecanismos de controlo (conjunto de acções e objectos que permitam ao utilizador alterar o estado do sistema); e a informação do utilizador de como os mecanismos de controlo irão afectar o estado do sistema [30].

O processo de desenho dos menus deve ser realizado iterativamente de acordo com a seguinte sequência de passos [28], [29]:

1. Análise e modelação dos utilizadores (do perfil de quem irá interagir com a aplicação – nível de educação, compreensão da tecnologia, gostos e preferências), das funcionalidades (tarefas que a interface disponibiliza para serem concretizados os objectivos da ferramenta em desenvolvimento) e dos ambientes de trabalho (onde a interface se irá situar fisicamente; a maneira como o utilizador estará posicionado ao utilizar o sistema – sentado, de pé, ou a desempenhar várias tarefas ao mesmo tempo; as necessidades do *hardware* que lhe está associado a nível de espaço, luz ou som; qualquer consideração especial que possa surgir ao nível humano devido a factores ambientais);

2. Desenho das janelas da interface;

3. Construção (implementação das ferramentas de funcionamento em *background*) da interface – geralmente começa com a criação de um protótipo que permite a avaliação dos cenários de uso;

4. Validação da interface – uma vez criado um protótipo da interface com o utilizador que esteja funcional, tem que se averiguar se ele vai de encontro às necessidades do utilizador: avalia-se a capacidade ou não da interface implementar correctamente todas as tarefas

selecionadas pelo utilizador, suportar possíveis variações ao longo da execução dessas tarefas, alcançar todos os requisitos usuais que um utilizador espera de uma ferramenta deste tipo, o grau de facilidade de utilização, aprendizagem e compreensão que lhe está associado, e a aceitação dos utilizadores da interface como uma ferramenta útil no seu trabalho.

Cada um destes passos é executado mais do que uma vez (ver Fig. 8), no sentido de se implementarem requisitos adicionais necessários a tornar a execução da interface livre de erros e fácil de compreender para o utilizador.

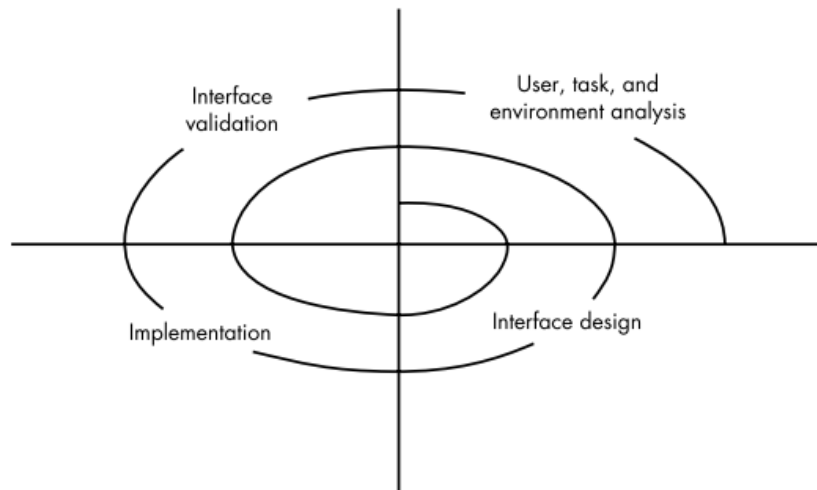


Figura 8 - O processo de *design* da interface com o utilizador. (Adaptado de: [28])

Pressman, 2001, defende também que o desenvolvedor da interface deve também ter em conta: como a interface irá ser implementada, o ambiente que será utilizado (a tecnologia/instrumento de visualização, o sistema operativo utilizado e as ferramentas/*software* de desenvolvimento); o tempo de resposta do sistema; a existência de ferramentas de ajuda para o utilizador, de um sistema que forneça informação acerca de erros que ocorram (mensagens que expliquem correctamente o que correu mal, e que deem sugestões ao utilizador de como resolver o problema), e da rotulagem de comandos (nota descritiva da acção desempenhada por um determinado comando). Estes aspectos, quando tomados em conta desde o início do processo de *design*, são fáceis de implementar e implicam custos reduzidos, mas geralmente o desenvolvedor da plataforma só se preocupa com eles num estágio mais avançado do projecto, levando a iterações desnecessárias, atrasos na entrega e frustração dos clientes.

É deveras importante que a interface com o utilizador, o único meio que este tem de aceder ao conteúdo do sistema (às suas ferramentas), se encontre bem desenhada, implementada, seja de fácil percepção e manipulação, porque é ela que vai construir a opinião do utilizador acerca da potencialidade da estrutura que está na sua raiz. Assim sendo, se esta for de má qualidade vai automaticamente denegrir a avaliação que é feita acerca do restante conteúdo da plataforma.

Capítulo 4

Equipamento Utilizado e Dados Recolhidos

4. Equipamento Utilizado e Dados Recolhidos

Recolhidos

As imagens e a meta-informação a elas associada a serem inseridas na BD são provenientes do *scanner* de PET/CT Gemini GXL da Philips® que existe no ICNAS (Fig. 9). Serão inseridos todos os ficheiros ‘.img’ e ‘.syn’, não comprimidos, armazenados no dispositivo de memória designado, genericamente, como ‘DBDISK’ que se encontra em rede no sistema informático do centro. Pretende-se, no entanto, que futuramente se possa expandir esta BD a dados recolhidos por outros tipos de equipamento de imagiologia médica existentes nesta instituição.



Figura 9 - Fotografia do *scanner* Gemini GXL da Philips® usado para adquirir imagens de PET/CT no ICNAS.

Para além dos dados imagiológicos propriamente ditos, este equipamento fornece vários tipos de ficheiros de acordo com os aspectos técnicos do exame (ficheiros de: registo de acções, calibração, resumos, etc.). Eles serão analisados mais à frente numa sub-secção deste capítulo.

Os ficheiros utilizados para o preenchimento da BD foram os correspondentes às imagens reconstruídas de PET e TAC (as já referidas extensões ‘.img’ e ‘.syn’, respectivamente). Estes ficheiros armazenam não só as imagens adquiridas no exame médico, mas também meta-informação relacionada com as condições sob as quais o exame foi executado, o paciente que o realizou, o médico que o requereu, configurações/características definidas para a aquisição, contagens estatísticas e outras informações específicas para cada um dos muitos cortes que o compõem. Todos estes dados são armazenados ao nível dos cabeçalhos do ficheiro. A estrutura base deste tipo de ficheiros é apresentada na Fig. 10.

<i>512 bytes</i>	Cabeçalho Principal
<i>512 bytes</i>	Primeira Estrutura de Directorias
<i>512 bytes</i>	Cabeçalho Alargado
<i>512 bytes</i>	Sub-Cabeçalho
<i>N×512 bytes</i>	Dados do Corte
<i>512 bytes</i>	Sub-Cabeçalho
<i>N×512 bytes</i>	Dados do Corte
<i>512 bytes</i>	Sub-Cabeçalho
<i>N×512 bytes</i>	Dados do Corte

Figura 10 - Imagem para ajudar o leitor a perceber a estrutura genérica do ficheiro de um estudo (volume de imagem reconstruído) no formato *imago*.

Com base na informação retirada de “GEMINI - Luminary Site Support Package,” 2005, a estrutura dos ficheiros em análise consiste num cabeçalho principal, um determinado número de registos de directorias (*directory records*) organizados numa lista ligada sequencialmente, e por registos de dados representando matrizes. Dados adicionais são guardados como um cabeçalho alargado que por sua vez é guardado no formato de uma entrada de directoria (*directory entry*). Devido ao facto do cabeçalho principal ser o primeiro conjunto de dados escrito no ficheiro, o cabeçalho alargado é criado como o primeiro registo de directorias no ficheiro [31].

A estrutura de dados fundamental neste formato é a matriz: uma concatenação de um sub-cabeçalho de *512 bytes* com um bloco de registos de dados (*data records* que são escritos como variáveis de *2 bytes – signed short*) que contém uma secção (um corte) de uma imagem ou sinograma. Esta matriz é indexada pelos números de corte, inclinação e *frame* e pode ocorrer em qualquer posição do ficheiro a partir dos primeiros *1024 bytes* (que são ocupados pela informação das duas únicas localizações pré-determinadas num ficheiro de imagem: a primeira pertence ao cabeçalho principal original que ocupa os primeiros *512 bytes* e a segunda pertence à primeira estrutura de directorias que é armazenada nos *512 bytes* seguintes).

São mantidas estruturas de directorias (*directory structures*) para indicar quais os registos de dados (registos de dados – *data records*) do ficheiro que são usados para armazenar uma determinada matriz (ou seja, a informação de um determinado corte). Isto significa que a organização dos blocos de dados pode variar. No entanto, assume-se que os dados pertencentes a uma mesma matriz são escritos de forma contígua. Informação mais detalhada sobre as estruturas utilizadas para organizar este tipo de ficheiros pode ser encontrada no Anexo A deste documento.

As imagens são tipicamente, mas não só, matrizes de dimensões 144×144 (para um campo de visão – FOV² – de 576) ou 128×128 (para um FOV de 256) pixéis, onde as colunas da matriz correspondem às linhas da imagem e o primeiro elemento da matriz corresponde ao pixel do canto superior esquerdo na imagem.

As imagens e os sinogramas podem ser comprimidos (manualmente ou automaticamente) no *software* PETView[®]. Neste caso, os dados são comprimidos antes de serem escritos no ficheiro. O cabeçalho principal, os sub-cabeçalhos e as directorias nunca são comprimidos. O algoritmo de compressão utilizado é o método de compressão usado por omissão pela biblioteca *zlib*.

Todos os dados são guardados usando a ordenação de *bytes* ‘Big Endian’.

Existem 13 formatos diferentes de ficheiros de imagem ou sinograma (tabela 10), correspondendo estas diferentes versões a uma diferente organização ou constituição do cabeçalho principal. Algumas das versões podem também estar logicamente associadas a uma determinada versão dos sub-cabeçalhos (tabela 11). As funções que leem o cabeçalho principal e os sub-cabeçalhos actualizam automaticamente os referidos cabeçalhos para a última versão.

Tabela 10 - Lista das versões de cabeçalhos principais. (Adaptado de: [31])

Versão	Razão
-1	Formato inicial.
0	Revisão inicial com o tipo de geometria do <i>scanner</i> e as revisões de <i>hardware</i> .
1	Nenhuma razão conhecida.
2	Adição da escala SUV ao sub-cabeçalho.
3	Nenhuma razão conhecida.
4	Adicionou-se o diâmetro do objecto, os <i>singles</i> como um tipo de aquisição, e os isótopos utilizados.
5	Adicionou-se a duração do <i>scan</i> ao sub-cabeçalho, a opção de aquisição de <i>singles</i> ao cabeçalho principal e o tipo de aquisição para o <i>scan</i> de “contaminação de emissão”.
6	Nenhuma razão conhecida.
7	Substituição da pré-escala inteira de <i>singles</i> por valores de vírgula flutuante.
8	Foi adicionada informação da geometria do <i>scanner</i> necessária para a reconstrução.
9	Foi adicionada informação do tipo de atraso (<i>delay</i>) usado na subtracção <i>online</i> de <i>randoms</i> e no <i>rebinning</i> das coincidências aleatórias.
10	Adição do cabeçalho alargado.
11	Mudança da estrutura ‘reconinfo’ do sub-cabeçalho para o cabeçalho principal.

² FOV – do inglês *Field of View*; refere-se ao campo de visão; a área ou ângulo sólido que pode ser visto pelos detectores do aparelho de PET/CT.

Tabela 11 - Lista das versões dos sub-cabeçalhos. (Adaptado de: [31])

Versão	Razão
0	Anterior à existência de qualquer campo de versão.
1	Primeiro campo de versão, ‘reconinfo’ movido para o cabeçalho principal, adição dos campos DICOM <i>gated</i> (aquisição sincronizada com algum tipo de sinal), cardíaco e de reconstrução.

Para possibilitar o acesso à meta-informação dos exames, a Philips® fornece uma série de ferramentas (designadas ‘Luminary Tools’) que permitem a extracção destes dados de acordo com a estrutura sequencial de campos, compostos por um determinado número de *bytes*, que se encontram por uma ordem definida (mapa de *bytes*) nos cabeçalhos de cada ficheiro. Estas rotinas encontram-se desenvolvidas em IDL®, MATLAB® e num *plug-in* baseado em linguagem Java designado ImageJ Pet Reader®, e serão abordadas no capítulo ‘Análise das Variáveis Existentes’.

A meta-informação foi assim recolhida de todos os ficheiros de imagem reconstruída dos estudos realizados até então, de acordo com o mapa de *bytes* dos cabeçalhos, para ser depois introduzida nos campos correspondentes da BD gerada. Note-se que, a cada campo do cabeçalho principal (um em cada ficheiro) e dos sub-cabeçalhos (um por cada corte, vários em cada ficheiro – tantos quantos o número de cortes desse exame) se fará depois corresponder uma variável na BD. Na Fig. 11 apresenta-se, a título de exemplo, parte da estrutura do mapa de *bytes* do cabeçalho principal de um ficheiro *imagio*.

	0	1	2	3	4	5	6	7	8	9	a	b	c	d	e	f
000x	01	00	00	00	00	16		file_fmt	scan_geom	hw_config	00	40	00	00	5B	
001x																
002x																
003x	00	07	FF	FF		filtyp									00	20
004x	FF	FF		daycre	mocre	yrcre		hrcr		micre		secre				duratn
005x		shdtyp		sngpscl		singopt		<===== pscale =====>		<===== detectorRadius =====>						virtualXtal
006x		phiMashing		polygonsides		xtalsPerSide		nXtalRows		<=== crystalThickness ===>		<===== xXtalPitch =====>				
007x		<===== zXtalPitch =====>				<===== axialFov =====>		rphiType		sliceType		delayType		00		1B
008x	FF	FF		pattyp		acntyp		numray		numang		slcthk		isotop		<== slope ==>
009x	<=====>			<===== intcpt =====>				injtim		<=== polygonVertAt0deg ===>				nslice		nframe
00Ax		bthday		bthmo		bthyr		<=====ssn[10]=====>								
00Bx		ntilt		petnum		00	21	FF	FF		<=====activity=====>		<=====weight=====>			
00Cx		hrinj		mininj		<===== srcRadius =====>		<===== srcZpos =====>				<===== halfLife =====>				
00Dx		<===== concfac =====>				<===== concfac_bgsb =====>		<===== dmax =====>				<===== dline =====>				
00Ex		<===== angmax =====>				<===== x0 =====>		<===== y0 =====>				<===== z0 =====>				
00Fx		<===== nevent =====>				<===== nsino =====>		eglob_low		eglob_up		eloc_low		eloc_up		

Figura 11 - Parte inicial do mapa de *bytes* do cabeçalho principal de um ficheiro *imagio*. (Adaptado de: [31])

4.1. Descrição dos Tipos de Ficheiros Gerados em cada Estudo

Segundo o “GEMINI GXL: Instructions for Use - System Operation,” 2007, é o programa de reconstrução que determina o tipo de estudo que é criado a partir da informação guardada nos cabeçalhos do ficheiro em análise [32]. Os tipos de estudos possíveis são:

- Emissão estática: uma única *frame* contendo um ou mais cortes.
- Emissão dinâmica: mais do que uma *frame*, correspondendo cada *frame* a um período diferente no tempo, mas mantendo sempre a mesma posição da mesa de exame.
- Corpo inteiro: várias *frames*, mas neste caso cada *frame* corresponde a uma posição diferente da mesa.
- Emissão *gated* (sincronizada com algum tipo de sinal – cardíaco, respiratório...): uma ou mais *frames*, com cada *frame* a corresponder a um intervalo de uma função periódica.

Ao ser realizado um exame de PET/CT a um determinado paciente, os ficheiros gerados durante e após a aquisição (reconstrução) são armazenados numa directoria definida para tal, numa pasta cujo nome é constituído pela letra ‘p’ (inicial da palavra ‘paciente’) seguida do número de ordem em que esse exame foi realizado na Instituição. Caso seja realizado mais que um estudo durante esse exame a esse mesmo paciente, cada um deles é armazenado em subdirectorias designadas ‘sX’ (‘s’ devido à palavra estudo - *study* do inglês - e ‘X’ representa o número de ordem do estudo recolhido durante esse exame, iniciando-se a contagem em 0, ou seja, para o primeiro estudo ‘s0’, para o segundo ‘s1’, e assim sucessivamente; ‘s1’ geralmente trata-se de um segundo estudo realizado pouco tempo depois do primeiro). Em cada estudo realizado são gerados pelo menos dois ficheiros de PET e um de TAC.

Inicialmente, durante a aquisição, são gerados os dados brutos, que são armazenados em ficheiros *scan* (extensão ‘.scn’). Não interessa perder esta informação, visto que a partir destes dados se pode voltar a reconstruir e obter todos os outros ficheiros. Na tabela 12 apresentam-se os tipos de ficheiros *scan* que podem ser gerados numa aquisição.

Tabela 12 - Tipo de ficheiros scan que podem ser gerados durante uma aquisição. (Adaptado de: [32])

Tipo de Ficheiro	Formato do Nome do Ficheiro
Sinograma de Emissão	pXsY_wb_dyn.scn (para <i>scans</i> de corpo inteiro) pXsY_dyn.scn (para <i>scans</i> dinâmicos) pXsY_wb.lst (para corpo inteiro em formato ‘ <i>list mode</i> ’) pXsY_resp.lst (para aquisições respiratórias em formato ‘ <i>list mode</i> ’)
<i>Delays</i> (estimativa de coincidências aleatórias em PET)	pXsY_wb_de.scn (para <i>scans</i> de corpo inteiro) pXsY_dyn_de.scn (para <i>scans</i> dinâmicos)
Sinograma de Emissão <i>Gated</i>	pXsY_gan.scn (normal para a selecção de um só batimento) pXsY_gas.scn e pXsY_gal.scn (curto e longo, respectivamente, para a selecção de um intervalo ou de uma percentagem de batimentos)

Adicionalmente, são criados três ficheiros ‘.log’ onde são guardados os *inputs* no ‘CT Host’ (computador da sala de controlo onde se controla a aquisição e onde ficam inicialmente

armazenados os dados de CT; depois são transferidos para outro computador, a "estação PET") e toda a actividade do sistema durante o *scan*.

Primeiro são criados, a cada aquisição, dois ficheiros '*scn*', um referente a eventos *prompt*³ e outro referente a eventos *delayed*⁴. O nome desses ficheiros é constituído primeiramente pelo número identificador do paciente a que pertence (a letra 'p' seguida desse número: exemplificando, para o paciente número 30 seria 'p30') seguido do número do estudo que está a ser realizado (por exemplo, 's0'). Após um *underscore* ('_') seguem-se especificações acerca das condições de realização do exame. O primeiro ficheiro gerado contém a sigla 'wb' - *whole body* - para designar um exame de corpo inteiro, e nele são registados os chamados eventos *prompts* (segundo a linha de exemplos que têm vindo a ser dados, o nome completo do ficheiro seria 'p30s0_wb.scn'). O segundo ficheiro contém também a sigla 'de' - *delayed events* - aqui é armazenada uma estimativa das coincidências aleatórias, a fim de depois serem subtraídas às restantes.

Para além destes são gerados ficheiros auxiliares à análise dos dados recolhidos. A informação relativa a alguns deles é apresentada no Anexo B desta tese de mestrado.

Em cada aquisição são gerados vários ficheiros no formato DICOM (extensão '.dcm'), geralmente com imagens de TAC. No entanto, ao nível de transferência de dados para outras estações, tanto os ficheiros PET como TAC são movidos neste formato. Este tipo de ficheiros, por enquanto, está somente a ser utilizado para visualização de imagens, como será descrito no Capítulo 8. Mas nos cabeçalhos dos ficheiros *imago* existem campos dedicados ao armazenamento de parâmetros característicos deste *standard*, e estes encontram-se armazenados nos campos da BD criada.

Optou-se então pelo uso dos ficheiros *imago* devido ao facto da BD ser primariamente para investigação, estudos estatísticos e processamento de imagem, e este ser o formato que melhor garante o acesso a todos os dados (imagem, dados brutos e meta-informação).

4.1.1. Ficheiros PET

Numa aquisição PET são tipicamente criados vários tipos de ficheiros *imago* (extensão '.img'). São estes ficheiros que têm as imagens propriamente ditas, e, por defeito, são criados mais ficheiros deste tipo do que aqueles que são necessários para a interpretação dos exames. Descrevendo os mais relevantes:

³ Eventos *prompt* são as coincidências detectadas (incluem coincidências verdadeiras, dispersas e aleatórias).

⁴ Eventos *delayed* são a estimativa das coincidências aleatórias em PET.

- Ficheiros com a expressão ‘ctac’ – *ct attenuation correction* – armazenam os dados reconstruídos com maior qualidade. São os dados realmente vistos pelo médico. Exemplificando ao nível do nome do ficheiro: ‘p30s0_wb_ctac.img’.

- Ficheiros com a expressão ‘nac’ – *not attenuation correction* – dados sem atenuação de radiação. Os médicos usam para tirar algumas dúvidas (p.e., no caso da existência de artefactos devido aos movimentos dos doentes). Seguindo os exemplos anteriores, um possível nome para este tipo de ficheiros seria: ‘p30s0_wb_nac.img’.

Note-se que, com outros protocolos de aquisição, podem ser criados outros ficheiros *imago* com nomes diferentes. No entanto, os descritos, são dois dos usados nos protocolos de oncologia implementados no ICNAS.

4.1.2. Ficheiros TAC

É gerado um só ficheiro TAC (extensão ‘.syn’), e nele são armazenados os dados correspondentes à imagem de CT na técnica de PET/CT. O nome do ficheiro é composto pelo identificador do paciente e do estudo já descritos para os ficheiros de PET, normalmente seguidos da sigla ‘ldCT’, que se refere a ‘*low dose CT*’, em português CT de baixa dosagem. No entanto existem mais dois tipos de ficheiros CT, que se encontram descritos na tabela 13.

Tabela 13 - Tipos de Ficheiros TAC que podem ser gerados numa aquisição. (Adaptado de: [32])

Tipo de Ficheiro	Formato do Nome do Ficheiro
CT de Baixa Dosagem (CTAC com FOV de 600 mm)	pXsY_ldCT<n>.syn
CT (CT com FOV menor que 500 mm)	CT<n>.syn
CT Respiratória de Baixa Dosagem	pXsY_resp_gatedCT<n>.syn

Capítulo 5

Estudo das Variáveis Existentes

5. Estudo das Variáveis Existentes

Começou-se por utilizar as rotinas em IDL[®] disponibilizadas pela Philips[®] para proceder a uma primeira extracção e análise dos dados, ou seja, das variáveis existentes no cabeçalho principal e sub-cabeçalhos de cada ficheiro. O nome, tipo e conteúdo de cada variável foram analisados a partir da informação fornecida em “GEMINI - Luminary Site Support Package,” 2005, e são também apresentados em anexo a este documento (ver Anexo C).

O IDL[®] (*Interactive Data Language*) é uma linguagem de programação usada para análise de dados, principalmente quando se tratam de grandes quantidades destes. As suas maiores potencialidades encontram-se ao nível do processamento de imagem e da concepção de interfaces gráficas, sendo muito utilizado na área científica, principalmente no que se refere à Astronomia e à Imagiologia Médica. A sua sintaxe tem como raiz muitas expressões de Fortran e C.

Seguidamente são apresentadas as rotinas que foram utilizadas nesta primeira fase e uma breve descrição do seu modo de funcionamento.

Estas rotinas têm por objectivo auxiliar os desenvolvedores de *software* de análise de dados PET que usam o IDL[®] como ferramenta de trabalho, fornecendo-lhes uma interface para o formato de dados PET da Philips[®]. No pacote de ferramentas ‘Luminary Tools’ existem rotinas com as seguintes funções: ler ficheiros de imagem, de sinogramas e em *List Mode*.

Para o nosso estudo das variáveis existentes apenas se utilizou a rotina ‘allo_volume_read.pro’ que permite a leitura de um ficheiro de imagem ou sinograma para uma matriz multidimensional em IDL[®]. Esta rotina não lê dados comprimidos. Lê sim o cabeçalho principal, os sub-cabeçalhos e as matrizes a eles associadas com os cortes de cada exame. Este programa chama internamente dois outros, designadamente o ‘idl_mhdr_struct.pro’ e o ‘idl_shdr_struct.pro’, que definem as estruturas do cabeçalho principal e dos sub-cabeçalhos, respectivamente [31]. A adaptação deste programa utilizada nesta análise inicial, concebida pelo Eng.º Nuno Ferreira, utiliza também a rotina ‘is_compressed.pro’ para averiguar se o programa que está a ser lido se encontra comprimido ou não.

Com esta análise inicial concluiu-se que nem todos os campos constantes dos ficheiros clínicos foram preenchidos aquando da criação do ficheiro. Muitos deles tratam-se simplesmente de campos vazios, que assim são mantidos para prevenir a necessidade de introduzir novas informações no futuro, e outros não têm simplesmente interesse para o estudo que se pretende desenvolver. Detectarem-se também algumas discrepâncias ao nível das variáveis extraídas relativamente à documentação que as descreve (o manual ‘Luminary Tools’). Para analisar especificamente as anomalias detectadas deve consultar-se o Anexo D deste trabalho.

A ideia inicial seria desenvolver a interface para a BD criada para o armazenamento da meta-informação retirada dos cabeçalhos de cada ficheiro em IDL[®], até porque já havia trabalho desenvolvido neste sentido nesta linguagem de programação. No entanto, não se conseguiu estabelecer o acesso à BD construída através do IDL[®], por questões técnicas e administrativas (não se conseguiu a licença de acesso ao IDL Dataminer[®]), e então optou-se por utilizar o MATLAB[®] como linguagem de programação base para este projecto, uma vez que: também existem rotinas de extracção da meta-informação desenvolvidas pela Philips[®] nesta linguagem de programação (e também de escrita de ficheiros deste tipo, o que pode ser uma mais valia no desenvolvimento de novas ferramentas de trabalho com este tipo de ficheiros e investigação no futuro); esta é uma poderosa linguagem de programação ao nível matemático, já bem conhecida, de fácil manipulação e a sua mais valia está nas *Toolboxes* específicas para desempenhar determinado tipo de tarefas (que vão sendo desenvolvidas e melhoradas, aumentando a abrangência das funcionalidades desta linguagem a cada dia), nomeadamente para o processamento de imagem que será em análise última uma das tarefas mais importantes a ser desempenhada pela interface em desenvolvimento; a ligação à BD através do MATLAB[®] foi estabelecida com sucesso, sem necessidade de incorrer em custos adicionais com a aquisição da *toolbox* ‘database’ desenvolvida para o efeito.

As rotinas em linguagem MATLAB[®] existentes no ‘Luminary[™] Site Support Package’ que foram usadas no desenvolvimento deste projecto encontram-se descritas detalhadamente no Anexo E deste documento.

Por fim, só referir o terceiro método de leitura que é disponibilizado, mas que não foi utilizado neste projecto. O ‘PET_Reader.jar’ é um *plug-in* que permite o acesso ao ‘ImageJ’, uma aplicação de manipulação de imagem baseada em linguagem Java que lê ficheiros de imagem PET e sinogramas formatados pela Philips[®]. Este *plug-in* reconhece a directoria e a estrutura do cabeçalho destes ficheiros e escala os dados para permitir análises quantitativas. Ele também trabalha com ficheiros de dados de PET que estejam comprimidos: o *plug-in* lê o subconjunto de parâmetros do cabeçalho necessários para a visualização dos dados [31].

Capítulo 6

Organização da Informação

6. Organização da Informação

Para a organização da informação na base de dados, foi seguida a estrutura/organização temática sugerida no manual ‘Luminary Tools’ da Philips[®], para se proceder à separação do tipo de informação que pode ser recolhida dos cabeçalhos dos ficheiros. Assim, os temas definidos correspondem aos nomes das 12 tabelas de variáveis apresentadas no referido manual, nomeadamente: ‘Informação do Paciente’ (*Patient Information*), ‘Informação do Estudo’ (*Study Information*), ‘Parâmetros de Configuração do Scanner’ (*Scanner Setup Parameters*), ‘Configuração do Scanner’ (*Scanner Setup*), ‘Estatísticas de Contagem’ (*Counting Statistics*), ‘Formato do Ficheiro’ (*File Format*), ‘Parâmetros de Alinhamento PET/CT’ (*PET/CT Alignment Parameters*), ‘Outros Valores do Cabeçalho Principal’ (*Other Mainheader Values*), ‘Parâmetros Globais dos Sub-Cabeçalhos’ (*Sub-Headers Global Parameters*), ‘Parâmetros da Frame’ (*Frame Parameters*), ‘Parâmetros do Corte’ (*Slice Parameters*) e ‘Parâmetros de Reconstrução’ (*Reconstruction Parameters*).

Tendo a informação extraída dos ficheiros de dados, é necessário armazená-la de forma estruturada mediante os temas definidos.

Assim sendo, procedeu-se à construção de uma BD com os campos/atributos necessários para a inserção de todas as variáveis extraídas dos cabeçalhos dos ficheiros de PET/CT (mas não a imagem em si, somente a directoria do ficheiro onde esta se encontra é armazenada num campo criado para o efeito na BD). Como esta informação se encontra em todos os ficheiros (é comum, trata-se de um conjunto de variáveis fixo, que existe sempre de forma sistemática), poderia ser introduzida numa estrutura fixa e imutável, pelo que, inicialmente, se optou por criar o tradicional Modelo Conceptual Relacional para conectar e guardar esta informação.

Mediante o tipo de informação disponibilizada, decidiu-se fazer corresponder uma entidade na BD a cada um dos temas definidos anteriormente, inserindo nessa entidade todas as variáveis extraídas correspondentes ao tema que lhes precedia. As questões que se colocavam neste ponto eram então se as entidades que continham a informação extraída seriam suficientes para concretizar as ambições deste projecto e como se relacionariam.

No primeiro caso a resposta era obviamente negativa, uma vez que nos faltava um meio de armazenar os dados resultantes do posterior processamento dos dados, as rotinas utilizadas para os processar e os dados dos utilizadores autorizados da BD. Por uma questão de comodidade para os utilizadores da BD e por questões práticas de implementação, verificou-se posteriormente que também seria necessário uma entidade que armazenasse os significados de todas as variáveis presentes na BD (este assunto será discutido de maneira mais aprofundada nos Capítulos 7 e 8 deste documento). Desta maneira foram criadas as entidades

‘Processamento’ (*Processing*), ‘Rotinas’ (*Routines*), ‘Acesso à Base de Dados’ (*DB Access*) e ‘Descrição de Variáveis’ (*Variables Description*) para esses efeitos.

Relativamente à segunda questão, identificaram-se as seguintes possibilidades de relacionamento (os tipos de relações estabelecidas serão clarificadas no Capítulo 7 desta tese):

1. Primeiro, dever-se-iam separar os dados relativos aos pacientes da restante meta-informação recolhida de cada exame, uma vez que a um mesmo paciente poderiam pertencer vários exames. Assim, estabelecer-se-ia uma relação entre a entidade ‘Informação do Paciente’ e os restantes dados do estudo.

2. Quanto a todos os outros dados recolhidos relativos ao estudo, poder-se-ia criar uma herança, de forma a identificá-los como pertencentes a um mesmo exame (mediante o identificador primário do estudo, que seria assim herdado por todas as entidades que lhe sucedessem na hierarquia). O mesmo para as entidades relativas aos dados de cada corte dentro de um mesmo estudo, que deveriam herdar o identificador primário da secção a que pertencem, ficando assim com a chave combinada do estudo e do corte que representam como identificador unívoco. Assim a primeira herança estabelecida seria entre a entidade ‘Informação do Estudo’ e todas as outras com dados relativos a esse estudo (‘Configuração do *Scanner*’, ‘Estatísticas de Contagem’, ‘Formato do Ficheiro’, ‘Parâmetros de Alinhamento PET/CT’, ‘Outros Valores do Cabeçalho Principal’ e ‘Parâmetros Globais dos Sub-Cabeçalhos’), e estabelecer-se-ia uma sub-herança a partir dos dados gerais dos cortes presentes nesta, a entidade ‘Parâmetros Globais dos Sub-Cabeçalhos’, e as restantes tabelas com dados específicos para cada corte (‘Parâmetros da *Frame*’, ‘Parâmetros do Corte’ e ‘Parâmetros de Reconstrução’).

3. Por fim, é necessário relacionar um determinado exame com todos os processamentos que lhe possam vir a ser aplicados, devendo-se assim criar uma relação entre a ‘Informação do Estudo’ e o ‘Processamento’.

4. Quanto às outras tabelas referidas até ao momento (‘Rotinas’, ‘Acesso à Base de Dados’ e ‘Descrição das Variáveis’) armazenam informação associada à gestão e implementação da ferramenta em desenvolvimento, e portanto surgem como entidades independentes, sem relação com qualquer outra.

No entanto, este modelo provou ser insuficiente para disponibilizar a fácil expansão da BD que era desejada: de forma a ser possível acomodar os novos dados gerados a partir de algoritmos de processamento de imagem (entre outros), e devido às características da informação médica, cujo volume e dinamismo está em permanente evolução. Assim, decidiram criar-se tabelas EAV associadas a cada uma das entidades ‘Informação do Paciente’, ‘Informação do Estudo’, ‘Processamento’ e ‘Parâmetros Globais dos Sub-Cabeçalhos’, de forma a poder associar novas variáveis a cada uma delas. Optou-se também por associar uma tabela EAV ao ‘Acesso à Base de Dados’ para permitir a inserção de novos dados associados aos utilizadores da BD (que a Instituição de operação da BD pode vir a necessitar e/ou exigir no

futuro). Juntaram-se desta forma as tabelas referidas ao modelo relacional descrito anteriormente, constituindo um Modelo Conceptual Híbrido, que se afigura como a melhor solução para o problema em estudo.

Na Fig. 12 apresenta-se um diagrama que representa sucintamente o modelo organizacional concebido.

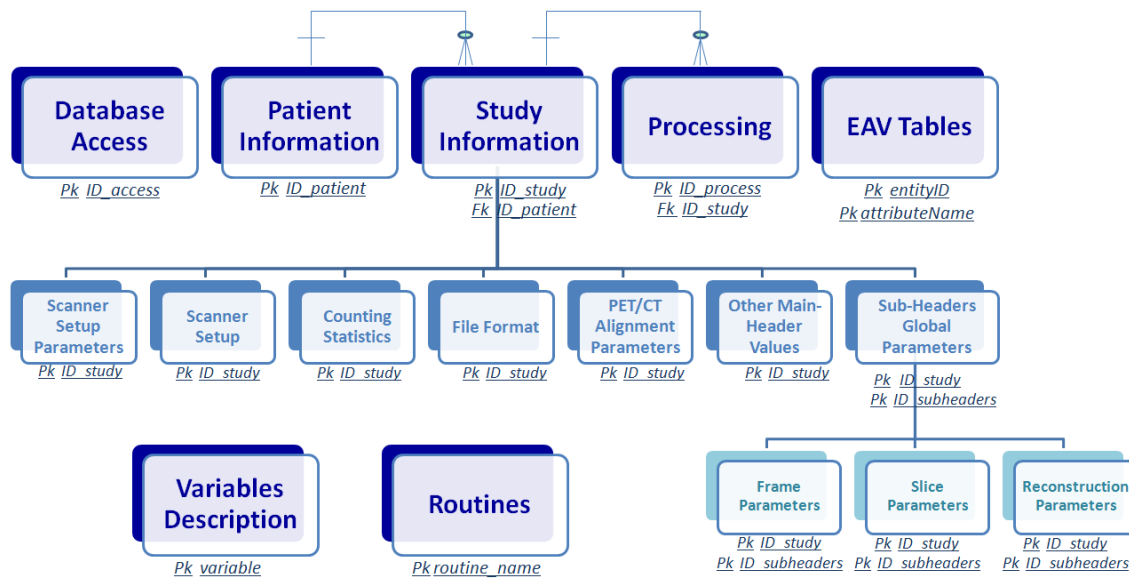


Figura 12 - Diagrama abreviado do Modelo Conceptual Híbrido concebido para a BD. Mostra apenas os nomes das entidades (no interior das caixas, exceptuando-se a designada 'EAV Tables') e as chaves primárias e estrangeiras (fora das caixas, a itálico), bem como uma representação apenas de todas as tabelas EAV que serão necessárias.

Foram mantidas as designações em inglês para a construção da BD, uma vez que a plataforma a ser desenvolvida irá funcionar numa entidade de pesquisa que actua a nível internacional (toda a ferramenta - BD, rotinas de carregamento e interface - foi concebida em inglês).

As imagens em si não são armazenadas na BD como dados binários porque este método não seria eficiente. Como os ficheiros já se encontram armazenados num dispositivo de memória à parte, é mais cómodo lê-las directamente da directoria onde elas já se encontravam armazenadas em vez de andar a ocupar mais memória do computador desnecessariamente, armazenando também esses dados na BD. Esta directoria é armazenada num campo da BD para facilitar a leitura da imagem por meio do acesso à BD (a directoria do ficheiro de imagem é armazenada na variável 'img_filepath' e a directoria onde se encontram os ficheiros associados a este exame é armazenada na variável 'directory_of_this_exam_files', ambas acrescentadas na entidade 'Informação do Estudo'). Assim, a meta-informação é guardada na BD para se poder efectuar pesquisas de forma eficiente, mas os dados originais encontram-se armazenados num

sistema de arquivo, organizado por pastas como foi descrito no Capítulo 4, num dispositivo externo (possivelmente em vários no futuro).

As variáveis que preenchem as tabelas que contêm a informação extraída dos cabeçalhos dos ficheiros foram já apresentadas no capítulo anterior. Falta então ver a constituição das tabelas associadas à gestão da plataforma e ao processamento de dados. Esta encontra-se resumida na tabela 14.

Tabela 14 - Descrição das restantes entidades (e atributos que as constituem) que fazem parte da BD.

Entidade	Atributo	Tipo de Dados	Descrição
DB Access	ID_access	Long Integer	Número identificador associado a cada utilizador da BD.
	Profile	Characters(256)	Perfil do utilizador. Pode assumir os valores: 'Administrator', 'Health Professional' e 'Visitor'
	Username	Characters(256)	Nome de utilizador para aceder à BD.
	Password	Characters(256)	Palavra-chave de acesso à BD.
	Name	Characters(256)	Nome próprio do utilizador.
	Surname	Characters(256)	Sobrenome do utilizador.
	Sex_db	Short Integer	Sexo do utilizador. Pode assumir os valores: 0 (desconhecido/outro) 1 (masculino) e 2 (feminino).
	Birth_Date	Characters(64)	Data de nascimento do utilizador. Definida como um conjunto de 8 dígitos, em que os dois primeiros se referem ao dia, os dois segundos ao mês, e os 4 últimos ao ano.
	Career	Characters(256)	Profissão do utilizador.
	Nationality	Characters(256)	Nacionalidade do utilizador.
	Adress	Characters(256)	Morada do utilizador.
	Email	Characters(256)	<i>E-mail</i> do utilizador.
	Marital_Status	Short Integer	Estado civil do utilizador. Pode assumir os valores: <i>Single</i> (solteiro(a)), <i>Married</i> (casado(a)), <i>Divorced</i> (divorciado(a)) e <i>Widowed</i> (viúvo(a)).
Phone_Number	Long Integer	Número de telefone/telemóvel do utilizador.	
Photo_Dir	Text	Directoria de acesso à foto que o utilizador selecionou como sua imagem de perfil.	
Processing	ID_Processing	Long Integer	Número identificador do processamento realizado.
	processing_date	Characters(256)	Data em que o processamento ocorreu.
	processing_time	Characters(256)	Hora em que o processamento foi efectuado.
	processing_description	Characters(256)	Descrição do tipo de processamento levado a cabo.
Routines	routine_name	Characters(256)	Nome da rotina.
	routine_description	Text	Descrição da(s) funcionalidade(s) desempenhada(s) pela rotina.
	entity	Characters(256)	Entidade da BD à qual o <i>output</i> da rotina deve ser associado.
	input_db_variables	Text	Variáveis de entrada da rotina cujo valor se encontra armazenado na BD (separadas por ';').
	input_external_variables	Text	Variáveis de entrada da rotina cujo valor não se encontra armazenado na BD (separadas por ';').
	external_types	Text	Tipos de dados associados às variáveis de entrada da rotina cujo valor não se encontra armazenado na BD (separados por ';', e colocados na mesma ordem que as variáveis correspondentes).
	output_variables	Text	Variáveis de saída da rotina (separadas por ';').
	output_types	Text	Tipos de dados associados às variáveis de saída da rotina (separados por ';', e colocados na mesma ordem que as variáveis correspondentes).
Variables Description	function_call	Text	<i>Callback</i> utilizado para chamar a função/rotina.
	Variable	Text	Nome da variável.
	Type	Characters(256)	Tipo de dados que a variável armazena.
	Variable_Entity	Characters(256)	Entidade onde a variável se encontra.
	Description	Text	Descrição do que ela representa.

As tabelas EAV adicionadas são constituídas por 3 atributos, sendo o primeiro o identificador da entrada correspondente na entidade com que se relaciona (tipo: inteiro longo), o segundo é o nome do atributo que se pretende inserir (tipo: caracteres, até um máximo de 256), e o terceiro é o seu valor (tipo: ponto flutuante ou texto, conforme a necessidade).

Capítulo 7

Método de Implementação da Base de Dados

7. Método de Implementação da Base de Dados

A implementação desta Base de Dados sofreu diversas etapas:

1. Primeiro analisaram-se os requisitos, as necessidades de utilização, que esta plataforma deveria satisfazer.

2. Depois desenvolveu-se a estrutura base mediante a qual se pretendia organizar a informação (o modelo conceptual, parcialmente descrito no capítulo anterior), e esta foi alterada ao longo de todo o processo de acordo com as conveniências de manipulação dos dados e sentido lógico das associações entre diferentes tipos de informação.

3. Aquando da implementação desta a nível informático, surgiram diversos problemas que foram sendo resolvidos mediante modificações a nível do tipo de dados, entre outros.

4. Após se ter a BD devidamente estruturada e carregada, desenvolveu-se uma interface gráfica para lhe aceder e manipular/visualizar a sua informação.

Neste capítulo ir-se-ão discutir aprofundadamente as três primeiras etapas, deixando-se a quarta para um capítulo subsequente devido à sua extensão e complexidade.

7.1. Análise de Requisitos

De acordo com os objectivos que foram definidos no início deste projecto, a BD criada deve satisfazer o seguinte conjunto de requisitos:

1. Estruturar a informação a ser armazenada, com uma organização que: permita associar vários exames a um determinado paciente, e garantir que um determinado exame corresponde a um único paciente; permita processar os dados diversas vezes e guardar os resultados do processamento; permita adicionar sempre novos dados a cada uma das entidades principais (estrutura dinâmica e flexível), de modo a satisfazer as necessidades associadas aos utilizadores, ao crescimento da informação médica e à instituição onde está a ser implementada.

2. Definir vários perfis de utilização, de modo que um utilizador tenha acesso a determinado tipo de informação mediante o perfil que lhe foi atribuído. Medida que serve para restringir o acesso aos dados, mantendo a informação segura, e disponibilizando informação privada dos pacientes e informação técnica dos exames unicamente a profissionais que estejam autorizados a aceder-lhes.

3. Permitir adicionar, remover, actualizar e alterar a informação nela constante.

4. Disponibilizar funções de visualização dos exames e da meta-informação a eles associada (visualizadores de exames com algumas funcionalidades de processamento de imagem e motores de pesquisa na BD), de processamento e tratamento estatístico da mesma.

5. Ser uma plataforma modular flexível, que permita sempre a integração de novas rotinas (adição de novas funcionalidades).

7.2. Casos de Uso

De seguida apresenta-se o Diagrama de Casos de Uso, segundo o *standard* UML (*Unified Modeling Language* – é a linguagem utilizada para facilitar o desenvolvimento de *software* e de ferramentas de apoio à sua implementação, por meio da realização de diagramas que recriem o problema de programação em questão, e com vista a uniformizar a concepção destes em todo o mundo [33]), usado para a concepção desta BD, e cujas restrições de funcionalidades foram implementadas ao nível da interface gráfica desenvolvida posteriormente.

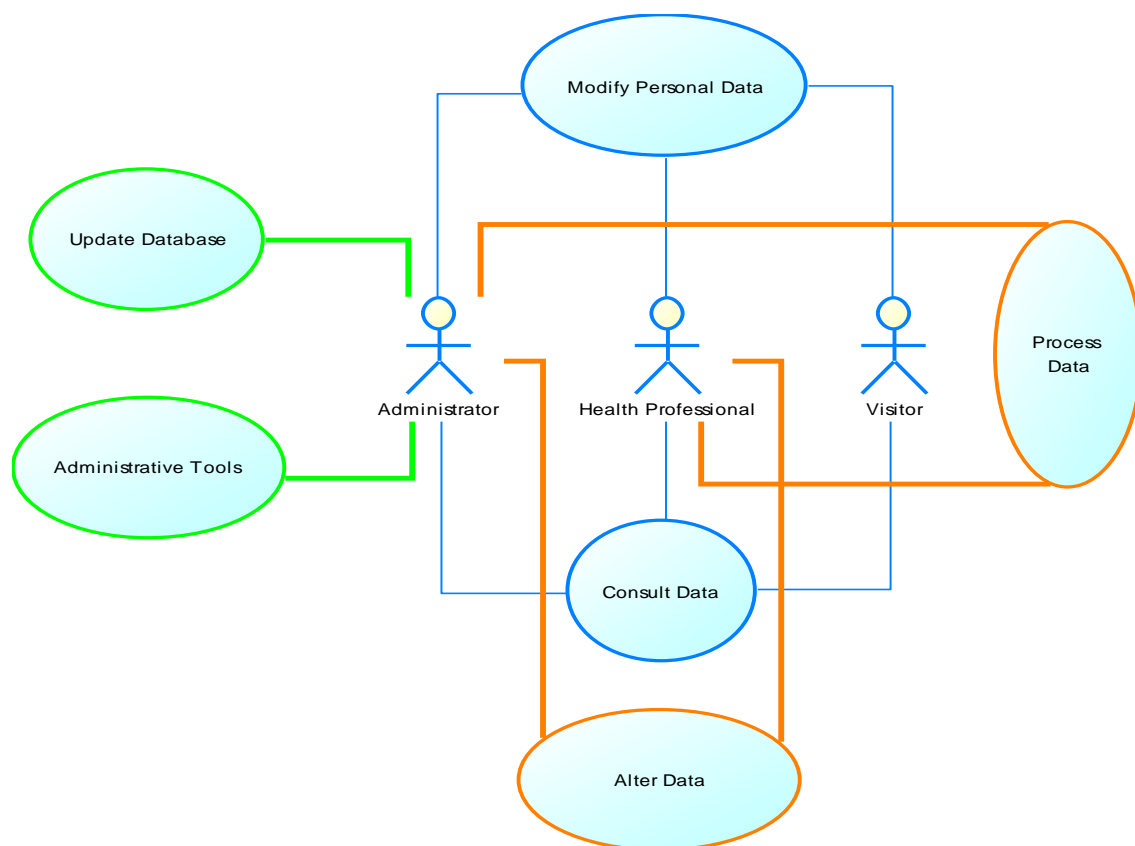


Figura 13 - Diagrama dos casos de uso definidos para a BD em desenvolvimento. (Diagrama desenvolvido na ferramenta Sybase Power Design 15[®])

Como podemos ver na Fig. 13, definiram-se três actores: o perfil de Administrador, o perfil de Profissional de Saúde e o perfil de Visitante. Todos eles podem desempenhar o papel de ‘Modificar a sua Informação Pessoal’ e ‘Consultar Dados’ (apesar de este último ter especificações diferentes para cada perfil como será discutido mais adiante nesta tese). Os papéis de ‘Alterar os Dados’ e ‘Processar Dados’ só podem ser interpretados pelo Administrador e o Profissional de Saúde, e o primeiro também tem diferente informação disponível para alteração para cada um dos perfis (para o Administrador é possível alterar todos os campos, enquanto que para o Profissional de Saúde, só aqueles que este tem competências para corrigir/alterar). Depois existem mais dois papéis que se encontram reservados ao Administrador, de modo a que este proceda à gestão e modificação da plataforma em si: ‘Actualizar a BD’ e ‘Ferramentas Administrativas’ (este último engloba uma série de funcionalidades para a gestão de utilizadores, informação existente e ferramentas implementadas).

7.3. Desenvolvimento da Base de Dados

Após se averiguar qual a organização que os dados deveriam ter, a estrutura base da BD, os requisitos que esta deve satisfazer e os casos de uso que lhes estão associados, desenvolveu-se um Modelo Conceptual a nível informático com todas as entidades, variáveis e relações referidas anteriormente. Este foi implementado recorrendo ao Sybase Power Designer 15[®]. Neste mesmo programa, gerou-se o Modelo Físico da Base de Dados a partir do modelo anterior, e por fim criou-se o *script* de geração da BD propriamente dito em linguagem SQL[®]. De seguida estudar-se-á cada um deles de maneira mais aprofundada.

7.3.1 Modelo Conceptual

Estudaram-se três modelos principais possíveis para a implementação da BD: o modelo Relacional, o modelo Entidade-Atributo-Valor (EAV) e o modelo Híbrido que é uma mistura dos dois anteriores. O primeiro é de estrutura fixa mas permite relacionar a informação de uma forma mais compreensiva, ao contrário do que acontece no modelo EAV, onde é necessário implementar uma interface com o utilizador que dê a percepção correcta da organização dos dados ao operador. No entanto, este modelo é o mais usado em ambientes hospitalares devido à

sua flexibilidade, isto é, a possibilidade de adicionar e remover variáveis sem modificar a estrutura raiz da BD [27].

No caso em estudo, a melhor solução seria a implementação de um modelo híbrido de forma a tirar vantagens dos outros dois, rectificando assim as suas falhas individuais. O modelo relacional individual é ideal para o armazenamento dos dados que são comuns a todos os ficheiros e que se correlacionam entre si hierarquicamente, enquanto que o modelo EAV permite a criação de tabelas onde novas variáveis podem ser adicionadas, relacionadas com as principais entidades ('Informação do Paciente', 'Informação do Estudo', 'Processamento', 'Parâmetros Globais dos Sub-Cabeçalhos' e 'Acesso à BD') incluídas no modelo relacional.

Assim, foi desenvolvido o modelo conceptual com a estrutura apresentada no Anexo F deste documento, que contém 342 campos/atributos, distribuídos por 36 entidades, que guardam a informação extraída dos ficheiros dos estudos e a informação relacionada com os utilizadores da BD, a descrição das variáveis nela incluídas, a descrição das rotinas que podem ser aplicadas para o processamento da informação armazenada dentro dela, os resultados desse processamento, e qualquer nova informação que possa surgir relacionando-se com um paciente, um exame ou um corte, quer seja do tipo numérico ou de caracteres.

A parte relacional do nosso modelo Híbrido é composta pelas entidades 'Informação do Paciente', 'Processamento' e 'Informação do Estudo', esta última mais as suas numerosas 'entidades-filhas' (sub-entidades mediante as heranças estabelecidas), de acordo com o modelo hierárquico da informação (elas herdaram a chave primária da 'entidade-mãe', o identificador do estudo e, no caso da 'sub-herança' desenvolvida a partir da entidade 'Parâmetros Globais dos Sub-Cabeçalhos', o identificador do sub-cabeçalho). Foi criada uma relação *one-to-many* (um para muitos) entre as entidades 'Informação do Paciente' e a 'Informação do Estudo', e assim o mesmo paciente pode estar associado a vários exames, enquanto cada estudo corresponde a um e um só paciente. Esta relação é obrigatória no sentido do estudo para o paciente, porque um estudo tem de pertencer necessariamente a um paciente, mas um paciente pode não ter realizado nenhum exame ainda. Para que um exame possa ser processado diversas vezes implementou-se uma relação *one-to-many* entre a entidade da 'Informação do Estudo' e a entidade 'Processamento', de maneira a que cada estudo possa estar associado a vários processamentos, e cada processamento se relacionar com um, e um só, estudo; cada processamento é forçado a estar ligado a um exame, mas um exame pode ou não ter um processamento a si associado – trata-se de uma relação obrigatória no sentido do processamento para o estudo.

As entidades adicionais, de carácter administrativo para guardar a informação relacionada com os utilizadores, as variáveis que lhe pertencem e as rotinas de processamento que podem ser aplicadas à meta-informação do estudo ou às suas imagens, são mantidas como entidades isoladas, independentes de todas as outras.

Também as tabelas EAV associadas a cada chave primária existente no modelo relacional, são mantidas como entidades independentes.

Vamos agora analisar as particularidades de cada uma das entidades presentes no modelo, tendo em conta os atributos que lhes pertencem e as chaves primárias que lhes foram atribuídas, para além de outras especificações resultantes do contexto do mundo real que envolve este projecto.

Começemos por abordar um tema comum a todas as tabelas: a chave identificadora primária que lhes foi atribuída. Por uma questão de evitar redundâncias na inserção de informação, e complicações nessa mesma inserção devido ao facto de se utilizar uma chave primária que possa ser facilmente repetida ou mal introduzida/definida, decidiu-se utilizar variáveis auxiliares para constituírem a chave primária. Estas chaves consistem simplesmente num número inteiro auxiliar que vai identificar uma nova entrada inserida na BD, pelo incremento de uma unidade do último identificador registado. Note-se que se definiu o tipo de dados desta variável de identificação como *Long Integer* (Inteiro Longo) para prevenir o facto de se vir a inserir um elevado número de entradas nas tabelas, pois não se sabe ao certo quantos exames a BD pode vir a armazenar no futuro [34]. Existem, no entanto, duas excepções a esta regra: as tabelas ‘Descrição das Variáveis’ e ‘Rotinas’. Nestes casos definiu-se o nome da variável e o nome da rotina, respectivamente, como chave primária, uma vez que não se podem inserir duas variáveis com o mesmo nome na BD, nem implementar duas vezes a mesma rotina. Como estes são nomes mais específicos, e que são verificados por ferramentas intrínsecas ao programa antes da sua introdução, podem ser usados como chaves primárias. Seria muito difícil aplicar uma identificação deste tipo nas outras tabelas, por causa dos erros que são cometidos ao nível de introdução de dados (por exemplo, nos nomes dos pacientes, ou mesmo uma qualquer descrição associada ao exame), e portanto é muito mais cómodo usar o tipo de identificação descrito anteriormente. Deve-se salientar também o facto das entradas de todas as tabelas relacionadas com os cortes, terem uma chave primária combinada (constituída por mais de um atributo; no caso em análise, por dois) constituída pelo identificador do estudo a que pertencem e, para distinguir qual o corte dentro desse estudo que lhes corresponde, o identificador do corte.

Falando agora de cada uma das tabelas em particular, na tabela ‘Acesso à BD’ colocaram-se as variáveis ‘Profile’, ‘Username’ e ‘Password’ como obrigatórias (<M> - *mandatory*), porque ao ser introduzido um utilizador é essencial que se defina qual o seu perfil para lhe fazer corresponder as funcionalidades que ele pode utilizar, e têm de se lhe fazer corresponder um nome de utilizador e uma palavra-chave para que ele possa aceder à plataforma. Outra especificação feita nesta tabela foi a atribuição do tipo de dados *Text* (texto) à variável ‘Photo_Dir’, que corresponde à directoria onde a foto de perfil do utilizador se encontra armazenada. Como as directorias geralmente envolvem longos conjuntos de caracteres, e não sendo possível estabelecer um valor como sendo o máximo para a extensão desse conjunto,

optou-se por definir todas as variáveis da BD que sirvam para armazenar directorias com este tipo de dados ('directory_of_this_exam_files' e 'img_filepath' na entidade 'Informação do Estudo'), que permite a introdução de um número variável de caracteres no campo, e geralmente é usado para grandes quantidades de texto [34]. Aplicou-se este critério também: para os valores das tabelas EAV que pretendem guardar caracteres, uma vez que não se sabe que novos valores podem vir a ser introduzidos e muito menos a sua extensão; para a variável 'Description' na tabela 'Descrição das Variáveis', visto que as definições das variáveis associadas aos exames costumam ser muito longas; e em todas as variáveis da entidade 'Rotinas', excepto o nome da rotina e da entidade onde os seus resultados devem ser guardados (que costumam ser muito curtos), por poderem ter que armazenar um elevado número de variáveis associadas à rotina em questão.

Na tabela 'Informação do Estudo' as directorias do exame em si, e da pasta onde os ficheiros a ele associados estão armazenados, foram declaradas como atributos de preenchimento obrigatório, pois é sempre necessário ter esta informação para proceder à contabilização e gestão dos ficheiros que já se encontram armazenados na BD, e para proceder à sua leitura quando necessário para qualquer tipo de processamento.

Na tabela 'Processamento' declararam-se as variáveis 'processing_date' e 'processing time' com carácter obrigatório para que se possa ter a noção temporal de tratamento dos dados em qualquer circunstância, e permitir pesquisas de quais os processamentos realizados num certo intervalo de tempo.

Por fim, no que respeita a todas as tabelas EAV, os três atributos que as constituem têm que ser sempre preenchidos (e foram para isso declarados como obrigatórios), uma vez que o intuito destas tabelas é sempre acrescentar um novo valor a uma dada entrada de uma das entidades principais do modelo relacional, e para isso é necessário sempre declarar qual o identificador da entrada na entidade à qual se está a adicionar nova informação, o nome do atributo que se está a inserir e o seu valor. No entanto, a chave primária para estas tabelas é somente a combinação entre o identificador da entrada e o nome do atributo, uma vez que, para uma dada entrada, um atributo não pode ter mais que um valor associado.

7.3.2. Modelo Físico

O modelo físico é um modelo lógico, onde já se retractam algumas restrições ao nível da nomenclatura e especificações da implementação informática; tem em conta as condições impostas pelo DBMS escolhido para a construção e execução da BD, no entanto é sem dúvida uma estrutura similar ao modelo conceptual, e criada com base nele [35]. O modelo físico para a nossa BD foi gerado automaticamente a partir do Modelo Conceptual anteriormente

apresentado, usando a ferramenta disponibilizada pelo Sybase Power Designer 15[®] para este efeito. Este modelo é apresentado no Anexo G desta tese.

7.3.3. Geração da Base de Dados

Fui usada a linguagem PostgreSQL[®] para geração e manipulação da BD, e usou-se a interface pgAdmin III[®] para a interacção com os dados *à priori* (na Fig. 14 apresenta-se um exemplo de uma das tabelas geradas, já preenchida, visualizada inicialmente por meio desta interface). O código utilizado para a criação da BD é apresentado em anexo a este trabalho (ver Anexo H).

	id_study [PK] bigint	nevent double precis	nsino double precis	ndelays double precis	reslice_ang1 double precis	reslice_ang2 double precis	reslice_ang3 double precis	scan_maxslc smallint	rebin_type smallint	concfac double precis	concfac_bgsu double precis	trailxists smallint	trailbeg bigint
1	3	0	2000536448	500486432	0	0	0	180	2	0	0	0	0
2	4	0	0	0	0	0	0	0	0	0	0	0	0
3	5	0	4808664	32531	0	0	0	180	2	0.089875	0.089875	0	0
4	6	0	4808664	32531	0	0	0	180	2	0.83	0.83	0	0
5	7	0	4808664	32531	0	0	0	180	2	0.83	0.83	0	0
6	8	491	315	0	0	0	0	0	0	0	0	0	0
7	9	491	315	0	0	0	0	0	0	0	0	0	0
8	10	0	4813828	32459	0	0	0	180	2	0.089875	0.089875	0	0
9	11	0	4813828	32459	0	0	0	180	2	0.83	0.83	0	0
10	12	0	4813828	32459	0	0	0	180	2	0.83	0.83	0	0
11	13	1067	891	0	0	0	0	0	0	0	0	0	0
12	14	1067	891	0	0	0	0	0	0	0	0	0	0
13	15	0	94628768	54014300	0	0	0	180	2	0.089875	0.089875	0	0
14	16	0	94628768	54014300	0	0	0	180	2	0.83	0.83	0	0
15	17	0	94628768	54014300	0	0	0	180	2	0.83	0.83	0	0
16	18	449	273	0	0	0	0	0	0	0	0	0	0
17	19	449	273	0	0	0	0	0	0	0	0	0	0
18	20	0	91172784	54221920	0	0	0	180	2	0.089875	0.089875	0	0
19	21	0	91172784	54221920	0	0	0	180	2	0.83	0.83	0	0
20	22	0	91172784	54221920	0	0	0	180	2	0.83	0.83	0	0
21	23	449	273	0	0	0	0	0	0	0	0	0	0
22	24	449	273	0	0	0	0	0	0	0	0	0	0
23	25	0	4288344832	288730368	0	0	0	180	2	0.089875	0.089875	0	0
24	26	0	4288344832	288730368	0	0	0	180	2	0.83	0.83	0	0
25	27	0	4288344832	288730368	0	0	0	180	2	0.83	0.83	0	0
26	28	455	279	0	0	0	0	0	0	0	0	0	0
27	29	455	279	0	0	0	0	0	0	0	0	0	0
28	30	1194320	1194769	449	0	0	0	180	2	0	0	0	0
29	31	1194320	1194769	449	0	0	0	180	2	0	0	0	0
30	32	461	285	0	0	0	0	0	0	0	0	0	0
31	33	461	285	0	0	0	0	0	0	0	0	0	0
32	34	0	416	5	0	0	0	180	2	0.071625	0.071625	0	0
33	35	0	51	2	0	0	0	180	2	0.83	0.83	0	0
34	36	0	51	2	0	0	0	180	2	0.83	0.83	0	0
35	37	1078	230	0	0	0	0	0	0	0	0	0	0
36	38	1078	230	0	0	0	0	0	0	0	0	0	0
37	39	0	42	0	0	0	0	180	2	0.089875	0.089875	0	0
38	40	0	42	0	0	0	0	180	2	0.83	0.83	0	0
39	41	0	42	0	0	0	0	180	2	0.83	0.83	0	0
40	42	1810	122	0	0	0	0	0	0	0	0	0	0
41	43	1810	122	0	0	0	0	0	0	0	0	0	0
42	44	0	169852320	75482408	0	0	0	180	2	0.10375	0.083125	0	0
43	45	452	276	0	0	0	0	0	0	0	0	0	0
44	46	452	276	0	0	0	0	0	0	0	0	0	0
45	47	0	169852320	75482408	0	0	0	180	2	0.10375	0.083125	0	0
46	48	0	169852320	75482408	0	0	0	180	2	0.10375	0.083125	0	0
47	49	0	169852320	75482408	0	0	0	180	2	0.071625	0.071625	0	0
48	50	0	169852320	75482408	0	0	0	180	2	0.83	0.665	0	0

100 rows.

Figura 14 - Representação parcial da tabela 'Estatísticas de Contagem', visualizada por meio da Interface pgAdmin III[®], utilizada inicialmente para a interacção com a BD após a sua geração, anteriormente à criação das rotinas em MATLAB[®] para o seu preenchimento e a sua manipulação (no entanto esta imagem foi adquirida já depois do preenchimento da BD). Recorreu-se a esta interface também, sempre que se necessitou de esclarecer qualquer tipo de dúvida relativa aos dados armazenados na BD até ao momento.

7.3.4. Carregamento da Base de Dados – Problemas e Restrições

Uma vez tendo a BD devidamente gerada, desenvolveram-se rotinas em MATLAB[®] para ler todas as directorias de armazenamento de ficheiros *imago* existentes na memória física externa onde os dados estão armazenados, chamar os programas de extracção fornecidos pela Philips[®], extrair e organizar a meta-informação dos cabeçalhos dos ficheiros, corrigindo certos parâmetros a fim de possibilitar a sua inserção na BD.

De seguida serão descritas sucintamente as funções desenvolvidas para o efeito:

- O programa 'ChargeDB.m' permite ao utilizador escolher o conjunto de dez directorias que lhe é possível escolher para proceder ao carregamento de ficheiros na BD (tenha-se presente que o nome dessas directorias se inicia sempre com 'p' de 'paciente', seguido pelo número de ordem em que os exames nelas armazenados foram executados na instituição), e chama a função 'read_charge_files.m' para proceder a esse carregamento. Nesta função o utilizador começa por declarar se é a primeira vez que estão a ser carregados dados na BD, escolhe o já referido conjunto, e no fim do carregamento executado, ela mostra-lhe uma lista com os ficheiros que por qualquer motivo não conseguiram ser abertos (por se encontrarem eventualmente corrompidos) e outra com os ficheiros que se encontravam comprimidos e portanto também não foi possível inseri-los na BD. Inseriram-se variáveis de controlo, designadamente 'no_files' e 'empty_db', para sinalizar se não existe nenhum ficheiro '.img' ou '.syn' nas directorias dos ficheiros a serem carregados e para declarar se esta é ou não a primeira vez que se inserem dados na BD, respectivamente. Este programa utiliza também as funções 'choose_elementX.m', 'choose_elementY.m' e 'choose_elementZ.m' para escolher o segundo elemento do nome no caso de se tratar de uma directoria constituída por 3 dígitos (exemplificando: 'p1XN', onde vão ser analisadas as directorias com N a variar de 0 a 9), para escolher o terceiro elemento de uma directoria com 4 dígitos (exemplificando mais uma vez: 'p1XYN, como N a variar da mesma forma que anteriormente) e para escolher o quarto elemento de uma directoria com 5 dígitos ('p1XYZN, como N a variar novamente de 0 a 9), respectivamente.

- A função 'read_charge_files.m' vai receber os parâmetros do programa 'ChargeDB.m' que lhe dão conta de quantos dígitos tem o conjunto de carregamento escolhido, o nome do conjunto de directorias seleccionado para o carregamento e se se trata ou não da primeira vez que a BD recebe informação. Escolhe-se previamente o comprimento do número identificador da directoria entre 2 e 5 dígitos associados à letra inicial 'p', e guarda-se essa escolha na variável 'option1'. Caso se tenha escolhido uma directoria com 2 dígitos, é chamada a função 'read_img.m' para proceder à extracção dos dados dos ficheiros. No caso de a directoria ter 3 dígitos é usada a função 'read_img3.m', no caso de ter 4 dígitos a função 'read_img4.m' e no caso de ter 5 dígitos a função 'read_img5.m. Caso não existam ficheiros novos para serem

carregados nas directorias analisadas, é mostrada uma mensagem ao utilizador na linha de comandos a informar deste facto. Em caso contrário, procede-se à correcção de possíveis dados corrompidos existentes entre os dados extraídos pelas funções listadas anteriormente (assunto que será abordado mais adiante nesta secção), e os dados são inseridos na BD mediante a estruturação implementada, e com identificadores incrementados a partir do valor numérico do identificador do último exame e do último processamento inserido na BD. Se se tratar do primeiro carregamento a ser efectuado à BD, são também inseridos os dados relativos aos três Administradores base da plataforma (os coordenadores deste projecto e a sua desenvolvedora) e as informações relativas a todas as variáveis que constituem a BD inicialmente.

- As funções ‘read_img.m’, ‘read_img3.m’, ‘read_img4.m’ e ‘read_img5.m’ funcionam da mesma maneira no processo de extracção dos dados dos ficheiros de exame, só mudando o conjunto de directorias por que procuram, como já foi explicado. Assim sendo, de um modo geral, estas funções começam por verificar qual o último identificador de paciente inserido, ou, sendo o primeiro carregamento, atribuem o valor 1 ao identificador do paciente inicial. Procede-se à incrementação de uma unidade desse valor, somente no caso de existirem ficheiros *imago* ou de extensão ‘.syn’ dentro das pastas dos estudos presentes na pasta do paciente que está a ser lida no momento. Depois vão listar todos os ficheiros ‘.img’ e ‘.syn’ que existem nas directorias em análise, mas só vão preparar para carregamento na BD aqueles que não se encontrarem já nas suas entradas. Criam uma lista com os ficheiros que não consigam abrir, e todos os outros são analisados para verificar se se encontram comprimidos, por meio da função ‘is_compressed.m’. Criam também uma lista com os ficheiros que se encontrem comprimidos, caso existam alguns, e os restantes são analisados para extracção dos dados dos seus cabeçalhos. Para tal são usadas as funções da ‘Luminary Tools’, já descritas no capítulo 5 deste trabalho, mediante o uso da função ‘rdimg.m’ e da função ‘rdimg2.m’ (adaptada da primeira, para resolver problemas de extracção dessa rotina – este assunto também será abordado mais adiante nesta secção).

Ao longo de todo o processo descrito, e também da interface que será apresentada no capítulo seguinte, são também utilizadas funções desenvolvidas para a conexão (‘connect.m’), execução de uma query (‘execute_sql.m’) e actualização (‘update.m’) dos campos da BD. São também utilizadas funções auxiliares para o processamento dos dados:

- ‘separate_by_semicolom’: separa os itens de um conjunto de caracteres mediante a existência do sinal de pontuação ‘;’ no seu interior.
- ‘remove_spaces’: remove os espaços existentes entre um conjunto de caracteres.
- ‘remove_spaces_char’: faz o mesmo que a função anterior, mas quando é necessário converter os dados para caracteres antes da operação de remoção dos espaços.

Tem de se desligar o MATLAB[®] entre cada carregamento devido à grande ocupação de memória que a extracção destes grandes volumes de meta-informação implica (consegue-se ter abertos simultaneamente somente à volta de 90 ficheiros *imagio*).

No decorrer deste processo descobriram-se vários problemas e falhas associados às rotinas em MATLAB[®] desenvolvidas pela Philips[®]. Seguidamente estes serão listados, bem como as acções empreendidas para resolvê-los.

1. Nem todas as variáveis presentes no mapa de *bytes* de cada ficheiro *imagio* (descrito no manual das 'Luminary Tools') estão parametrizadas nas rotinas 'getmhdr.m' e 'getshdr.m'. Assim, foram criadas novas versões destas rotinas onde as variáveis em falta foram adicionadas.

2. Mesmo após a adição, ficariam em falta as variáveis 'sop_uid' e 'recon_method' pertencentes aos sub-cabeçalhos. Para se proceder à extracção destas duas variáveis criou-se a rotina 'rdimg2.m' que vai chamar a rotina 'getshdr_sop_recon.m', que consiste numa versão da extracção dos sub-cabeçalhos que permite a extracção das duas variáveis mencionadas.

3. Alterou-se a rotina de identificação de ficheiros comprimidos desenvolvida em MATLAB[®] pela Philips[®], designada 'is_compressed.m', devido ao facto de esta identificar aquisições dinâmicas sempre como ficheiros comprimidos. Corrigiu-se esta rotina mediante o programa desenvolvido pelo Eng.º Nuno Ferreira para executar esta função em IDL[®], que funcionava bem para todos os casos (independentemente de se tratar de um exame dinâmico ou não).

4. As mensagens de erro pré-definidas para quando os ficheiros se encontram comprimidos foram removidas dos ficheiros 'rdimg.m' e 'rdimg2.m', bem como a possibilidade de descomprimir os ficheiros por falta do *software* que o permite (programa 'imagio_uncompress').

5. Alguns campos extraídos dos cabeçalhos encontravam-se corrompidos com caracteres estranhos. Para ser possível a sua inserção na BD, restringiram-se a códigos ASCII entre o valor 32 e o valor 127. As variáveis de caracteres (ou seja, as passíveis de apresentarem caracteres estranhos) foram todas corrigidas a fim de evitar possíveis erros nos carregamentos de dados futuros.

6. Devido a haver truncamento dos dados ao serem inseridos na BD, foi necessário alterar o tipo de dados de diversas variáveis. Alterou-se o tipo de dados da variável 'duratr' de *Short Integer* para *Long Integer*, da variável 'scnlen' de *Short Integer* para *Long Integer* na tabela 'Parâmetros da *Frame*' e das variáveis 'mseclen' e 'bedpos' do tipo *Short Integer* para *Integer*. Mudaram-se todas as variáveis presentes na tabela 'Parâmetros de Alinhamento PET/CT' do tipo *Short Integer* para o tipo *Integer* por haver também truncamento de dados aquando do carregamento.

7. Alteraram-se os tipos das variáveis definidas inicialmente como float[2], float[3] e float[6] para float[8] devido a estes formatos não serem suportados pela ferramenta de PostgreSQL[®] utilizada.

8. Na tabela 'Parâmetros do Corte' mudaram-se os nomes de duas variáveis que eram designadas, segundo a Philips[®], por nomes iguais noutras tabelas: as variáveis 'sop_uid' e 'decay_corr' que se encontram nas tabelas 'Informação do Estudo' e 'Parâmetros de Reconstrução', respectivamente. Passaram, assim, a ser designadas 'sop_uid_slice' e 'decay_corr_slice' na tabela 'Parâmetros do Corte'.

Capítulo 8

Desenvolvimento da Interface de Comunicação com a Base de Dados

8. Desenvolvimento da Interface de Comunicação com a Base de Dados

Tendo a BD devidamente preenchida, começou-se a explorar a melhor maneira de estabelecer o contacto entre ela e o utilizador, para possibilitar a sua fácil consulta, manipulação e compreensão, assim permitindo estudos e pesquisas nos dados e, futuramente, retirar conclusões estatísticas que ajudem a prevenir os erros/problemas passados.

Como já foi descrito, as rotinas para a extracção e carregamento dos dados encontram-se desenvolvidas em MATLAB[®]. Devido a este facto e também às potencialidades desta linguagem de programação ao nível do Processamento de Imagem (entre outras funcionalidades, devido às suas numerosas *Toolboxes*), optou-se por se desenvolver a interface de comunicação com a BD utilizando este *software*.

8.1. Aceder e Partilhar a Informação

O acesso à BD encontra-se restrito a utilizadores com credenciais de acesso que estejam armazenadas na entidade ‘Acesso à Base de Dados’. Ao nível da Interface implementou-se um menu de Boas-Vindas onde o utilizador deve inserir essas credenciais. Quanto à partilha de informação, foram criados três perfis, para que cada um apenas pudesse aceder à informação que o seu estatuto na Instituição ou grau de conhecimento lhe permita tratar. Os perfis desenvolvidos já foram referidos (ver, por exemplo, Fig. 13). Os administradores da BD têm tarefas que lhe são exclusivas, relacionadas com a gestão e actualização da plataforma. Depois partilham com os Profissionais de Saúde a possibilidade de processar dados, consultar dados dos pacientes e alterar/corrigir os mesmos, bem como a informação relacionada com as aquisições feitas a esses pacientes. Os Profissionais de Saúde não podem, no entanto, alterar as informações relativas ao processamento de dados e às descrições das variáveis existentes na BD, estas alterações são exclusivas do perfil de Administrador. Todos os três podem alterar os seus dados pessoais, e consultar os dados existentes na BD, no entanto, ao nível da consulta, os visitantes têm restrições de acesso aos dados dos pacientes para preservar a privacidade destes.

De seguida apresenta-se um esquema das funcionalidades disponíveis para cada um dos perfis (Fig. 15).

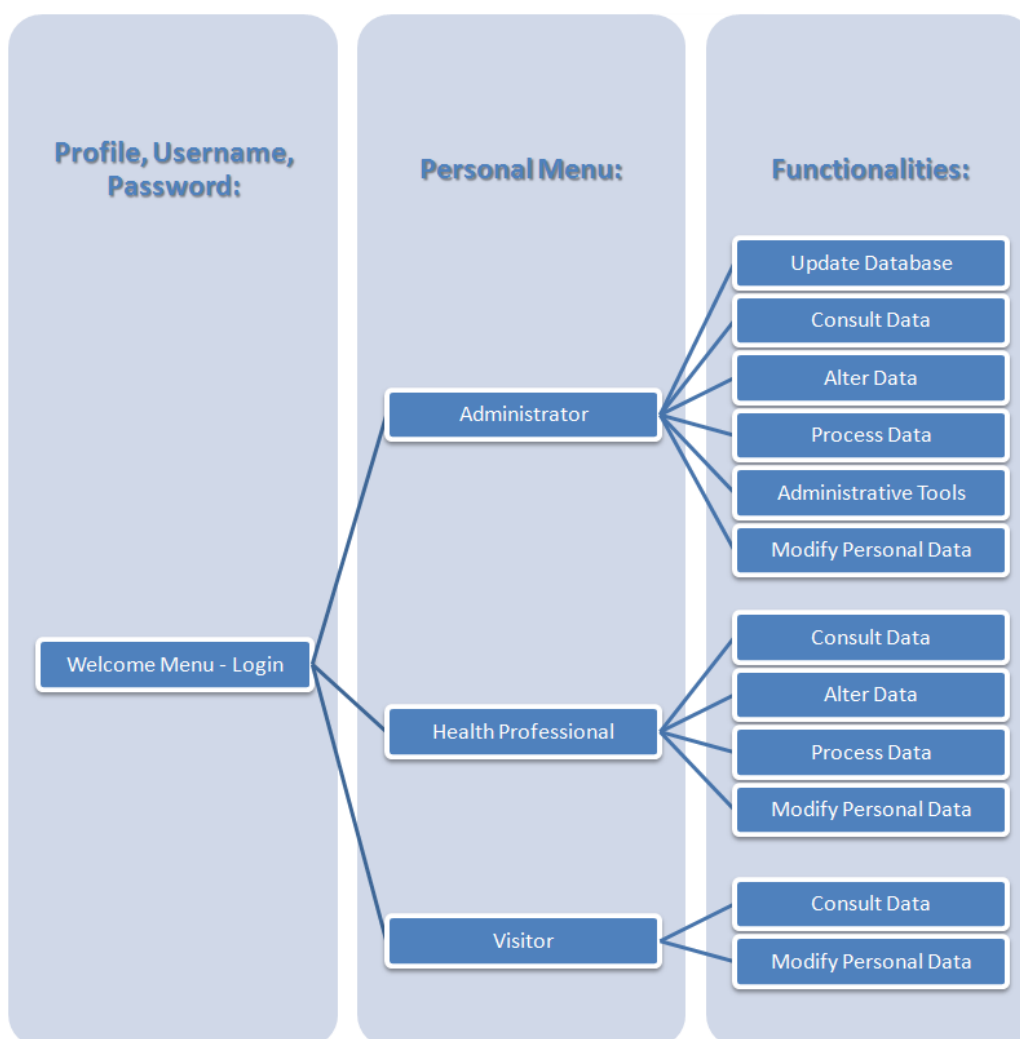


Figura 15 - Esquema representando as funcionalidades da BD disponíveis para cada perfil.

Até ao momento, a BD apenas permite o acesso de um único utilizador de cada vez, e a execução de uma única tarefa continuamente (isto é, se a execução for interrompida, os resultados do processamento de dados são ignorados).

No entanto, a sua grande mais valia prende-se com o facto de, após o processamento de dados, realização de uma pesquisa ou processamento de uma imagem, os resultados serem guardados na BD ou em ficheiros (cuja directoria se encontra guardada num campo da BD, ou é mostrada ao utilizador depois da geração do ficheiro), e depois permanecem disponíveis para estudos adicionais/futuros e para utilização dos dados por outros operadores. De ressaltar que essa disponibilidade desaparece no que respeita à informação pessoal dos pacientes, que nunca é armazenada em ficheiros gerados para guardar a informação processada, de maneira a preservar a sua privacidade.

Utilizando o perfil de Administrador como exemplo, na Fig. 16 é apresentado um diagrama que resume as acções específicas que podem ser executadas pelo utilizador através da interface desenvolvida.

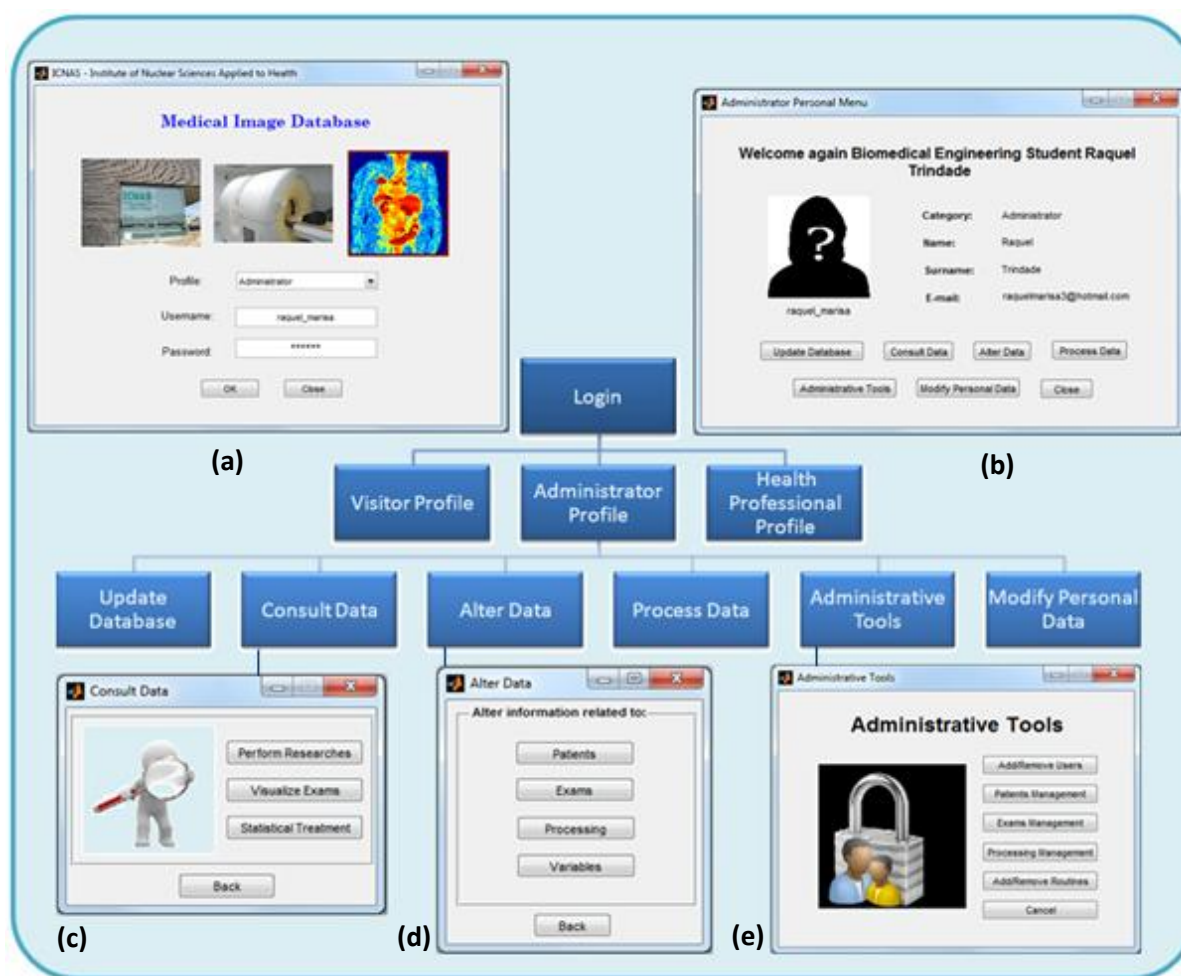


Figura 16 - Diagrama ilustrando as funcionalidades disponibilizadas pela Interface desenvolvida.

Podemos ver que o processo se inicia com a introdução por parte do utilizador das suas credenciais de acesso (perfil, nome de utilizador e palavra-chave) no menu de Boas-Vindas (Fig. 16 – (a)). Se estas estiverem correctas, o menu pessoal do utilizador aparece. As funcionalidades disponibilizadas neste vão depender do perfil do utilizador (que define o grau de permissão e competência do utilizador para aceder e manipular os dados presentes na BD), como vimos anteriormente. Para o Administrador todas as ferramentas estão disponíveis (Fig. 16 – (b)). A rotina ‘Update Database’ disponibiliza-lhe duas acções: verificar se algum ficheiro foi apagado da memória física, e apagá-lo também da BD; e carregar novos ficheiros, previamente seleccionados pelo Administrador (um máximo de 10 pastas de ficheiros em cada carregamento). O menu ‘Consult Data’ é constituído por três opções (Fig. 16 –(c)): a realização de pesquisas na BD de acordo com os valores de variáveis que o utilizador seleccione; a visualização de exames e também algum processamento das imagem neles contidas; e proceder ao tratamento estatístico de alguns dos parâmetros contidos na BD (funcionalidade ainda não implementada – ver secção 10.1). O menu disponibilizado pelo botão ‘Alter Data’, permite seleccionar entre alterar a informação presente nos campos da BD respeitante aos pacientes, aos exames, aos processamentos ou a descrição das variáveis dela constantes (Fig. 16 – (d)). De

ressalvar que para o perfil de Profissional de Saúde o menu de alteração de informação só inclui os dois primeiros botões. O ‘Process Data’ disponibiliza uma ferramenta para implementar um conjunto de rotinas a um grupo de exames, sendo a constituição de ambos escolhida pelo utilizador; permite visualizar e/ou guardar o resultado deste processamento. A funcionalidade ‘Administrative Tools’ apresenta uma colecção de sub-menus para realizar a gestão da informação armazenada: adicionar e remover utilizadores, pacientes, exames, processamentos e rotinas (Fig. 16 – (e)). Por fim, a opção ‘Modify Personal Data’ possibilita ao utilizador modificar a sua informação pessoal.

Terminar esta secção referindo só o modo como a inicialização da Interface é realizada. Colocou-se somente os programas respeitantes ao menu inicial da interface no ambiente de trabalho da rede, permitindo o acesso a estes ficheiros por todos os utilizadores da rede computacional do ICNAS. No entanto, as restantes funções da plataforma, bem como outras ferramentas necessárias para a sua execução foram armazenadas na pasta ‘Medical Image Database’, cujo acesso deve ser partilhado somente com os administradores da BD, para que possam adicionar-lhe novas rotinas. Ao se introduzirem credenciais válidas para o acesso à BD, a directoria de trabalho do MATLAB[®] é estabelecida na pasta ‘Medical Image Database’. Os seus programas são executados, mas não são vistos pelo utilizador – o utilizador comum não precisa de ter conhecimento dos detalhes internos à execução da aplicação.

8.2. Interface para o Processamento e Análise de Dados

Nesta secção ir-se-ão apresentar cada um dos menus que constituem a interface criada, descrevendo as funcionalidades de cada item que os constitui, ou seja, as ferramentas de processamento e análise de dados que estão na base das acções disponibilizadas pela plataforma.

Como já foi referido, a interface é iniciada com um menu de Boas-Vindas ao utilizador, onde lhe são pedidas as suas credenciais de acesso à plataforma (Fig. 17).

O utilizador deve introduzir o seu nome de utilizador e a sua palavra-chave nos campos destinados para o efeito, e seleccionar o seu perfil de utilizador de entre os três possíveis (‘Administrador’, ‘Profissional de Saúde’ e ‘Visitante’). Note-se que, a palavra-chave vai sendo mascarada com asteriscos à medida que é introduzida. Após o devido preenchimento das suas credenciais de acesso, o utilizador deve validá-las premindo o botão ‘OK’. Após esta acção o sistema vai procurar a combinação perfil+nome de utilizador+palavra-chave introduzidos numa das entradas da tabela ‘Acesso à Base de Dados’. Caso esta não seja encontrada por falha no nome de utilizador introduzido, é mostrada a mensagem de erro apresentada na Fig. 18.

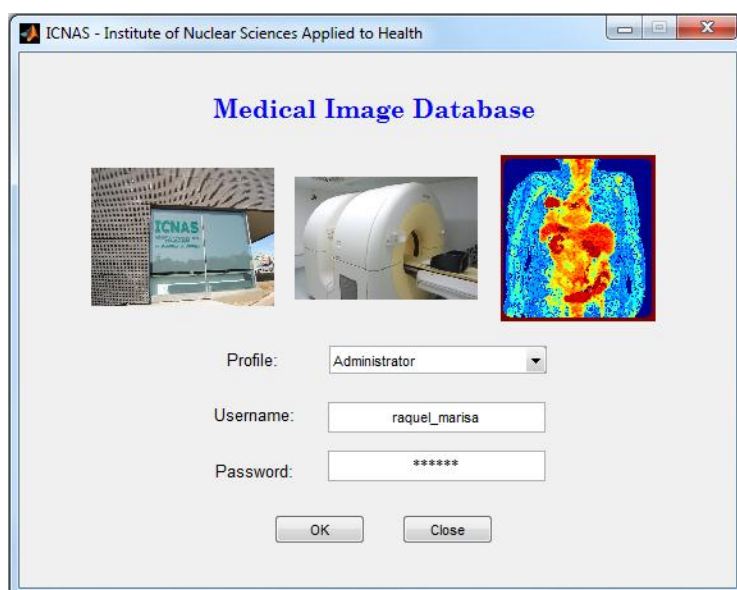


Figura 17 - Menu de Boas-Vindas da Interface desenvolvida para se aceder à BD criada.

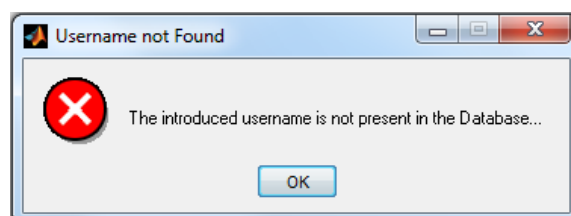


Figura 18 - Mensagem de erro apresentada ao utilizador da BD quando o nome de utilizador por si introduzido não se encontra presente entre os autorizados a aceder à plataforma.

No caso do nome de utilizador introduzido constar da referida tabela, mas se a palavra-chave a ele associada não for igual à introduzida pelo utilizador que está a tentar aceder à BD, o utilizador é informado que a palavra-chave introduzida está incorrecta (Fig. 19), e não lhe é permitido o acesso à plataforma enquanto não introduzir a combinação correcta nos campos do menu inicial.

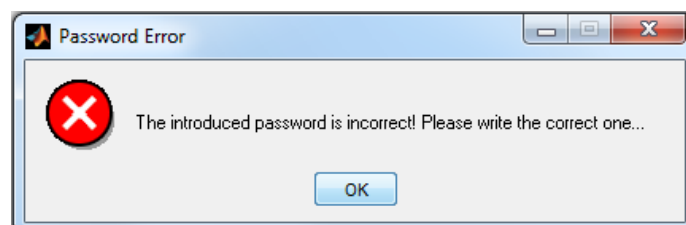


Figura 19 - Mensagem de erro mostrada quando é introduzida a palavra-chave incorrecta para um dado nome de utilizador constante da BD.

No caso da combinação de credenciais de acesso estar correcta, o utilizador acede à plataforma, sendo-lhe mostrado o seu menu pessoal, com funcionalidades disponíveis de acordo com o perfil que lhe corresponde (que já foram explicitadas na secção anterior deste trabalho).

De salientar alguns aspectos relacionados com o funcionamento geral de toda a interface: as mensagens de erro, de aviso e de informação (por exemplo, do sucesso de execução de uma determinada operação) mostradas contêm um botão *OK* que apenas serve para fechar a mensagem, após a sua leitura por parte do utilizador; o botão *Close* apresentado procede sempre ao encerramento da aplicação por inteiro (de todos os menus); para se encerrar apenas o menu actual e voltar ao menu anterior devem seleccionar-se os botões *Cancel* ou *Back*, conforme a situação adequada.

Para se analisarem os restantes menus desenvolvidos para a construção da Interface, vai utilizar-se o perfil de um Administrador da BD como exemplo.

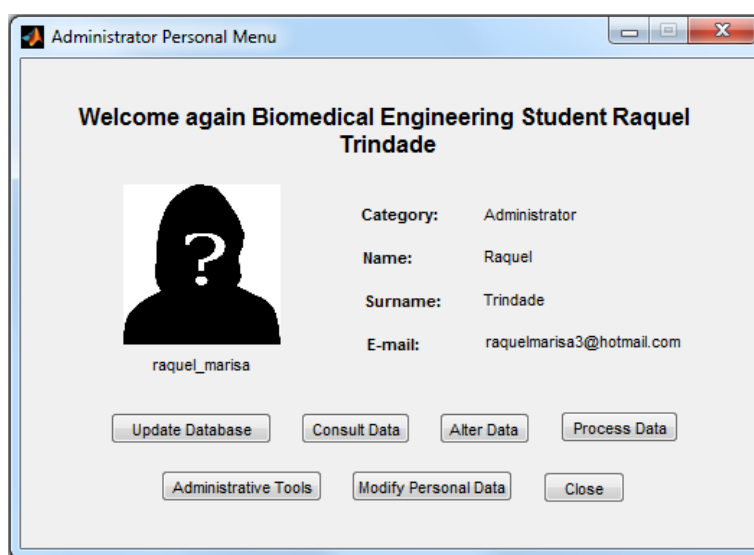


Figura 20 - Menu pessoal de um utilizador da BD com o perfil de Administrador.

Como se pode ver na Fig. 20, no menu de perfil de cada utilizador é apresentado um pequeno resumo da sua informação pessoal, caso esta tenha sido introduzida previamente nos campos destinados para o efeito na BD (processo que será explicado mais à frente). Caso essa informação se encontre parcialmente preenchida, os campos cuja informação não se encontre na BD são apresentados em branco, os outros preenchidos e o título ou é semi-preenchido com a informação existente, ou apresenta simplesmente a mensagem ‘Welcome again...’ se nenhuma existir. Caso nenhuma informação tenha sido ainda introduzida, apenas o campo destinado ao nome de utilizador (abaixo da foto de perfil) e o campo ‘Category’ (respeitante ao perfil do utilizador) serão preenchidos, visto tratarem-se de informações de preenchimento obrigatório para a introdução de um utilizador na BD. Analisando especificamente o que acontece com a foto de perfil, primeiro averigua-se se na BD está inserida a directoria correspondente à imagem que deve ser apresentada. Se esta existir, a imagem guardada nessa directoria é mostrada no menu, caso contrário vai-se analisar a informação relativa ao sexo do utilizador. Se se tratar do sexo feminino é apresentada a imagem de uma silhueta feminina, para todos os outros

(incluindo a possibilidade de a informação relativa ao sexo também não se encontrar preenchida) é apresentada a imagem de uma silhueta masculina.

Vão-se agora analisar os menus desenvolvidos para que um utilizador (neste caso, um Administrador) possa desempenhar cada uma das tarefas disponibilizadas pela BD.

8.2.1. Processo de Actualização da Base de Dados

Começamos então pela funcionalidade de ‘Actualização da BD’ (*Update Database*). Primeiramente é mostrado ao Administrador um menu onde ele selecciona o tipo de actualização que pretende desempenhar (Fig. 21). Deve seleccionar pelo menos um tipo (caso contrário será informado que o deve fazer pela mensagem presente na Fig. 22), podendo realizar os dois de uma vez. Caso se cancele a operação, é mostrada a mensagem de erro presente na Fig. 23, e o menu de actualização é encerrado.

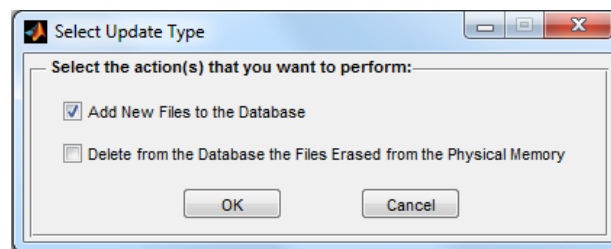


Figura 21 - Menu de selecção do tipo de actualização que o Administrador pretende realizar.

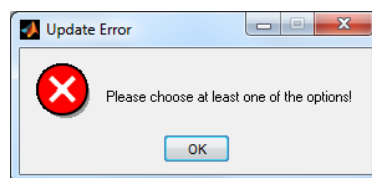


Figura 22 - Informação ao utilizador da necessidade de seleccionar pelo menos um tipo de actualização.

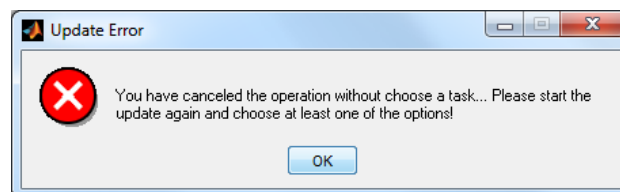


Figura 23 - Mensagem de erro mostrada quando se cancela a operação.

Para realizar qualquer dos tipos de actualizações à BD, é necessário estabelecer previamente a conexão ao dispositivo físico de memória onde a informação se encontra

armazenada ('BDDISK'), se tal não for realizado será mostrada a mensagem de erro mostrada nas Fig. 24.

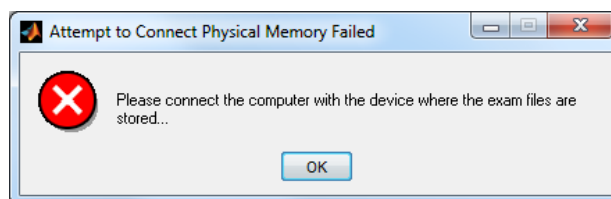


Figura 24 - Mensagem de erro mostrada quando se pretende proceder à actualização da BD e a conexão ao dispositivo de memória onde os ficheiros se encontram não foi previamente estabelecida.

Se a opção seleccionada for a de adicionar novos ficheiros à BD, é utilizada uma adaptação da rotina de carregamento da BD para realizar esta tarefa, e então o utilizador é avisado que deve trabalhar ao nível da linha de comandos do MATLAB® (Fig. 25).

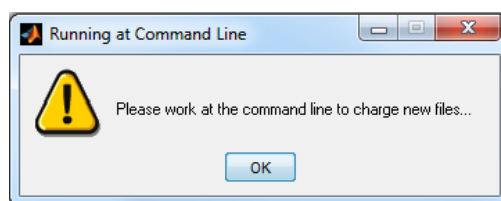


Figura 25 - Mensagem de aviso para que o utilizador trabalhe ao nível da linha de comandos.

Deve então escolher o tipo de nome das pastas associadas aos novos estudos que pretende inserir na BD. Primeiro deve seleccionar o número de dígitos presentes nos nomes das directorias onde pretende procurar os novos ficheiros (2, 3, 4 ou 5 dígitos, correspondendo às opções 1, 2, 3 e 4, respectivamente), para lá da letra 'p' pela qual se iniciam (Fig. 26). De salientar que existem restrições para limitar a escolha do utilizador ao número de opções existentes, e que o avisam caso tentem introduzir outros valores.

```
Command Window
New to MATLAB? Watch this Video, see Demos, or read Getting Started.
>> Medical_Image_Database_Welcome

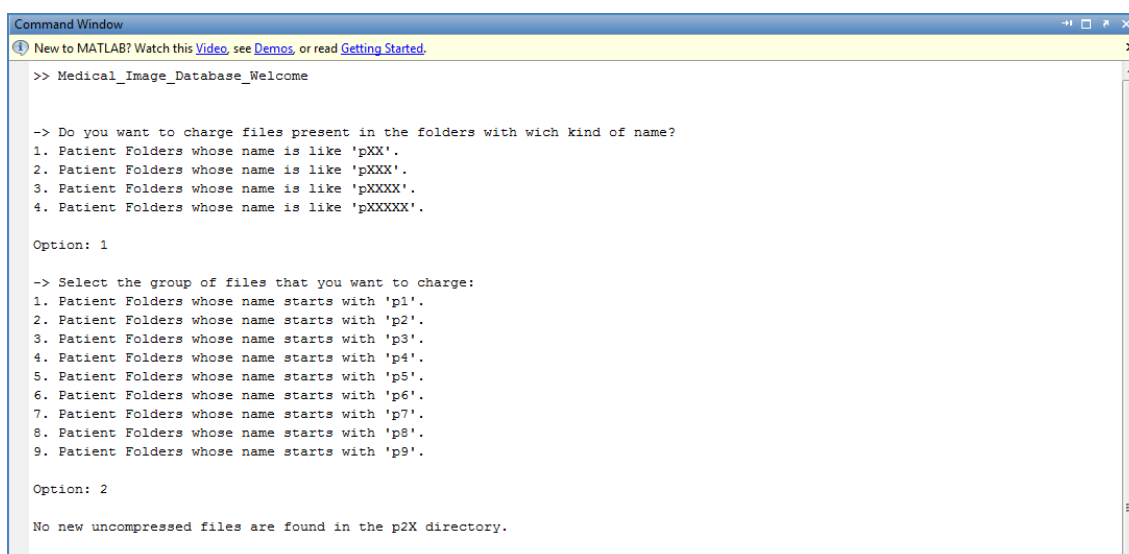
-> Do you want to charge files present in the folders with wich kind of name?
1. Patient Folders whose name is like 'pXX'.
2. Patient Folders whose name is like 'pXXX'.
3. Patient Folders whose name is like 'pXXXX'.
4. Patient Folders whose name is like 'pXXXXX'.

fx Option:
```

Figura 26 - Apresentação parcial da linha de comandos do MATLAB® onde é pedido ao utilizador para inserir a sua opção.

De seguida um sistema de opções semelhante é apresentado para o utilizador escolher on números que pretende que estejam no formato do nome da directoria seleccionado anteriormente. Na Fig. 27 apresenta-se um exemplo no qual o utilizador selecciona as

directorias do tipo 'pXN', e depois escolhe o número dois para substituir o primeiro dígito do nome. Assim serão procurados novos ficheiros nas dez directorias de 'p20' até à 'p29'. De maneira genérica, atribuem-se sempre valores fixos aos primeiros dígitos da directoria, e o último dígito vai variar de 0 a 9, carregando os ficheiros que ainda não se encontrem na BD existentes nessas dez directorias. Restringiu-se este procedimento ao carregamento de apenas dez pastas de cada vez devido ao facto dos ficheiros a carregar serem muito pesados, de modo que um grande volume de ficheiros deste tipo abertos leva a memória utilizada pelo MATLAB® à exaustão, levando à ocorrência de erros que páram o processo impedindo o carregamento. Devido ao processo de carregamento ser muito demorado, foram implementadas barras de espera no seu decorrer, como a mostrada na Fig. 28. O tempo de carregamento é variável: vai depender do número de ficheiros de extensão '.img' que estão nas dez directorias escolhidas para carregamento, e também do facto de esses ficheiros se encontrarem comprimidos ou não (quando há muitos ficheiros comprimidos nessas pastas, o processo não demora quase nada porque esses ficheiros não são lidos). Mediante a observação directa do processo, estima-se que o carregamento das dez directorias seleccionadas possa variar entre 10 a 40 minutos. Como estão sempre a ser realizados novos exames de PET/CT no ICNAS, a actualização da BD deve ser feita semanalmente, tendo em conta as designações das pastas que foram adicionadas.



```
Command Window
New to MATLAB? Watch this Video, see Demos, or read Getting Started.
>> Medical_Image_Database_Welcome

-> Do you want to charge files present in the folders with wich kind of name?
1. Patient Folders whose name is like 'pXX'.
2. Patient Folders whose name is like 'pXXX'.
3. Patient Folders whose name is like 'pXXXX'.
4. Patient Folders whose name is like 'pXXXXX'.

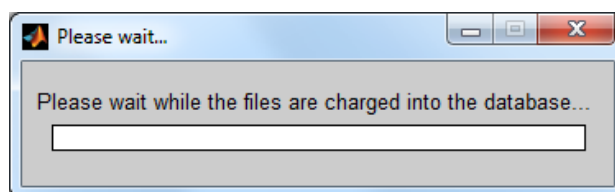
Option: 1

-> Select the group of files that you want to charge:
1. Patient Folders whose name starts with 'p1'.
2. Patient Folders whose name starts with 'p2'.
3. Patient Folders whose name starts with 'p3'.
4. Patient Folders whose name starts with 'p4'.
5. Patient Folders whose name starts with 'p5'.
6. Patient Folders whose name starts with 'p6'.
7. Patient Folders whose name starts with 'p7'.
8. Patient Folders whose name starts with 'p8'.
9. Patient Folders whose name starts with 'p9'.

Option: 2

No new uncompressed files are found in the p2X directory.
```

Figura 27 - Exemplo da selecção de uma série de directorias para carregamento na BD.



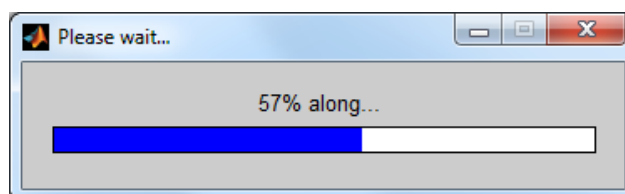


Figura 28 - Exemplo de uma das barras de espera implementadas, para que o utilizador tenha a noção do progresso da execução do carregamento de dados.

Caso não existam novos ficheiros na série de directorias seleccionadas, o utilizador é avisado com uma mensagem semelhante à representada no final da linha de comandos do exemplo mostrado na Fig. 27. Se existirem novos ficheiros a serem carregados em alguma das directorias seleccionadas o utilizador recebe um aviso a informar do sucesso da operação (Fig. 29). A mensagem mostrada nesse pode variar de acordo com o facto de existirem na série de ficheiros examinados algum ficheiro comprimido, ou algum ficheiro que por qualquer motivo não tenha conseguido ser aberto para leitura pelo MATLAB[®]. No caso de existirem ficheiros nestas condições, no final da execução do programa, é apresentada uma lista com as directorias dos ficheiros em cada uma das situações referidas.

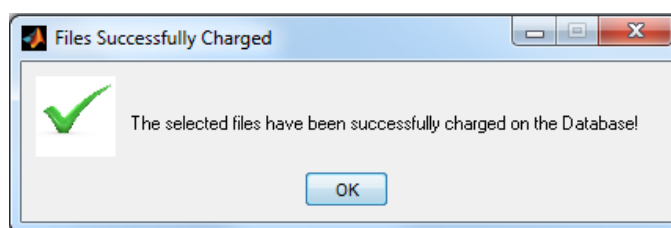


Figura 29 - Aviso para notificar o utilizador do sucesso da operação de carregamento.

Devido à referida questão da exaustão de memória devido ao enorme tamanho dos ficheiros imagiológicos, no final do processo de carregamento o utilizador é aconselhado a reiniciar o MATLAB[®] pela janela mostrada na Fig. 30.

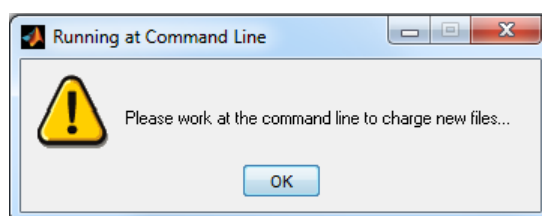


Figura 30 - Aviso mostrado para que o utilizador reinicie o MATLAB[®] no final do carregamento de dados.

Se se optar pela remoção da BD de ficheiros que tenham sido apagados da memória física, primeiro o programa faz uma listagem dos ficheiros de imagem currentemente existentes no dispositivo 'BDDISK' (que por ser um processo demorado, lhe foi implementada uma barra

de espera – Fig. 31), e depois vai verificar se existem ficheiros na BD que não constem dessa lista.

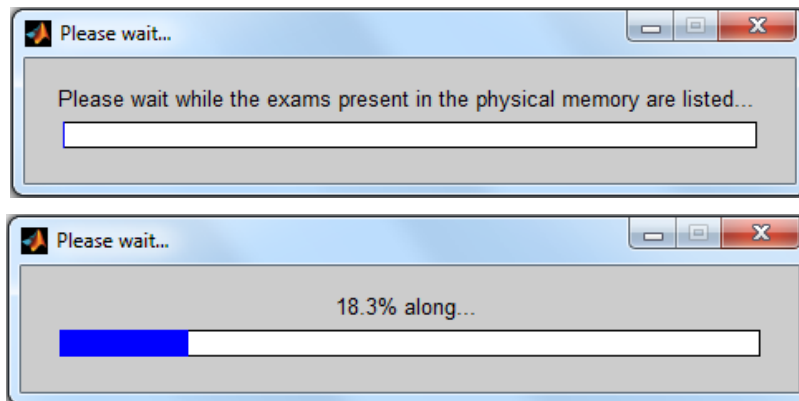


Figura 31 - Barra de espera para notificar o utilizador da progressão da leitura dos exames presentes na memória física.

Se nenhum ficheiro tiver sido apagado, e mostrada uma mensagem com esta informação ao utilizador (representada na Fig. 32). No caso de haver ficheiros na BD que não constem do dispositivo de armazenamento, esses serão apagados também das entradas da BD (no fim é também apresentada uma mensagem ao utilizador a informá-lo do término da operação de remoção – Fig.33).

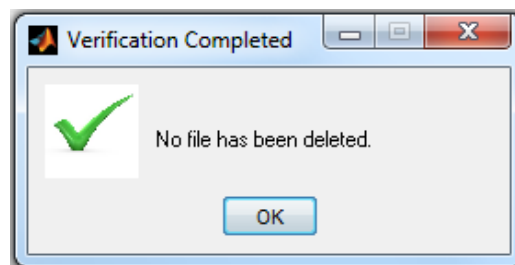


Figura 32 - Mensagem a informar o utilizador do facto de nenhum ficheiro ter sido apagado do 'BDDISK'.

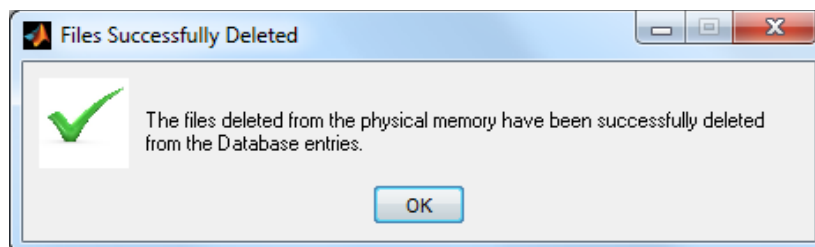


Figura 33 - Mensagem a informar o utilizador do fim da operação de remoção dos ficheiros da BD que já haviam sido removidos da memória física do sistema.

8.2.2. Funcionalidade de Consulta de Dados da Base de Dados

Passando agora à ferramenta de consulta de dados (*Consult Data*) ela inicia-se com um menu que permite ao utilizador escolher entre a realização de pesquisas, visualização de exames e tratamento estatístico (Fig. 34). Só as duas primeiras funcionalidades estão implementadas. A terceira pretende-se que seja implementada futuramente pelos prosseguidores deste projecto.

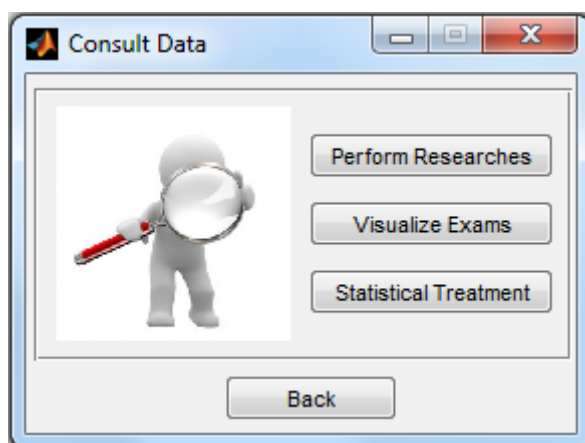


Figura 34 - Menu com as funcionalidades de consulta de dados disponibilizadas pela plataforma.

A capacidade de realização de pesquisas (*Perform Researches*) desta BD é espelhada no menu da interface representado na Fig. 35. Nesta janela começa-se por se ter uma secção para escolher se se pretende realizar uma pesquisa ao nível de todas as variáveis existentes numa dada tabela da BD, ou simplesmente de um conjunto específico delas, seleccionado pelo utilizador. Mediante a escolha de uma ou outra opção, torna-se disponível no menu o painel que permite prosseguir com a acção escolhida, e o painel correspondente à outra acção permitida torna-se inoperacional (ver figuras 35 e 36).

No caso da pesquisa de todas as variáveis de uma determinada tabela, o painel disponibilizado permite escolher qual a tabela que se pretende pesquisar. Após a selecção do *Radio Button* associado ao nome da tabela escolhida, o utilizador de premir o botão *Select*.

Se o pretendido for pesquisar apenas um conjunto de variáveis, o painel disponibilizado apresenta uma tabela com o nome e a descrição de todas as variáveis presentes na BD passíveis de serem sujeitas a pesquisa (não apresenta, por exemplo, as que guardam a informação respeitante aos utilizadores da BD, às rotinas de que ela disponibiliza, às tabelas EAV, etc.), e uma lista vazia, para onde irão os nomes das variáveis que o utilizador seleccionar, clicando nas linhas da tabela que a elas corresponda (Fig. 38). Para remover uma variável da lista, basta clicar sobre ela. Caso se tente seleccionar duas vezes a mesma variável da tabela, é apresentada uma mensagem de erro ao utilizador a informá-lo que essa já se encontra na listagem das variáveis seleccionadas para pesquisa (Fig. 37). Este tipo de mensagem é apresentado ao longo

de toda a interface para a BD, quando ocorre a situação descrita anteriormente envolvendo uma tabela e uma listagem.

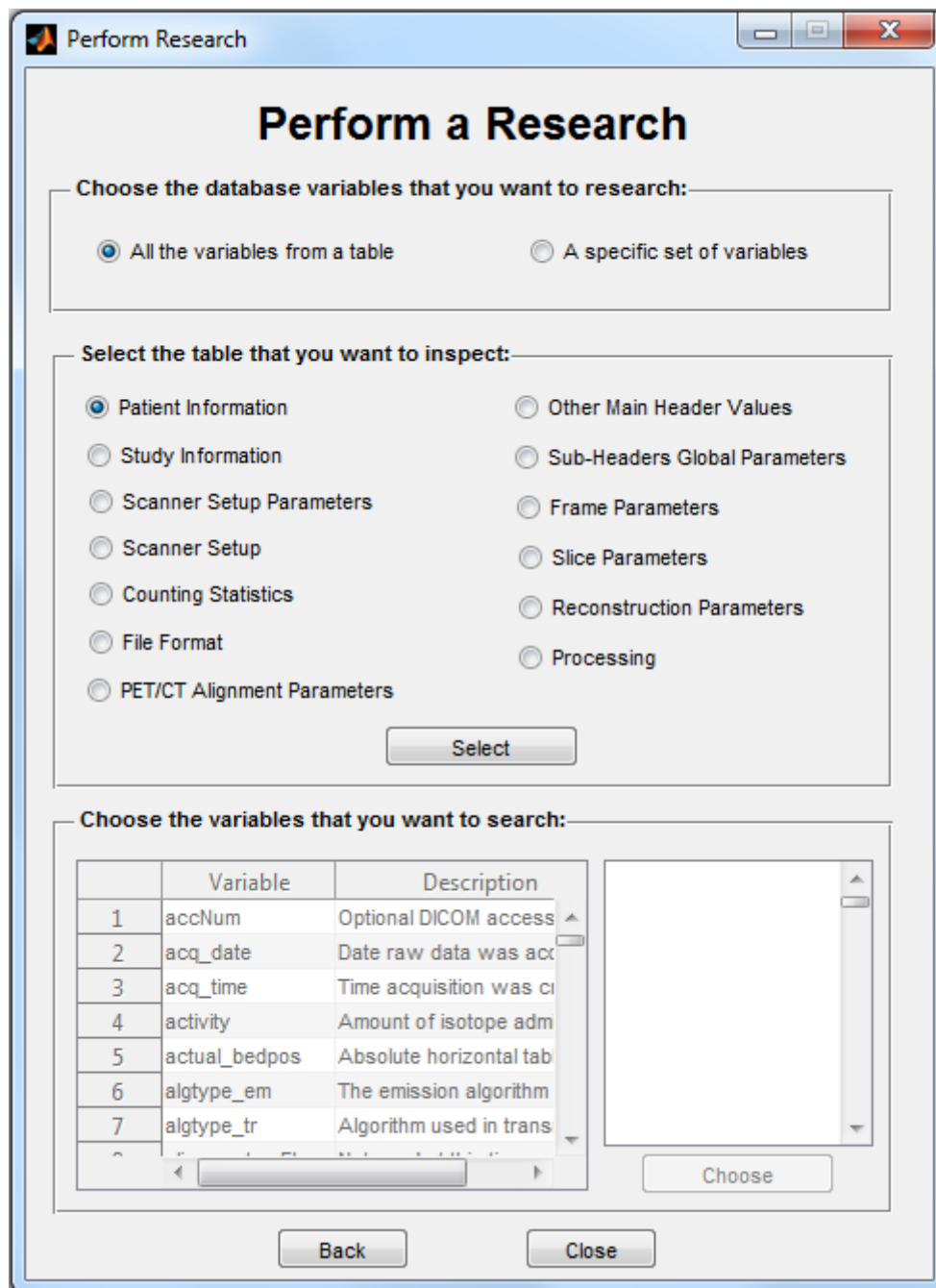


Figura 35 - Menu inicial que permite definir o tipo de pesquisa que se pretende realizar. Corresponde à opção de pesquisa de todas as variáveis associadas a uma janela.

Utilizando as variáveis seleccionadas na Fig. 38 como exemplo, vamos ver a sequência de menus desencadeada pelo tipo de pesquisa que a selecção de um conjunto de variáveis acarreta.

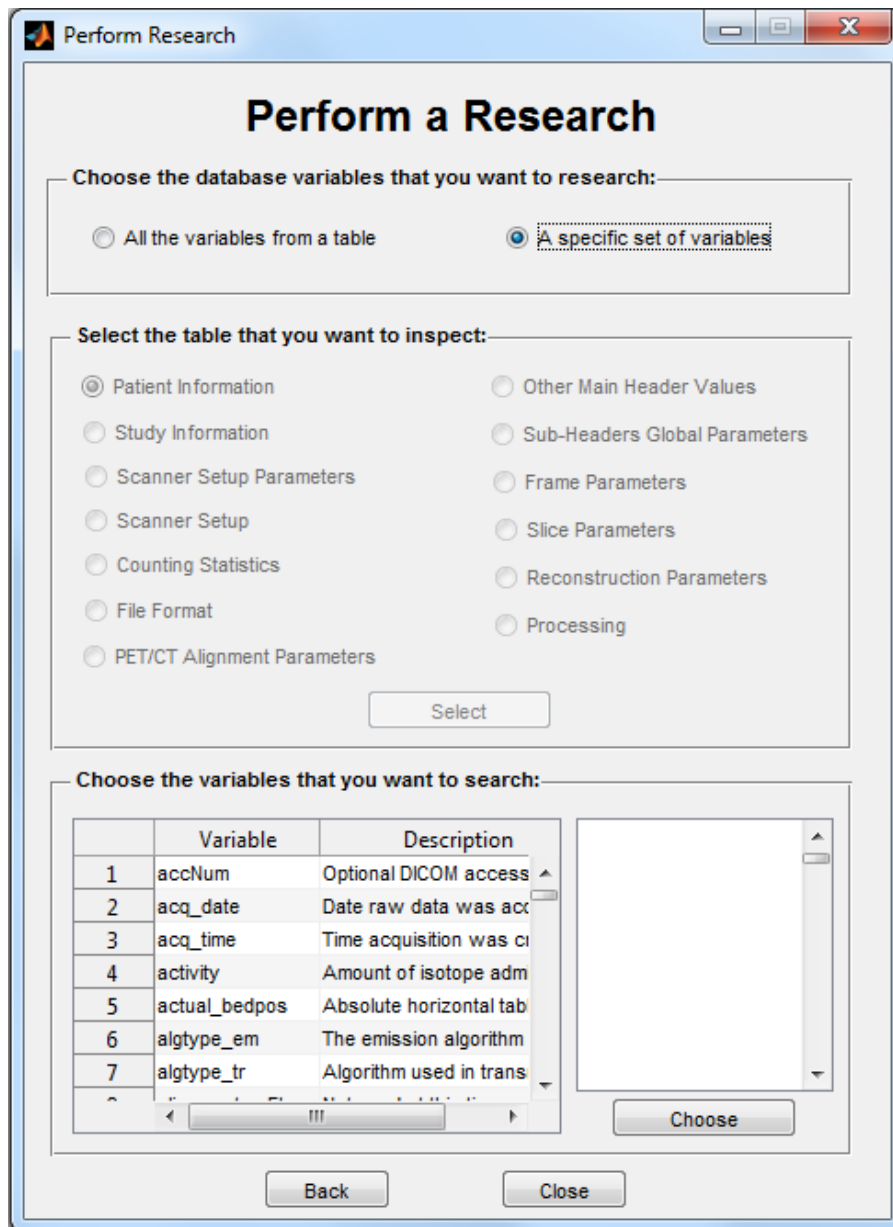


Figura 36 - Aspecto do menu quando seleccionada a escolha de um conjunto específico de variáveis.

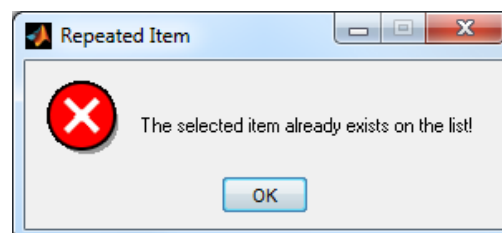


Figura 37 - Mensagem apresentada quando se selecciona de uma dada tabela, um item que já havia sido seleccionado e introduzido na lista de selecções.

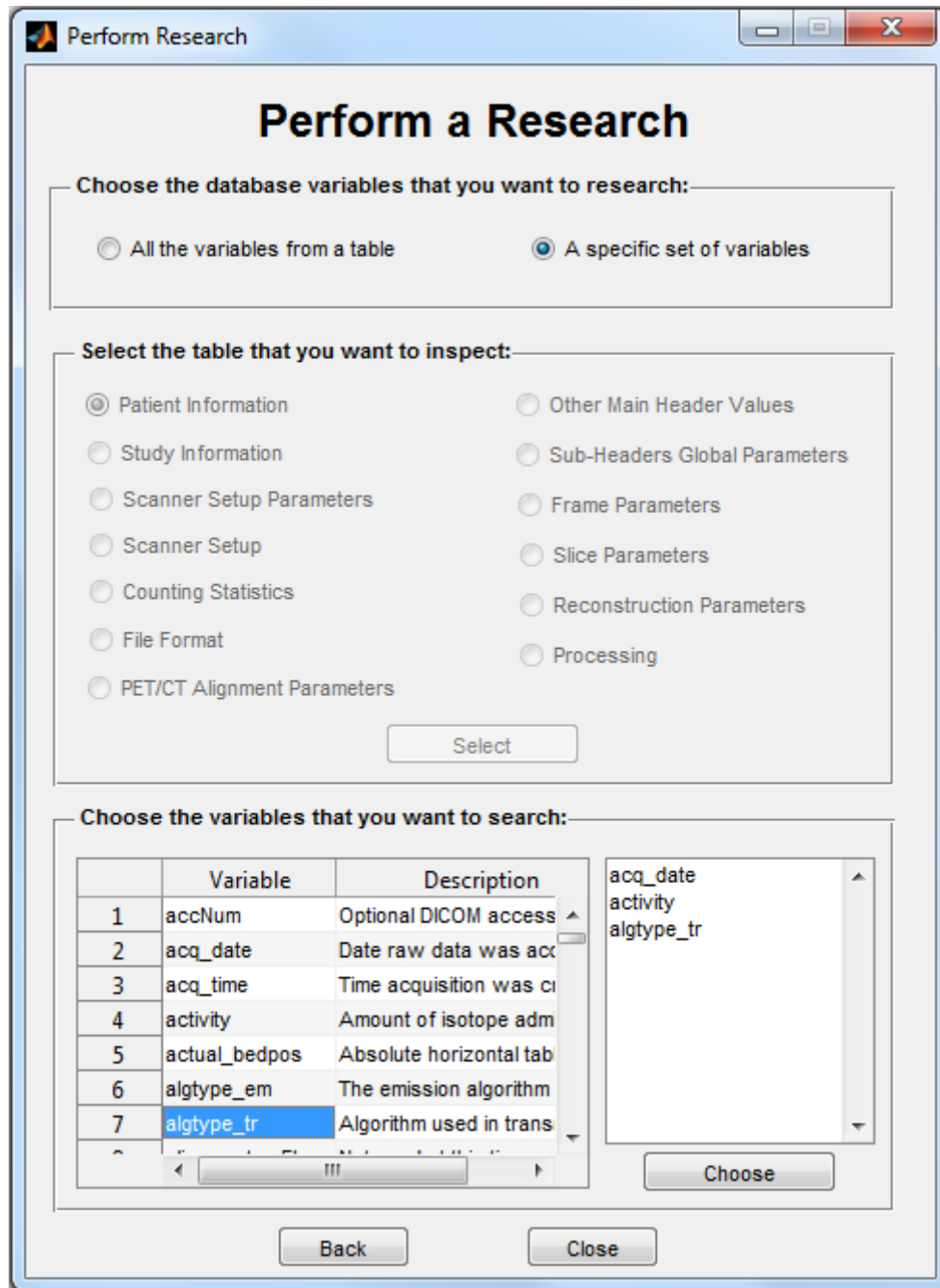


Figura 38 - Apresentação da forma de selecção de variáveis para pesquisar.

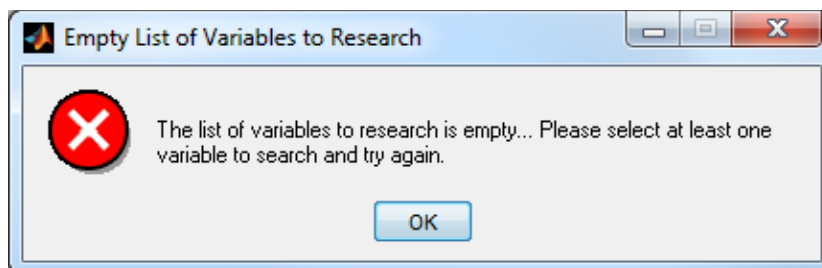


Figura 39 - Mensagem de erro apresentada quando se tenta fazer uma pesquisa de um conjunto de variáveis, sem ter seleccionado nenhuma variável das existentes na BD para o efeito.

Após se pressionar o botão *Choose*, e se a lista de variáveis seleccionadas não estiver vazia (se tal acontecer é mostrada uma mensagem de erro ao utilizador – Fig. 39), vai surgir um menu com a estrutura apresentada na Fig. 40. No primeiro painel desse menu são apresentados os nomes, os tipos, as entidades a que pertencem e as descrições de cada uma das variáveis seleccionadas para a pesquisa. No segundo painel o utilizador deve escolher se pretende ou não aplicar restrições na pesquisa das variáveis escolhidas. Caso pretenda fazê-lo, só pode restringir os valores de variáveis que pertençam às mesmas entidades das variáveis que estão a ser pesquisadas (exceptuando-se, claro, os valores dos identificadores primários dessas tabelas). Os nomes dessas variáveis que podem ser restringidas, tipos, entidades a que pertencem e descrição são apresentados na tabela da terceiro painel. No exemplo dado, na tabela do terceiro painel apenas aparecerão todas as variáveis pertencentes às entidades ‘Informação do Estudo’ e ‘Parâmetros de Reconstrução’, pois as variáveis em pesquisa (‘acqdate’, ‘activity’ e ‘alctype_tr’) fazem parte destas. O modo de selecção das variáveis da tabela para a lista de itens seleccionados é o mesmo do descrito no menu anterior (Fig. 41). No quarto painel é proporcionada ao utilizador a possibilidade de escolher a variável pela qual deseja ordenar os resultados, e o tipo de ordem (ascendente ou descendente) que pretende. Definiu-se que os resultados apenas podem ser ordenados segundo uma das variáveis que está a ser pesquisada, pelo que no menu *Pop-up* só aparecem os nomes das variáveis que estão a ser pesquisadas (Fig. 41). Por fim, existe uma opção que o utilizador pode seleccionar caso queira que os resultados da sua pesquisa sejam apresentados na linha de comandos do MATLAB®.

Após se escolher o tipo de restrições que se pretende implementar à pesquisa, deve pressionar-se o botão *Search*.

Caso se tenha escolhido não implementar qualquer restrição na pesquisa, serão apresentados apenas os resultados com identificadores comuns das variáveis seleccionadas. Como este pode ser um processo bastante demorado, devido à enorme quantidade de meta-informação armazenada na BD, é mostrada uma barra de espera ao utilizador para lhe dar conta da evolução do processo (Fig. 42). No final da pesquisa, os resultados são armazenados num documento no formato ‘.txt’, estruturado do seguinte modo: um cabeçalho de quatro linhas, em que na primeira são apresentados os nomes das variáveis pesquisadas, na segunda as restrições que foram implementadas à pesquisa (neste caso nenhuma), na terceira a variável que foi utilizada para a organização dos resultados e o tipo de ordem escolhido, e na quarta o número de ocorrências encontradas com as características definidas. De seguida são apresentados os resultados propriamente ditos, onde a primeira linha apresenta os nomes das variáveis pesquisadas, e nas restantes, nas colunas correspondentes ao nome adequado, os valores recolhidos para essas variáveis (será dado um exemplo de um ficheiro deste tipo mais à frente).

Caso se tenha optado por implementar restrições, e estando a lista de variáveis a restringir devidamente preenchida (caso não esteja é apresentada uma mensagem semelhante à

da Fig. 39), o processo será o mesmo, mas com a adição da escolha dos valores permitidos para esta pesquisa a cada uma das variáveis da lista (a forma de selecção será apresentada mais à frente neste capítulo).

Define Characteristics

Define the Selected Variables Characteristics to Perform the Research

List of variables whose values you want to research:

	Variable	Type	Entity	Description
1	acq_date	int	study_inform...	Date raw data was acquired
2	activity	float	study_inform...	Amount of isotope administered
3	algtype_tr	short	reconstructio...	Algorithm used in transmission

Do you want to restrict the values of the variables that will be shown?

No Yes

Please select the variables whose values you want to restrict:

	Variable	Type	Entity	Description
1	accnum	char[16]	study_info...	Optional DI
2	acq_date	int	study_info...	Date raw c
3	acq_time	long	study_info...	Time acqui
4	activity	float	study_info...	Amount of
5	algtype_em	short	reconstruc...	The emissi
6	algtype_tr	short	reconstruc...	Algorithm u
7	analy_swrel	char[7]	reconstruc...	Version of

Please select the variable by which you want to sort the results:

acq_date

Ascending Order Descending Order

Show the Search Results on the MATLAB Line of Comands

Search Back

Figura 40 - Menu que permite a definição de restrições para a pesquisa de um conjunto de variáveis.

Define Characteristics

Define the Selected Variables Characteristics to Perform the Research

List of variables whose values you want to research:

	Variable	Type	Entity	Description
1	acq_date	int	study_inform...	Date raw data was acquired
2	activity	float	study_inform...	Amount of isotope administe
3	alctype_tr	short	reconstructio...	Algorithm used in transmissi

Do you want to restrict the values of the variables that will be shown?

No Yes

Please select the variables whose values you want to restrict:

	Variable	Type	Entity	Descript
1	accnum	char[16]	study_info...	Optional DI
2	acq_date	int	study_info...	Date raw c
3	acq_time	long	study_info...	Time acqui
4	activity	float	study_info...	Amount of
5	alctype_em	short	reconstruc...	The emissi
6	alctype_tr	short	reconstruc...	Algorithm u
7	analy_swrel	char[7]	reconstruc...	Version of

3) acq_time
5) alctype_em
7) analy_swrel

Please select the variable by which you want to sort the results:

acq_date
acq_date
activity
alctype_tr

Show the Search Results on the MATLAB Line of Comands

Figura 41 - Menu que permite a definição de restrições para a pesquisa de um conjunto de variáveis com ilustração de dois pormenores: modo de selecção de itens da tabela para a lista e variáveis que podem ser escolhidas para ordenar os resultados.

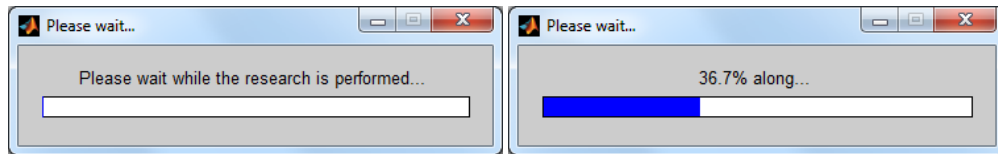


Figura 42 - Barra de espera implementada para a monitorização do processo de pesquisa.

Vamos agora ver o modo de funcionamento da pesquisa de todas as variáveis associadas a uma dada tabela. Para tal, e por esta ter diversos tipos de dados, escolheu-se a tabela ‘Parâmetros de Configuração do *Scanner*’ como exemplo (Fig. 43).

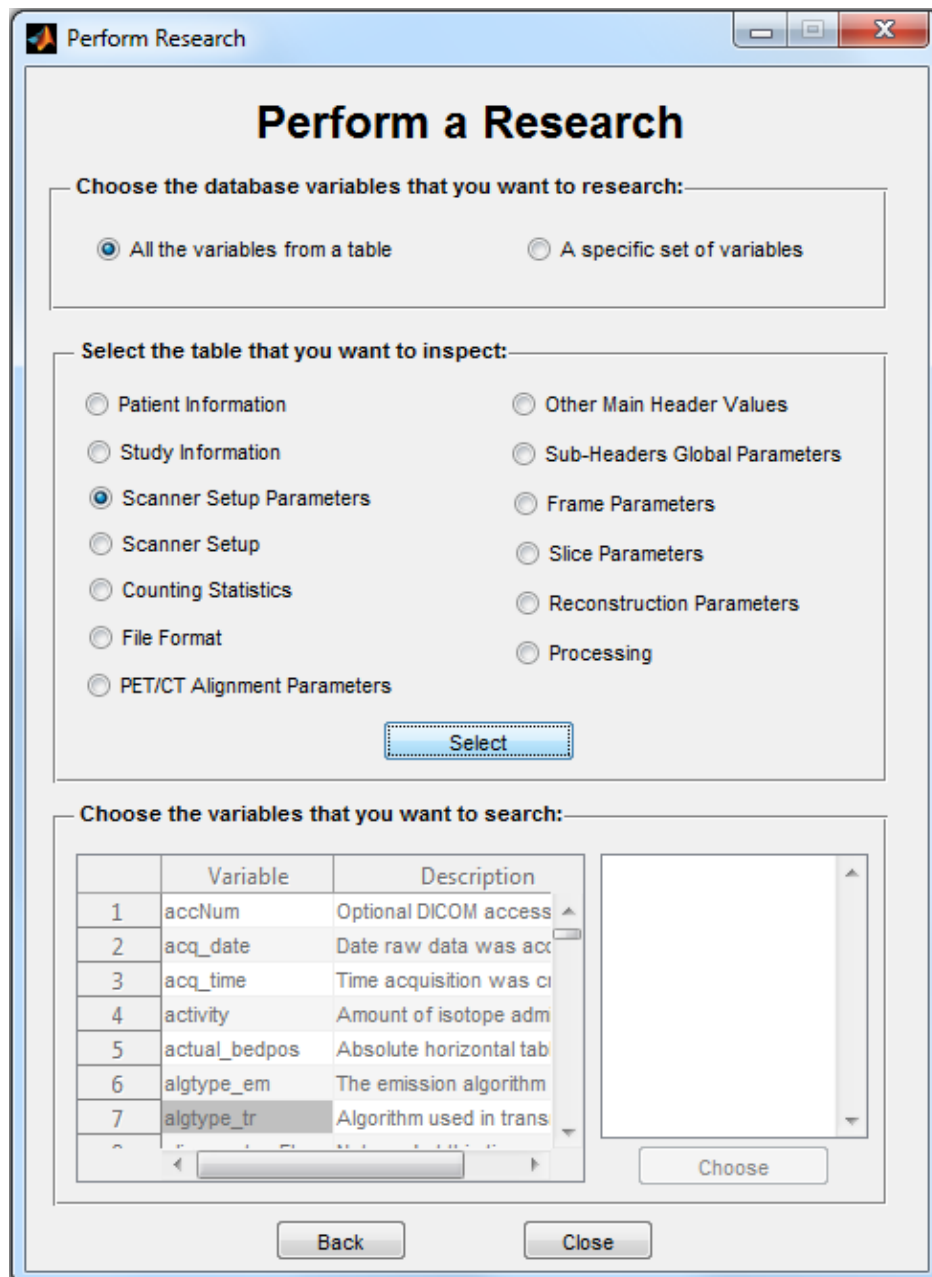


Figura 43 - Seleção da tabela ‘Parâmetros de Configuração do *Scanner*’ para efectuar a pesquisa de todas as variáveis nela existentes.

Após a selecção, surge um menu que permite escolher se se pretendem aplicar restrições a esta pesquisa ou não (primeiro painel do menu apresentado na Fig. 44). Se não se pretender tal, o utilizador apenas precisa de accionar o botão *Search*, após ter escolhido a variável da tabela segundo a qual pretende organizar os resultados (atenção que aqui está incluída também a variável (ou variáveis) de identificação primária da tabela em questão – Fig. 45). Assim que a pesquisa estiver completa, é mostrada uma mensagem de sucesso ao utilizador, semelhante à mostrada na Fig. 46, que inclui também o nome com que os resultados foram armazenados (que se definiu como sendo a expressão ‘SearchResult’ seguida da data e da hora de geração do ficheiro) e a directoria onde o ficheiro ‘.txt’ se encontra.

	Variable	Description
1	angmax	Maximum axial acceptanc
2	axialFov	Useful axial length of the
3	backLeadDiamet...	The diameter of the back
4	crbTstampPeriod	Period in msec between C
5	crystalThickness	Detector crystal thickness
6	delayType	Type of delayed event us
7	detectorRadius	Inscribed scanner radius

Figura 44 - Menu que permite escolher entre a implementação ou não de restrições à pesquisa de todas as variáveis existentes numa tabela, e caso se pretenda restringir os valores pesquisados, permite a selecção das variáveis presentes nessa tabela às quais se pretende aplicar as restrições.

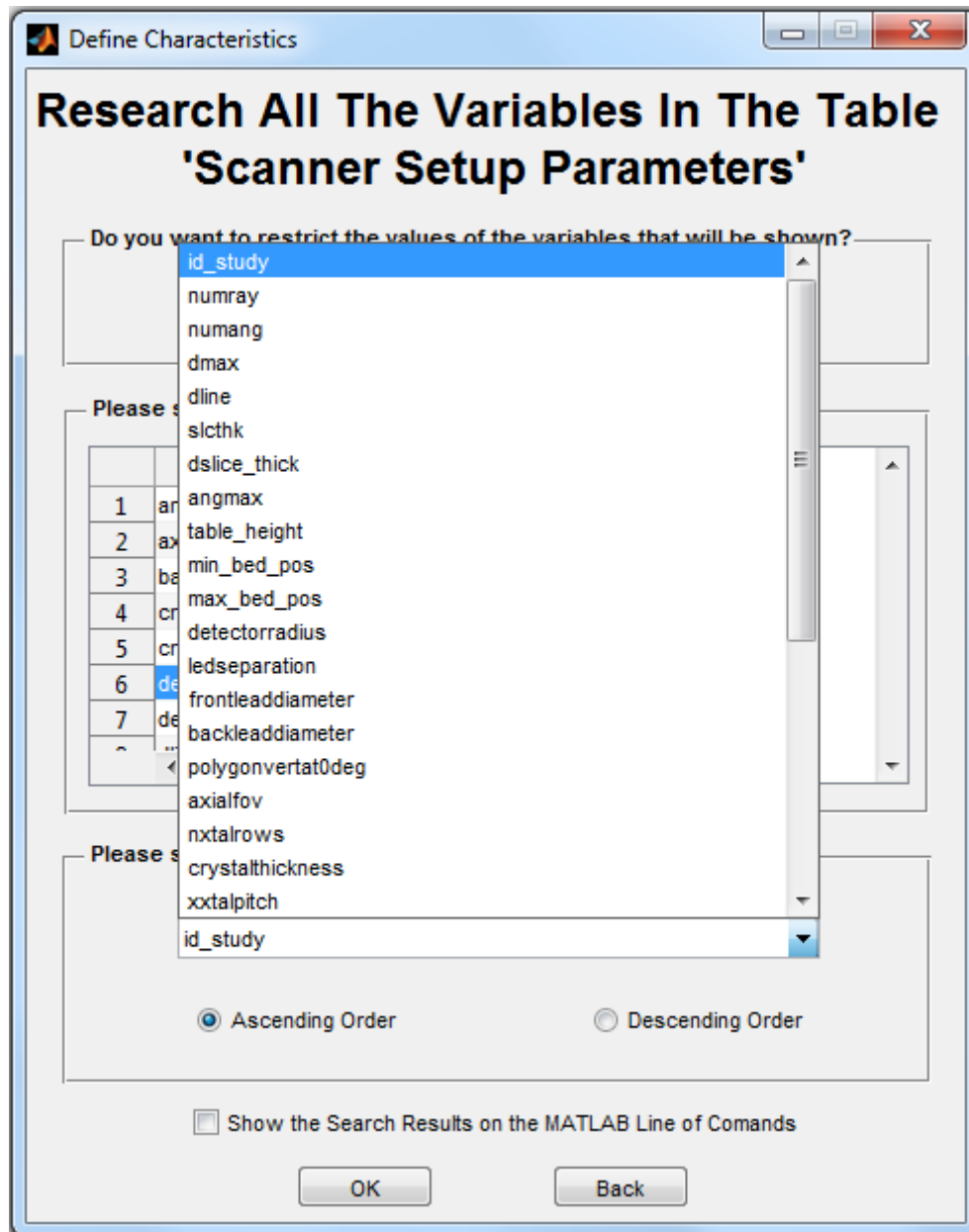


Figura 45 - Menu de aplicação de restrições, mostrando o pormenor das variáveis disponíveis para organização dos resultados.

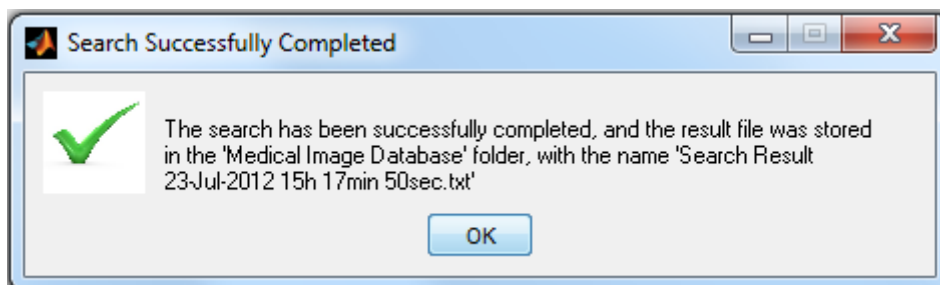


Figura 46 - Mensagem de notificação do fim da pesquisa, bem como do nome do ficheiro e da directoria onde esta foi armazenada.

Na Fig. 47 apresenta-se o resultado da pesquisa efectuada mediante o exemplo apresentado.

Search Result 23-Jul-2012 15h 54min 17sec.txt - Bloco de notas

Fichero Editar Formatar Ver Ajuda

--> Researched variables: all the variables in the table 'Scanner Setup Parameters'.
 --> defined characteristics: no restrictions have been applied on this search.
 --> order of the Results Presentation: results presented in the ascendent order, ordered by the variable ''.
 --> Number of occurrences Found: 10236.

id_study	numray	numang	dmax	dline	slcthk	dslice_thick	anamax	table_height	min_bed_pos	max_bed_pos	detectorradius	ledseparation
3	128	128	256	2	2	10	0	547	431.94999999999999	179.80000000000001	711.70000000000005	
4	128	128	256	2	0	10	0	0	193.00999999999999	609.60000000000002	588.7699999999	
5	144	144	576	4	4	10	0	313	431.94999999999999	179.80000000000001	711.70000000000005	
6	144	144	576	4	4	10	0	313	431.94999999999999	179.80000000000001	711.70000000000005	
7	144	144	576	4	4	10	0	313	431.94999999999999	179.80000000000001	711.70000000000005	
8	144	144	576	4	4	0	185	6	201	0	0	0
9	144	144	576	4	4	0	185	6	201	0	0	0
10	144	144	576	4	4	10	0	889	889	431.94999999999999	179.80000000000001	711.70000000000005
11	144	144	576	4	4	10	0	889	889	431.94999999999999	179.80000000000001	711.70000000000005
12	144	144	576	4	4	10	0	889	889	431.94999999999999	179.80000000000001	711.70000000000005
13	144	144	576	4	4	0	185	-3.5	191.5	0	0	0
14	144	144	576	4	4	10	0	271	271	431.94999999999999	179.80000000000001	711.70000000000005
15	144	144	576	4	4	10	0	271	271	431.94999999999999	179.80000000000001	711.70000000000005
16	144	144	576	4	4	10	0	271	271	431.94999999999999	179.80000000000001	711.70000000000005
17	144	144	576	4	4	10	0	271	271	431.94999999999999	179.80000000000001	711.70000000000005
18	144	144	576	4	4	0	185	39.5	224.5	0	0	0
19	144	144	576	4	4	0	185	39.5	224.5	0	0	0
20	144	144	576	4	4	10	0	271	271	431.94999999999999	179.80000000000001	711.70000000000005
21	144	144	576	4	4	10	0	271	271	431.94999999999999	179.80000000000001	711.70000000000005
22	144	144	576	4	4	10	0	271	271	431.94999999999999	179.80000000000001	711.70000000000005
23	144	144	576	4	4	0	185	39.5	224.5	0	0	0
24	144	144	576	4	4	0	185	39.5	224.5	0	0	0
25	144	144	576	4	4	10	0	277	277	431.94999999999999	179.80000000000001	711.70000000000005
26	144	144	576	4	4	10	0	277	277	431.94999999999999	179.80000000000001	711.70000000000005
27	144	144	576	4	4	10	0	277	277	431.94999999999999	179.80000000000001	711.70000000000005
28	144	144	576	4	4	0	185	36.5	216.5	0	0	0
29	144	144	576	4	4	0	185	36.5	216.5	0	0	0
30	128	128	256	2	2	10	0	284	284	431.94999999999999	179.80000000000001	711.70000000000005
31	128	128	256	2	2	10	0	284	284	431.94999999999999	179.80000000000001	711.70000000000005
32	144	144	576	4	4	0	185	26.5	211.5	0	0	0
33	144	144	576	4	4	0	185	26.5	211.5	0	0	0
34	144	144	576	4	4	10	0	1220	1892	431.94999999999999	179.80000000000001	711.70000000000005
35	144	144	576	4	4	10	0	1220	1892	431.94999999999999	179.80000000000001	711.70000000000005
36	144	144	576	4	4	10	0	1220	1892	431.94999999999999	179.80000000000001	711.70000000000005
37	144	144	576	4	4	0	185	-1594.5	-689.5	0	0	0
38	144	144	576	4	4	0	185	-1594.5	-689.5	0	0	0
39	144	144	576	4	4	10	0	488	2000	431.94999999999999	179.80000000000001	711.70000000000005
40	144	144	576	4	4	10	0	488	2000	431.94999999999999	179.80000000000001	711.70000000000005
41	144	144	576	4	4	10	0	488	2000	431.94999999999999	179.80000000000001	711.70000000000005
42	144	144	576	4	4	0	185	-1714	6	0	0	0
43	144	144	576	4	4	0	185	-1714	6	0	0	0
44	144	144	576	4	4	10	0	274	274	431.94999999999999	179.80000000000001	711.70000000000005
45	144	144	576	4	4	0	185	40	217	0	0	0

Figura 47 - Resultado da pesquisa efectuada a todas as variáveis da tabela 'Parâmetros de Configuração do Scanner' sem aplicar qualquer tipo de restrição à procura.

No caso de se seleccionar a opção de aplicar restrições às variáveis da tabela escolhida, o segundo painel do menu torna-se operacional, e o utilizador deve colocar as variáveis que pretende restringir na lista para o efeito (Fig. 48). Após o preenchimento da lista deve premir o botão *OK*. Esta acção desencadeia o aparecimento de um menu para restringir os valores de cada uma das variáveis seleccionadas. O menu que aparece depende do tipo de dados que a variável em questão engloba, porque obviamente é necessário aplicar diferentes tipos de restrições para uma variável de caracteres comparativamente a uma numérica. Neste sentido foram criados três tipos de menus para restrição de variáveis: o primeiro para variáveis do tipo 'int' (*Integer*), 'long' (*Long Integer*) e 'float'; o segundo para variáveis do tipo 'char' (*Characters*); e o terceiro para variáveis do tipo 'short' (*Short Integer* - este tipo foi separado dos restantes tipos numéricos porque muitas vezes representa variáveis do tipo nominal, tendo cada um dos seus elementos um significado lógico). Vamos pegar nas variáveis seleccionadas na Fig. 48 para exemplificar este processo de restrição.

Como a variável 'axialFov' é do tipo 'float' é mostrado o primeiro tipo de menu desenvolvido, representado na Fig. 49. Dentro deste menu são permitidos dois tipos de restrição (apresentados para escolha no primeiro painel do menu): no segundo painel pode-se escolher que os valores da variável sejam maiores, menores ou iguais a um determinado número inserido pelo utilizador (ver Fig. 49); no terceiro painel o utilizador pode introduzir dois valores entre os

quais deseja que os valores das variáveis se encontrem (ver Fig. 51). Para auxiliar o utilizador nesta tarefa, existe um quarto painel onde são mostrados os valores mínimo e máximo armazenados na BD para a variável em questão. No caso de mesmo assim o utilizador não respeitar os limites estabelecidos, e pressionar o botão *OK* para prosseguir com a operação, são lhe mostradas mensagens de erro como as representadas nas figuras 50 e 52.

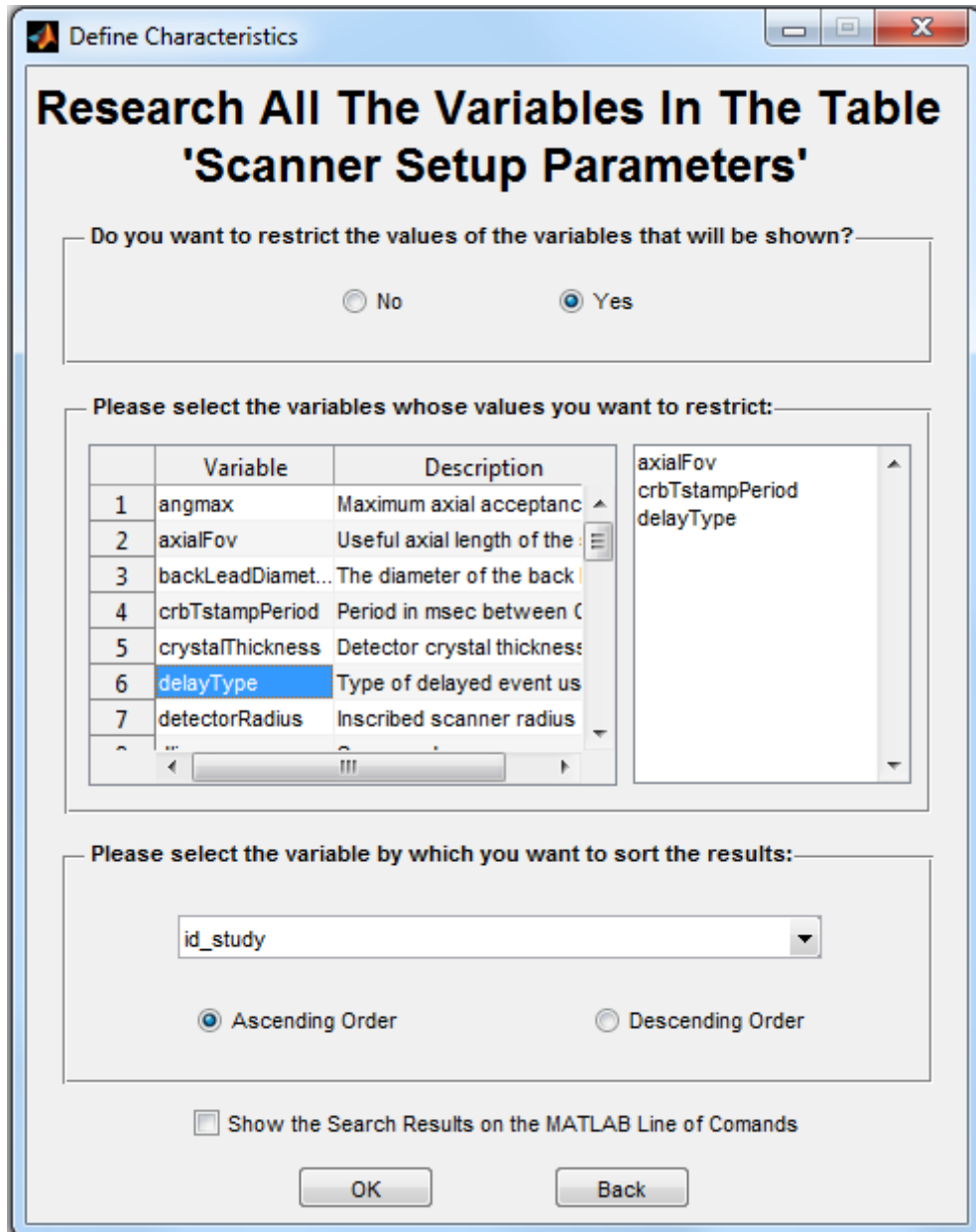


Figura 48 - Menu de definição das variáveis cujos valores se pretende restringir para a realização da pesquisa, especificando o modo de selecção das variáveis a restringir (preenchimento da lista).

Figura 49 - Tipo de menu desenhado para restringir os valores a serem pesquisados para variáveis do tipo *float*, *int* e *long*. Pode-se ver o pormenor dos sinais lógicos permitidos para o primeiro tipo de restrição a aplicar.

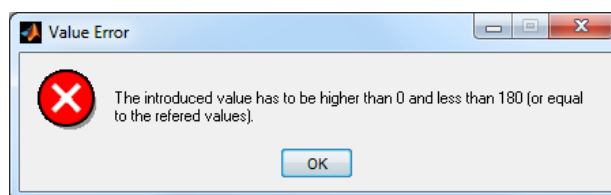


Figura 50 - Mensagem de erro apresentada quando o utilizador insere no segundo painel um valor que não esteja entre o intervalo definido pelos valores mínimo e máximo para a variável em questão, ou então que não seja igual a esses.

Figura 51 - Tipo de menu desenhado para restringir os valores a serem pesquisados para variáveis do tipo *float*, *int* e *long*, mostrando a activação do terceiro painel mediante a escolha do utilizador de 'procura entre dois valores'.

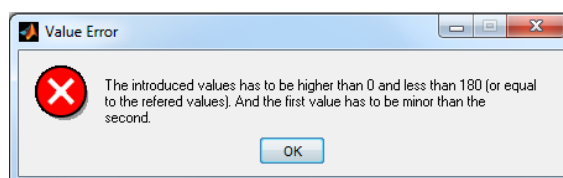


Figura 52 - Mensagem de erro apresentada quando o utilizador insere no terceiro painel um valor que não esteja entre o intervalo definido pelos valores mínimo e máximo para a variável em questão, ou então que não seja igual a esses, e caso o valor definido como mínimo não seja menor que o máximo.

Definidas as restrições para a primeira variável, prossegue-se para os critérios da segunda. Sendo a 'csbTstampPeriod' do tipo 'short', é apresentado o terceiro tipo de menu descrito, cuja estrutura é a da Fig. 53. Este tipo de menu disponibiliza três tipos de restrições, a selecção da que se pretende implementar deve ser feita ao nível do primeiro painel do menu (conforme a seleccionada, será colocado em modo operacional o painel correspondente). As duas últimas (correspondentes aos painéis 4 e 5 do menu) são semelhantes às já apresentadas no menu anterior, e servem para o caso de se tratar simplesmente de um número inteiro curto (sem qualquer associação lógica aos seus valores). Devido ao significado que pode estar associado a variáveis deste tipo, decidiu-se implementar uma forma do utilizador ver o seu significado e dos valores lógicos que lhe estejam associados, para poder escolher aqueles que são da sua conveniência, como se pode ver no segundo painel do menu.

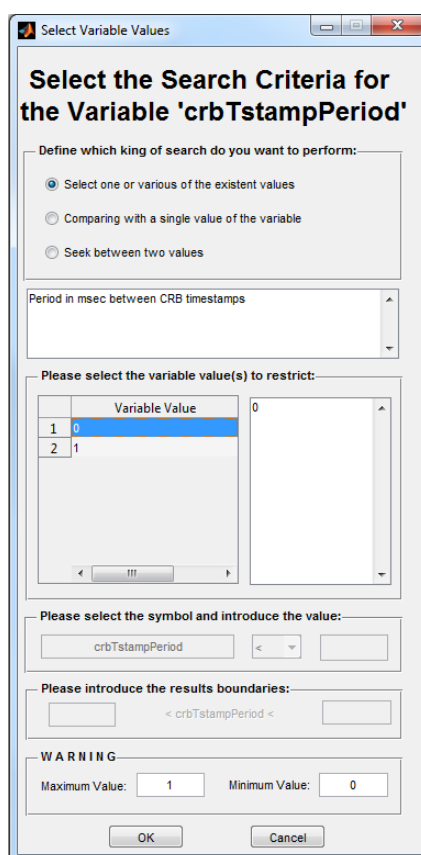


Figura 53 - Menu de implementação de restrições para variáveis do tipo *short* (que podem referir-se a valores lógicos).

No terceiro painel é apresentada uma tabela com todos os valores disponíveis, e a partir desta o utilizador pode seleccionar os que quer que façam parte da pesquisa para a lista ao lado (modo de selecção semelhante ao já explicado). Após esta definição, haveria ainda a implementação de restrições à variável 'delayType', mas devido a esta ser do tipo 'short', seria também mostrado o terceiro tipo de menu. No final da definição de restrições a todas as variáveis seleccionadas, o operador deve fazer *OK* para proceder à pesquisa por si definida anteriormente, e no final é-lhe mostrada uma mensagem semelhante à da Fig. 46 a dar conta do sucesso da pesquisa. Caso o utilizador interrompa a operação de definição das variáveis, pressionando o botão *Cancel* num dos menus anteriores, é-lhe mostrada uma mensagem de erro como a representada na Fig. 54.

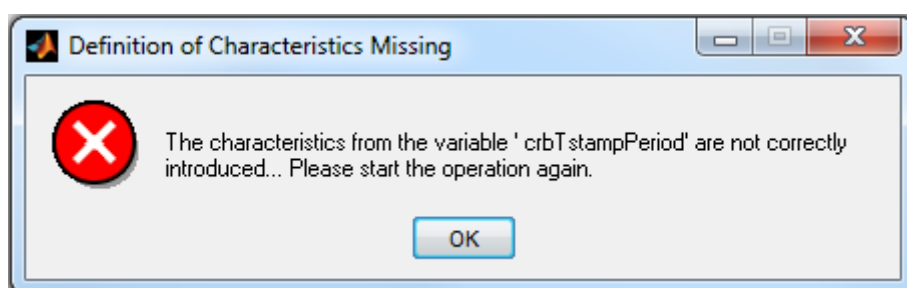
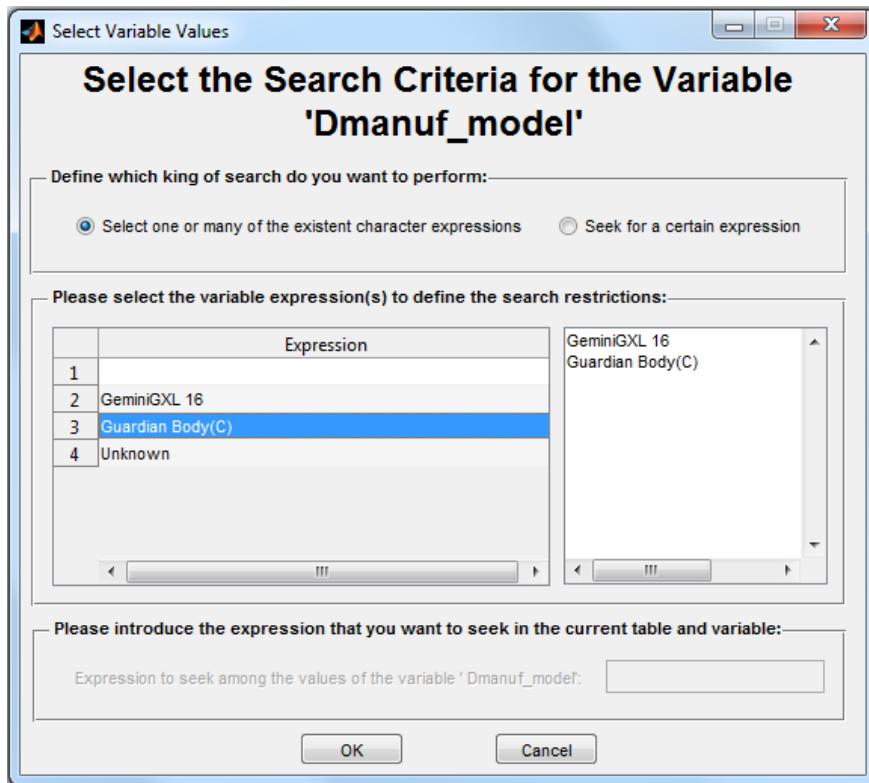
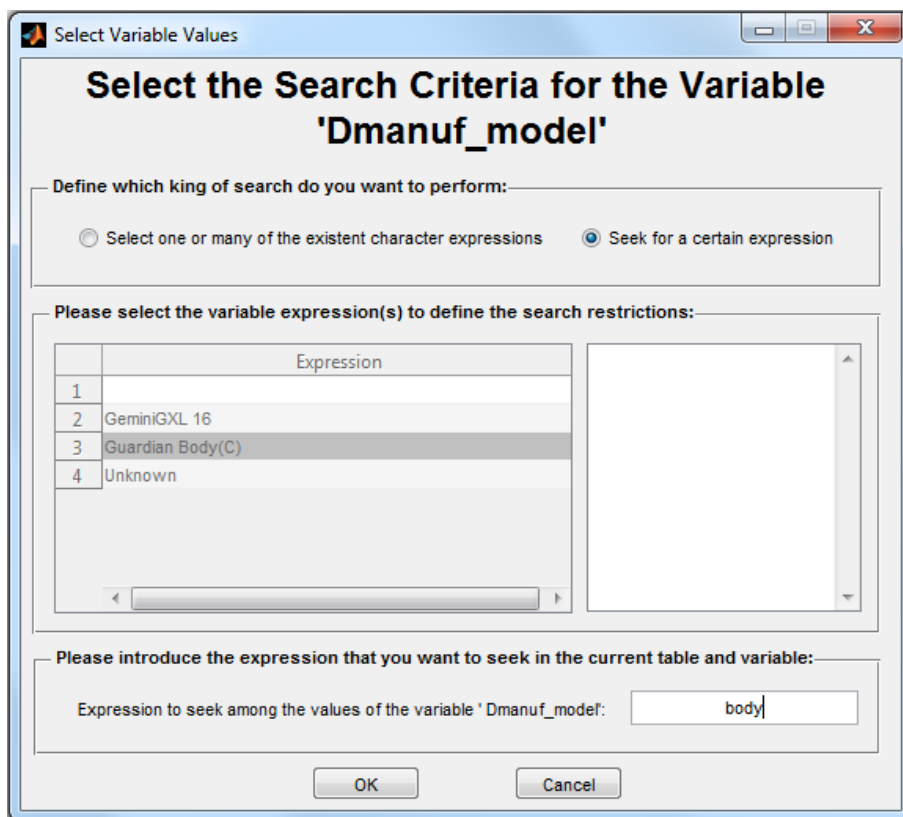


Figura 54 - Tipo de mensagem de erro apresentada caso o utilizador interrompa a operação de definição das restrições a aplicar à pesquisa a meio. No exemplo, seria a mensagem mostrada caso o utilizador premisse o botão *Cancel* no menu de implementação de restrições da variável 'crbTstampPeriod'.

Faltou no entanto ver o segundo tipo de menus de implementação de restrições, dedicado a variáveis do tipo 'char'. Como não existem variáveis desse tipo na tabela seleccionada para o exemplo, fez-se o mesmo tipo de pesquisa para a tabela 'Configuração do Scanner' e seleccionou-se uma variável de caracteres para a implementação de restrições (por ser um menu semelhante a outros já apresentados, ele foi colocado em apêndice a este documento, no Anexo I.1). O menu que surge neste caso é apresentado na Fig. 55. Este disponibiliza a selecção de um ou vários dos valores disponíveis para a variável em questão (painel 2), ou a procura de uma dada expressão (conjunto de caracteres introduzidos pelo utilizador) nos valores dessa variável (Fig. 56). De notar que esta última funcionalidade foi implementada de forma a pesquisar a expressão tal e qual como for inserida pelo utilizador, toda em minúsculas, toda em maiúsculas, e com a primeira letra em maiúscula e as restantes minúsculas (para prevenir situações de diferentes introduções dos dados no que respeita à forma de escrita de maiúsculas e minúsculas).

Figura 55 - Menu para restrição dos valores de variáveis do tipo *char*.Figura 56 - Menu para restrição dos valores de variáveis do tipo *char*, mostrando a ferramenta de pesquisa de uma dada expressão dentro dos valores da variável.

Apresentadas as ferramentas de pesquisa, vai agora proceder-se à descrição dos métodos de visualização de exames. Ao seleccionar o botão *Exams Visualization* no menu *Consult Data*, surge então uma janela onde são listados todos os tipos de visualizadores implementados na plataforma até ao momento (Fig. 57).

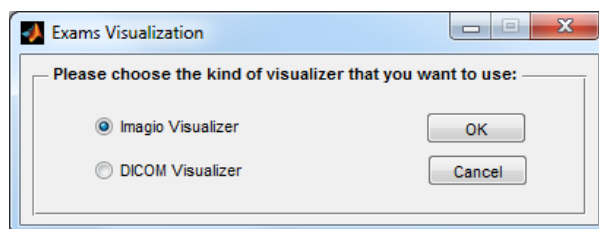


Figura 57 - Menu apresentando o tipo de visualizadores de que a BD disponibiliza até ao momento.

Para já, apenas existem dois: um visualizador de ficheiros de imagem no formato *imagio* e outro de ficheiros de imagem no formato DICOM. Após o operador seleccionar o que lhe interessa utilizar (note-se que, por enquanto, o de DICOM na maior parte das vezes apenas permite a visualização das imagens de TAC, que geralmente são as únicas com representação neste formato armazenadas nas directorias dos exames) deve premir *OK*. Se tiver seleccionado o formato *imagio* ser-lhe-á então mostrada uma janela com a listagem de todos os exames existentes na BD (Fig. 58), se, pelo contrário, se escolheu o visualizador DICOM aparecerá uma lista com as directorias de armazenamento dos ficheiros de imagem (Fig. 59). De entre as entradas de uma destas listas o utilizador deve seleccionar o item que lhe interessa visualizar e pressionar *Select*. Note-se que: é necessário que se estabeleça à conexão ao dispositivo de memória onde os exames estão armazenados antes de tentar visualizá-los (caso não se leve a cabo esta acção, aparecerá uma mensagem de erro semelhante à mostrada na Fig. 60); até ao momento, apenas se permite a visualização de um exame de cada vez, podendo no entanto visualizar-se todas as secções desse exame ao mesmo tempo no visualizador *imagio* (o de DICOM permite a navegação entre estas, mas não a visualização em simultâneo).

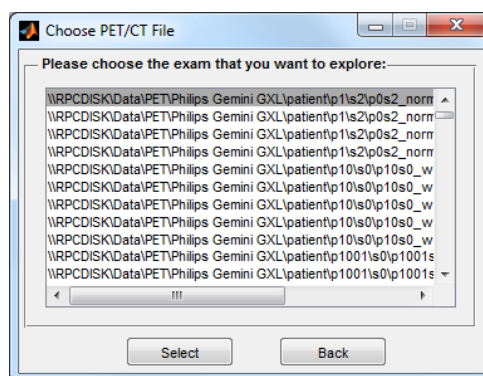


Figura 58 - Janela com a listagem de todos os exames presentes na BD (as dimensões do menu foram alteradas).

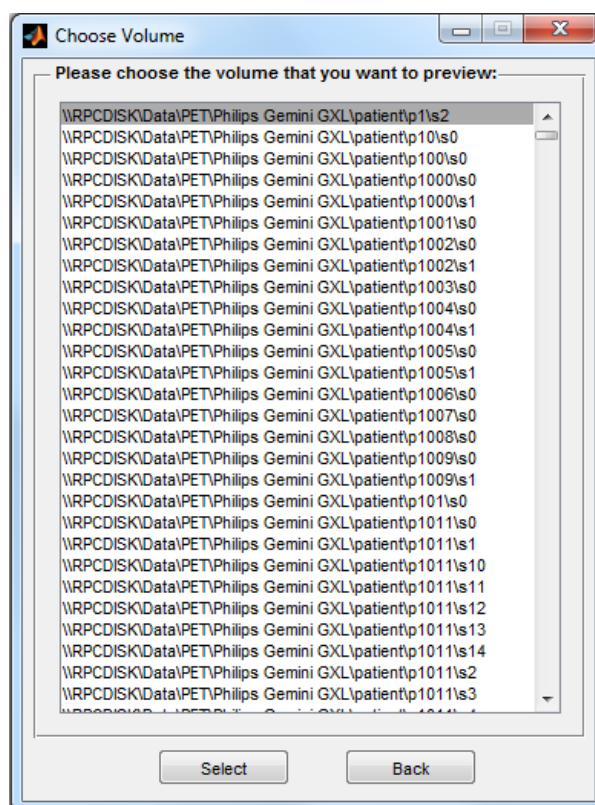


Figura 59 - Janela com a listagem das directorias de pastas de armazenamento de ficheiros de imagem presentes na BD.

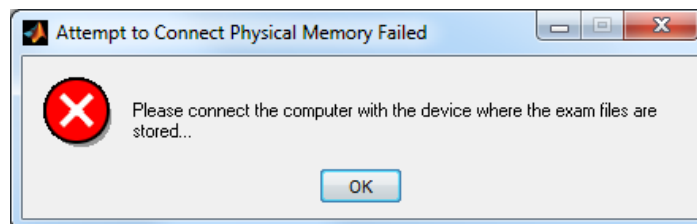


Figura 60 - Mensagem de erro notificando o utilizador da necessidade de estabelecer o acesso ao 'BDDISK' antes de tentar visualizar os ficheiros.

Se a escolha recaiu sobre o visualizador *imago*, e se o ficheiro '.img' tiver o respectivo '.syn' inserido nas entradas da BD (caso contrário será mostrada a mensagem de erros da Fig. 61) será aberta uma ligação à máquina virtual do IDL, e visualizar-se-á a interface gráfica mostrada na Fig. 62.

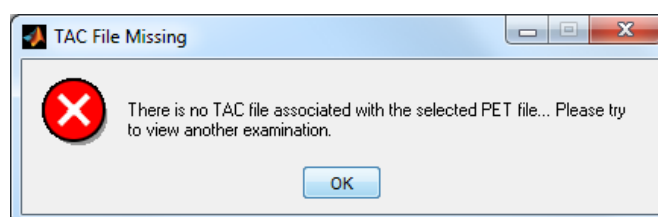


Figura 61 - Mensagem de erro apresentada caso o ficheiro de TAC associado ao '.img' seleccionado não seja encontrado nos dados imagiológicos constantes das entradas da BD.

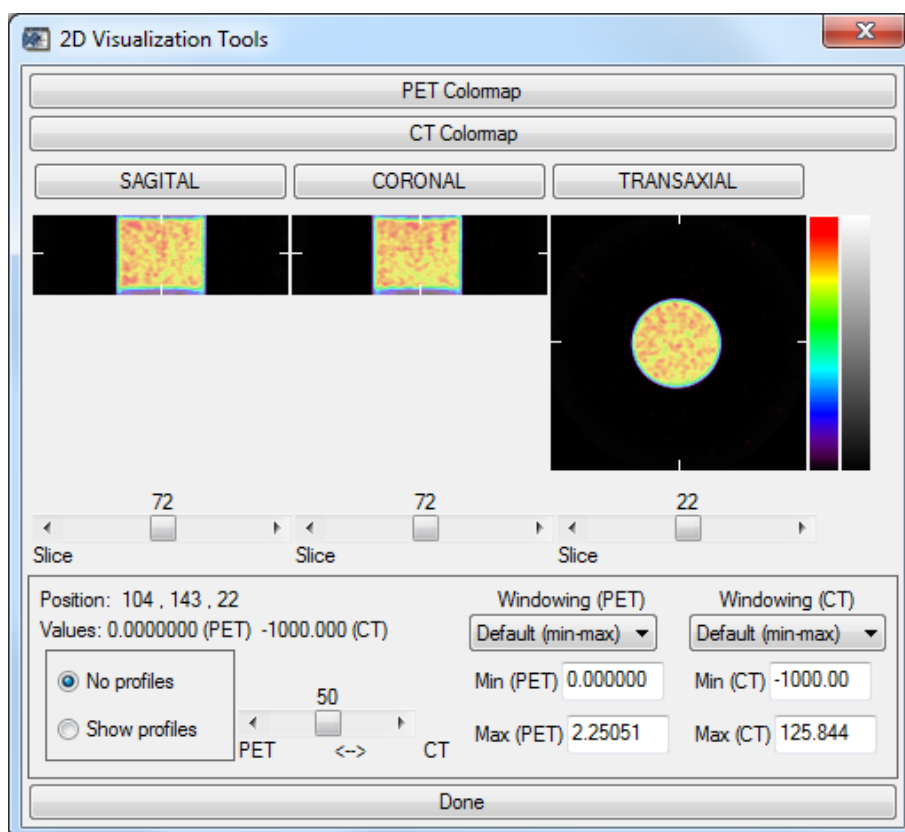


Figura 62 - Interface de visualização dos ficheiros *imago* implementada em IDL®.

Informação mais detalhada sobre o visualizador desenvolvido em IDL® é apresentada no Anexo I.2 desta tese.

No caso de se ter escolhido visualizar os ficheiros DICOM, será aberto o menu de visualização representado na Fig. 64 em *background* (se o botão *Back* do menu de escolha do volume for premido, em vez de se proceder à selecção de um dos itens da lista, é apresentada uma mensagem de erro ao utilizador, que é mostrada na Fig. 63), onde deve premir-se o botão direito do rato para porcerder ao carregamento dos ficheiros DICOM presentes na directoria seleccionada previamente (menu de escolha mostrado na Fig. 64). Caso não existam ficheiros neste formato dentro dessa pasta, será apresentada a mensagem de erro mostrada na Fig. 65, e a aplicação de visualização de exames deverá ser encerrada, e proceder-se à escolha de uma outra directoria na listagem apresentada.

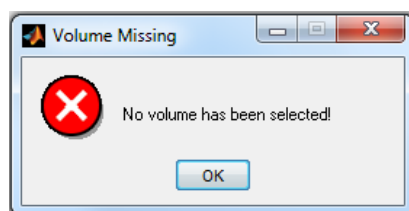


Figura 63 - Mensagem de erro mostrada ao utilizador quando a operação de selecção de um volume de dados é cancelada (é seleccionado o botão *Back* no menu apresentado na Fig. 59).

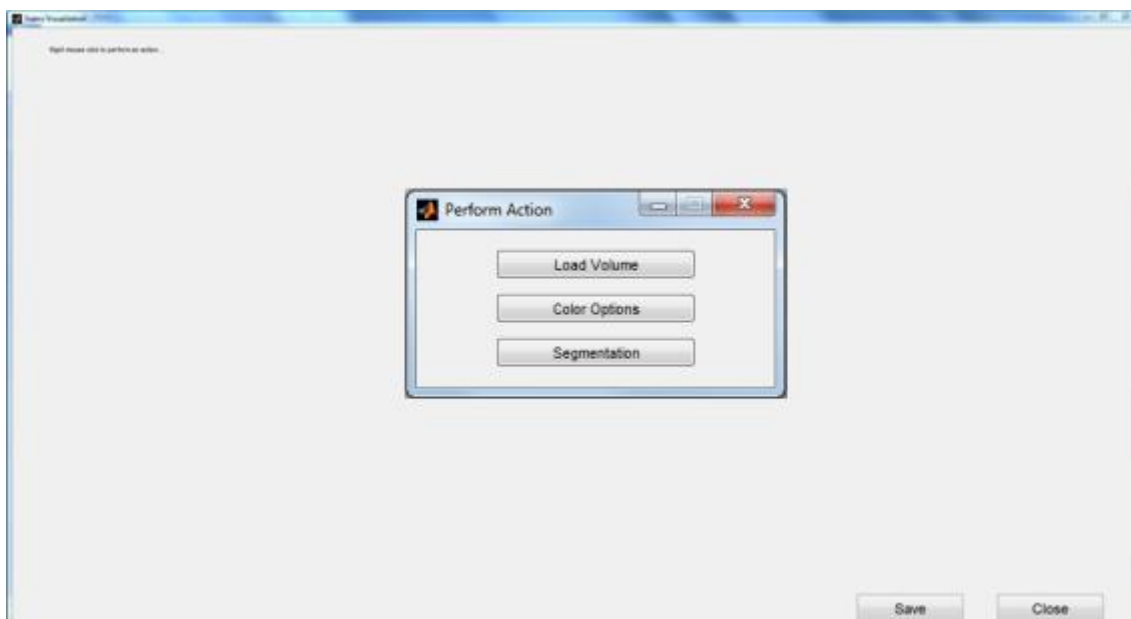


Figura 64 - Menu inicial do visualizador DICOM em *background* e menu de escolha de tarefas disponibilizado após se pressionar o botão direito do rato (a imagem não se encontra na escala normal).

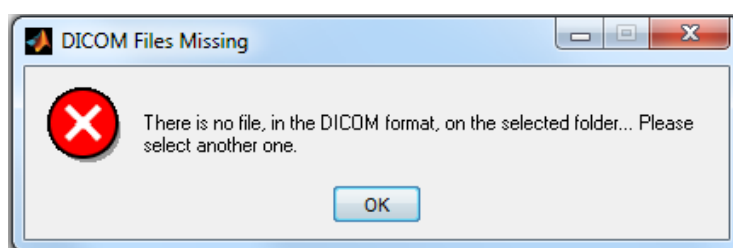


Figura 65 - Mensagem de erro que informa o utilizador do facto de não existirem ficheiros no formato DICOM na directoria por si seleccionada.

Para mais informações sobre o modo de funcionamento do visualizador de ficheiros DICOM, consultar o Anexo I.3 deste documento.

Deve-se esclarecer que o visualizador DICOM foi desenvolvido pelo Eng.º Francisco Caramelo, e adaptado para a integração nesta Interface pela sua desenvolvedora (adaptações ao nível da janela do menu de visualização, aspecto gráfico e descrições de alguns menus, e inserção da opção de guardar imagens). O visualizador *imago* foi desenvolvido integralmente em IDL[®] pelo Eng.º Nuno Ferreira, e foi simplesmente implementada na interface a linha de comandos para se aceder à IDL Virtual Machine[®].

8.2.3. Ferramentas que Permitem a Alteração dos Dados Constantes na Base de Dados

Passando agora à descrição dos menus associados à alteração de dados, na Fig. 66 podemos ver os itens aos quais se definiu que se poderiam aplicar alterações.

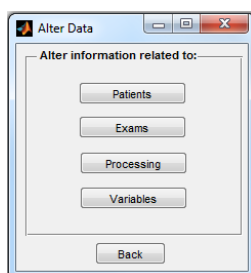


Figura 66 - Menu de selecção do tipo de informação que se pretende alterar (informação relativa aos pacientes, exames, processamento ou variáveis).

Seleccionando a opção de alterar dados relativos aos pacientes, é mostrado o menu apresentado na Fig. 67.

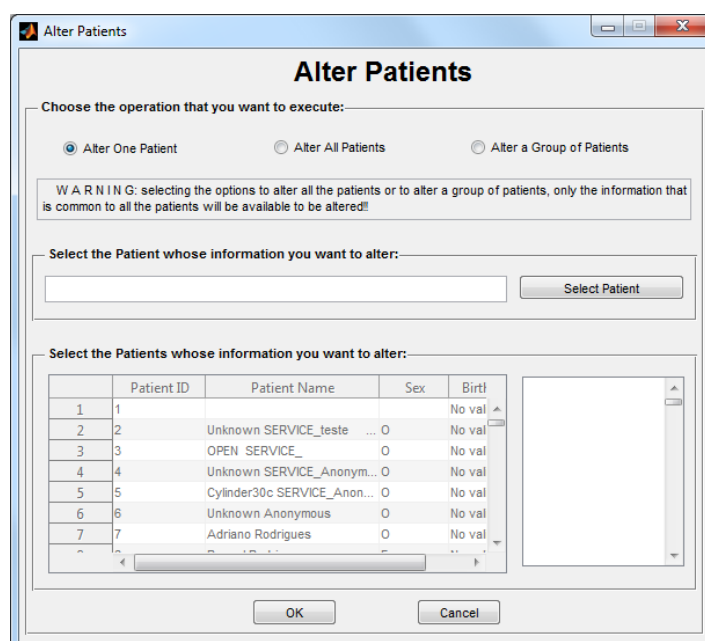


Figura 67 - Menu apresentado para que o utilizador possa alterar os campos da BD que contêm informação relativa aos pacientes.

Podemos ver que este menu disponibiliza 3 opções: alterar os dados relativos a um só paciente, alterar os dados de todos os pacientes da BD, ou alterar os dados de apenas um grupo de pacientes. Se se seleccionar a primeira opção, o primeiro painel fica activo. Pressionando o botão *Select Patient* é mostrado o menu apresentado na Fig. 68, com uma lista de todos os pacientes inseridos na BD, e com um resumo das informações referentes a cada um deles (nome, sexo, data de nascimento). O botão *OK* apenas fica disponível após a selecção de um (e apenas a selecção de um é permitida) dos pacientes da lista (Fig. 69), e após ser pressionado, o

identificador e o nome do paciente em causa são automaticamente colocados na caixa de texto do menu principal da alteração de dados dos pacientes (Fig. 71). Caso contrário, se se cancelar a operação sem seleccionar um paciente, ou se se tentar efectuar uma pesquisa sem se quer se abrir a caixa de diálogo de selecção de um paciente, será mostrado um aviso ao utilizador a informar que nenhum paciente foi seleccionado (Fig. 70).

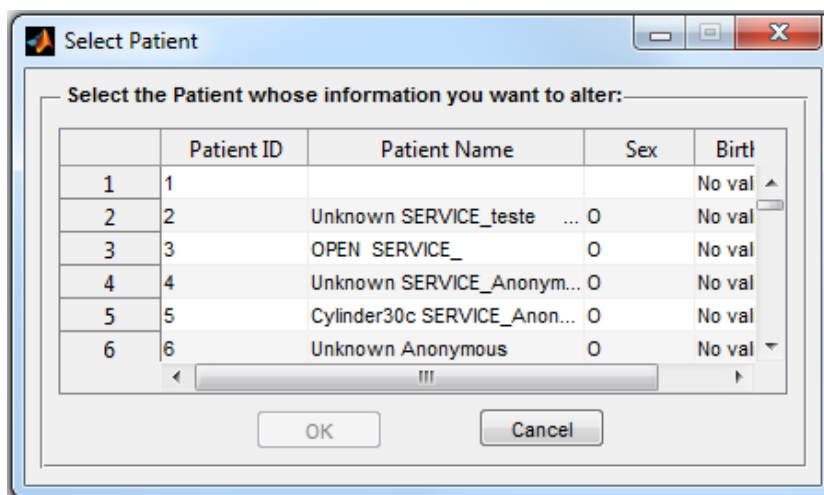


Figura 68 - Menu que permite a selecção do paciente ao qual se pretende alterar os dados.

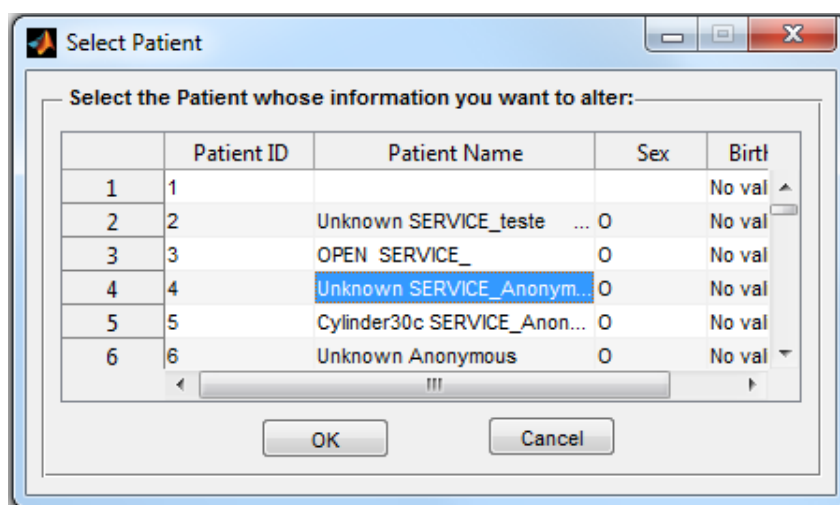


Figura 69 - Pormenor da activação do botão *OK* apenas após a selecção de um dado paciente da lista.

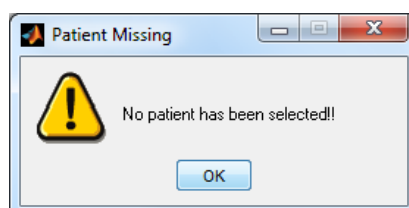


Figura 70 - *Warning* mostrado ao utilizador quando ele tenta alterar os dados de um dado paciente, sem o ter escolhido previamente.

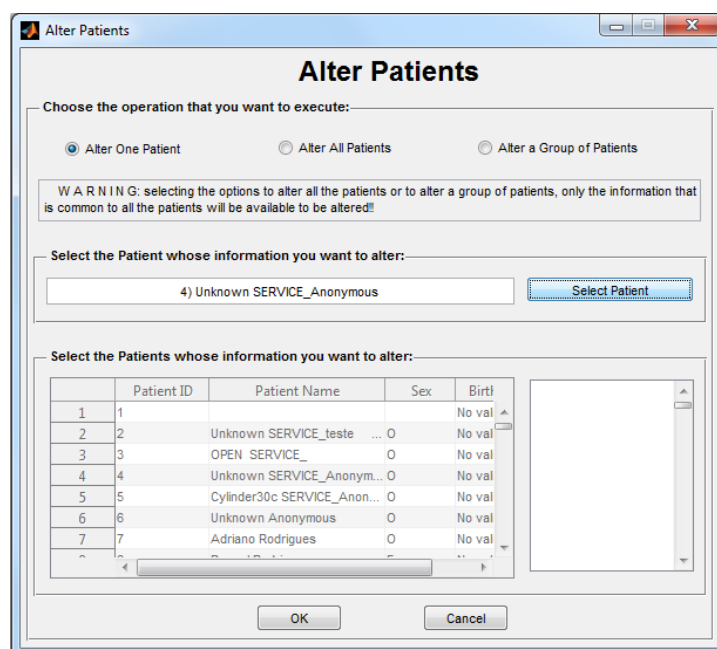


Figura 71 - Inserção do identificador e nome do paciente escolhido no menu principal de alteração de dados de pacientes.

Deve então seleccionar-se o botão *OK* do menu principal, para se proceder à alteração dos dados do paciente seleccionado. Ao executar esta acção, surge o menu de alteração dos dados de um paciente em concreto (Fig. 72), que mostra não só a informação retirada do ficheiro do exame que foi realizado a esse paciente, como também qualquer outro tipo de informação que lhe pertença e tenha sido adicionada *à posteriori* nas tabelas EAV. Este menu apresenta assim uma lista com o nome e o valor de todas as variáveis associadas ao paciente seleccionado, que se encontrem na BD. Para proceder à alteração da informação de qualquer um dos campos apresentados, basta clicar duas vezes sobre o valor da variável que se pretende modificar, e editar o conjunto de caracteres que nele constam (ver as figuras 72 e 73). Mas atenção: apenas serão efectuadas alterações que estejam de acordo com o domínio definido para cada atributo, ou seja, que respeitem o tipo de dados que podem ser associados a uma determinada variável da BD (por exemplo: não se permite a inserção de números menores que 0 – valor permitido unicamente por causa dos casos em que as datas não foram preenchidas – e maiores que 31, para um campo que se refira ao dia do mês de uma determinada data). Caso essas restrições não sejam respeitadas, é mostrada uma mensagem de erro ao utilizador com o nome do campo, e os valores que lhe são permitidos, e no final é apresentada uma tabela com todas as variáveis que foram mal alteradas, ou seja, onde se tentaram inserir valores que não fazem parte do seu domínio (ver Fig. 74). Apenas são modificadas ao nível da BD as variáveis que foram alteradas de forma correcta, respeitando as restrições, as restantes permanecem com o valor que tinham inicialmente (este facto é explicitado na mensagem também no menu que lista os nomes das variáveis em que a alteração não correu bem).

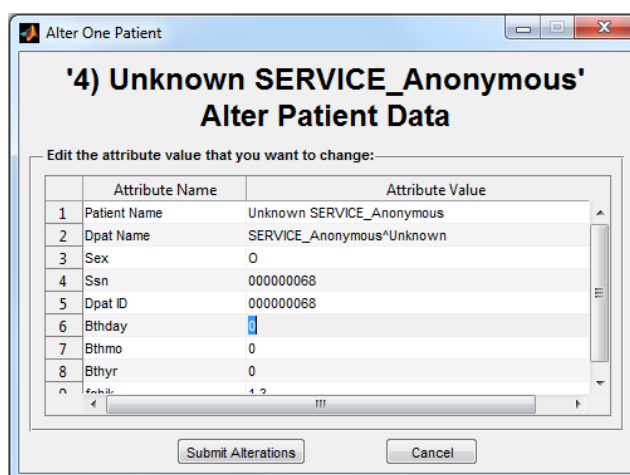


Figura 72 - Menu que apresenta o nome e o valor de todas as variáveis associadas a um dado paciente (inclui os atributos que foram adicionados em tabelas EAV).

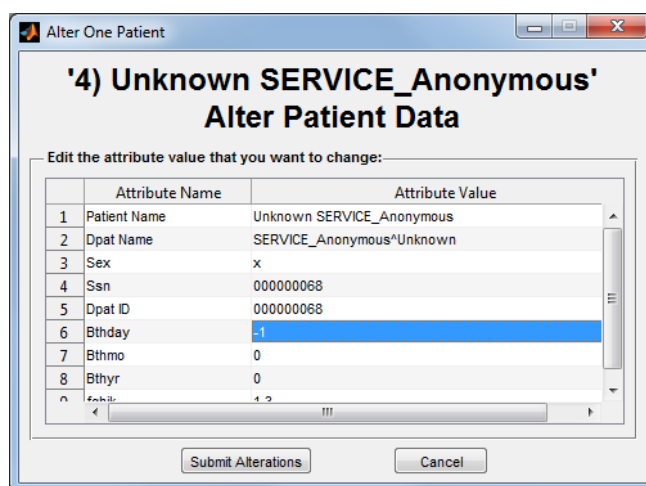


Figura 73 - Edição do campo que havia sido seleccionado para modificação no menu apresentado na Fig. 72.

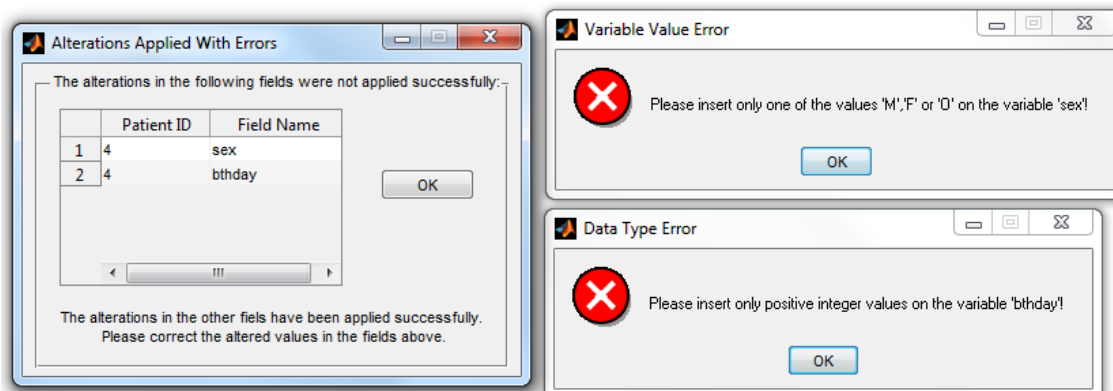


Figura 74 - À direita vemos as mensagens que vão surgindo ao longo da execução do programa de verificação dos valores das variáveis alteradas, no caso de os valores inseridos não se encontrarem concordantes com as restrições definidas para elas. À esquerda podemos ver a mensagem apresentada no final da execução referida, no caso de terem havido variáveis mal alteradas. Ele apresenta uma lista com os identificadores dos pacientes onde esse erro ocorreu (note-se que na selecção de alteração dos dados de um dado paciente, este será sempre o

mesmo, mas no caso da alteração de todos os pacientes ou de um determinado grupo destes, isto já não acontece) e os nomes das variáveis que foram mal alteradas.

Seleccionando a opção de alterar os dados referentes a todos os pacientes constantes da BD, todos os restantes painéis da BD ficarão inactivos (Fig. 75), e deve ser pressionado o botão OK para avançar com o processo de alteração. Surge então o menu de alteração dos dados de todos os pacientes, que é apresentado na Fig. 76.

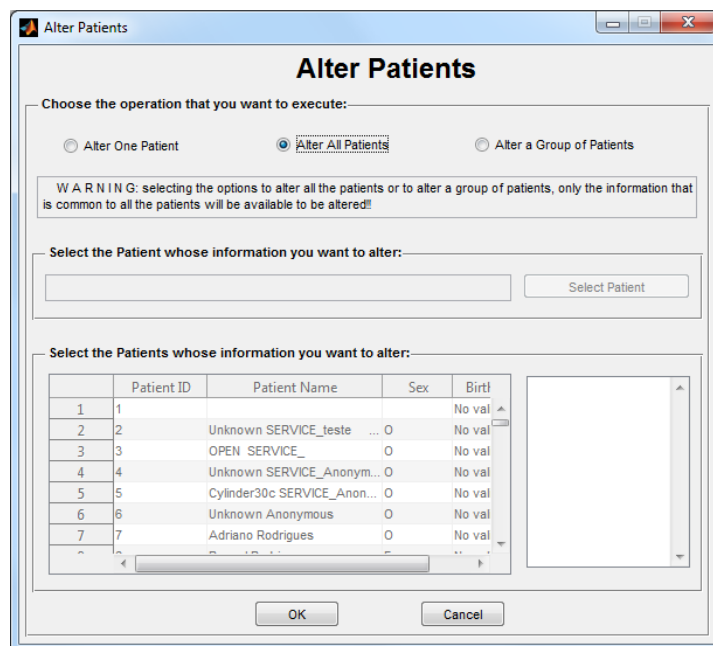


Figura 75 - Selecção da opção de alteração dos dados referentes a todos os pacientes da BD.

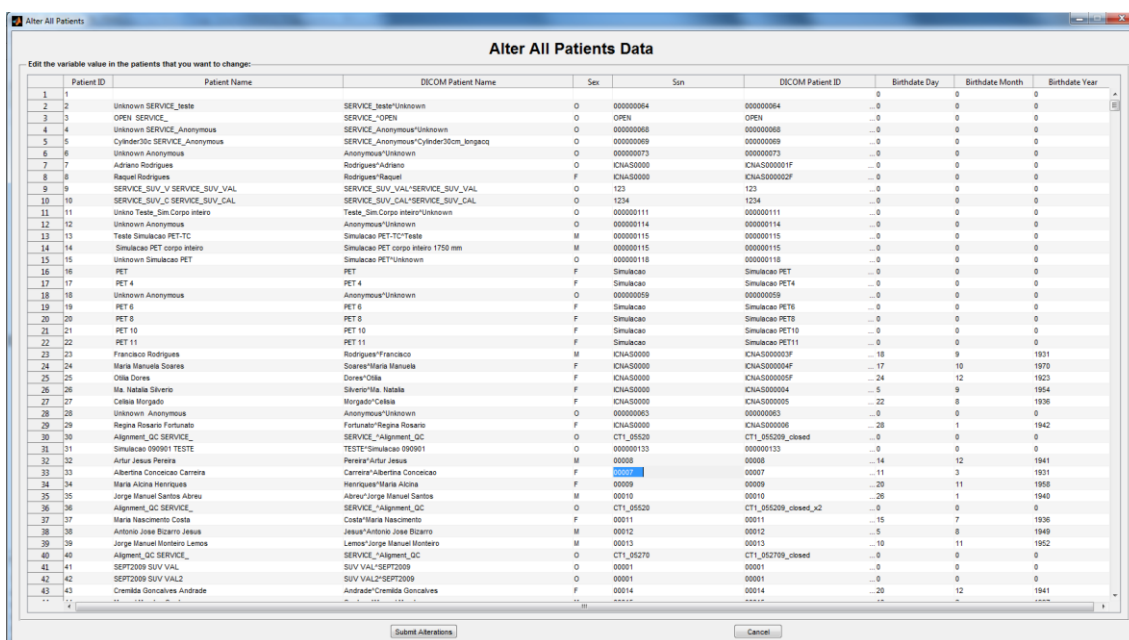


Figura 76 - Menu de alteração dos dados referentes a todos os pacientes constantes da BD.

Neste menu são apresentados os atributos que são comuns a todos os pacientes (ou seja, as variáveis que são retiradas de todos os ficheiros dos exames, e que constam da parte relacional do modelo conceptual da BD; as variáveis constantes das tabelas EAV não são mostradas; para se proceder à sua alteração deve-se executar a alteração dos dados dos pacientes um a um, como foi descrito anteriormente), sendo que não são permitidas alterações no número identificador do paciente, que só é mostrado na tabela por uma questão de organização dos dados e comodidade para o utilizador (todos os outros campos podem ser seleccionados e alterados do mesmo modo que se explicou para o caso de um só paciente). No fim de realizar todas as alterações pretendidas o utilizador deve pressionar o botão *Submit Alterations*, e caso não se tenham respeitado as restrições impostas pra qualquer uma das variáveis, são mostradas mensagens semelhantes às apresentadas na Fig. 74 (este processo também ocorre para a alteração de um grupo de pacientes que será apresentada de seguida).

Caso se tenha optado por alterar a informação relativa apenas a um dado conjunto de pacientes, o terceiro painel do menu principal de alteração de pacientes fica disponível (Fig. 77), e nele é mostrada uma tabela com um resumo dos dados de todos os pacientes da BD (identificador, nome, sexo e data de nascimento). Devem seleccionar-se aqueles que se pretendem alterar para a lista ao lado. Terminada a selecção, deve pressionar-se o botão *OK*, e será então mostrado um menu de alteração semelhante àquele que surgiu para a alteração de todos os pacientes, mas apenas com os pacientes seleccionados anteriormente nas entradas da sua tabela (Fig. 78). Tirando este pormenor, este menu funciona exactamente do mesmo modo que o de alteração de todos os pacientes da BD.

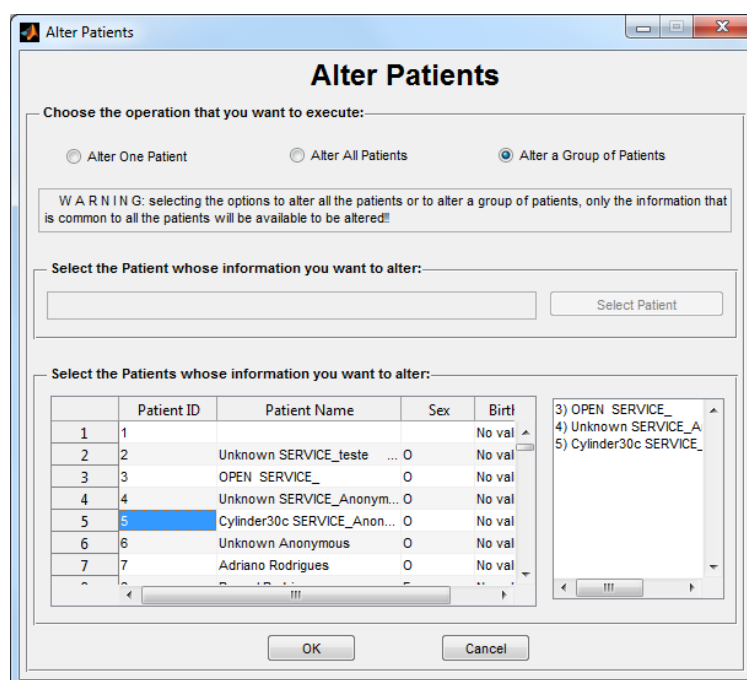


Figura 77 - Disponibilização do painel de selecção de um grupo de pacientes, quando é esta a opção escolhida.

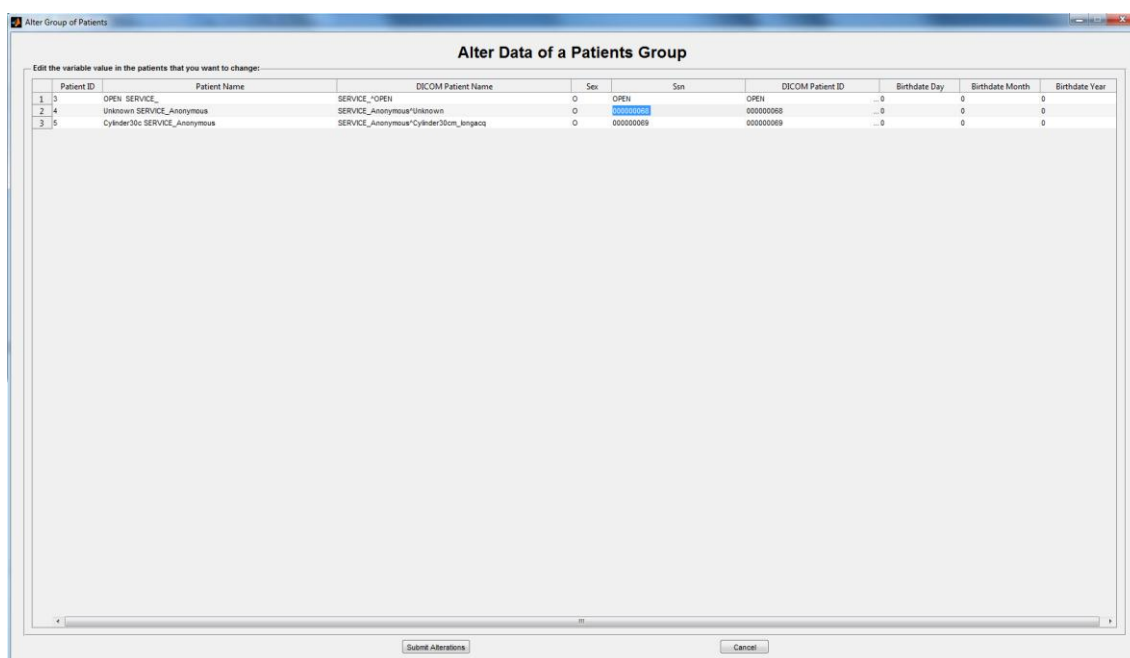


Figura 78 - Menu que permite a alteração apenas de um conjunto de pacientes.

Informar só que, quando as alterações são todas aplicadas com sucesso (não houve nenhum erro nos valores inseridos relativamente ao domínio dos atributos), é apresentada uma mensagem que notifica o utilizador do término e do sucesso da operação realizada (Fig. 79).

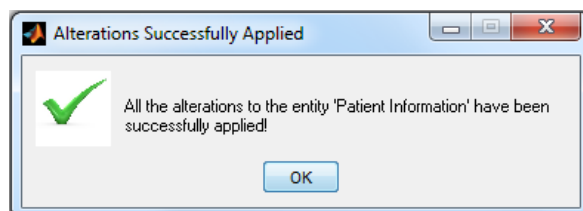


Figura 79 - Mensagem de sucesso apresentada ao utilizador quando todas as alterações são aplicadas sem qualquer erro. São apresentadas mensagens semelhantes para os outros casos de alteração que serão analisados: exames, processamentos e variáveis.

A alteração de exames funciona de modo análogo à de pacientes, com um grau de complexidade um pouco maior devido ao facto da informação dos exames se encontrar estruturada e hierarquizada em diferentes entidades, e devido a cada exame ser constituído por muitos cortes, cuja informação também tem de ser disponibilizada para alteração. Para ver detalhadamente o modo de funcionamento deste tipo de alteração, consultar o Anexo I.4 desta tese.

Quanto à alteração de processamentos, uma vez que as únicas informações que se repetem pra todos os processamentos são a data e hora da sua realização e a descrição do processamento que foi executado, e como não há interesse em alterar as primeiras duas informações (que apenas servem para fins de gestão da BD), e apenas importa alterar a

descrição do que foi realizado e a informação resultante desse processamento, e estando esta última armazenada na tabela EAV associada ao processamento de dados, sendo os dados armazenados nas tabelas deste tipo mais convenientemente extraídos mediante a pesquisa de um único identificador, definiu-se que apenas se permitiria a alteração de um processamento de cada vez. Assim sendo, ao seleccionar-se a opção *Processing* no menu da Fig. 66, é automaticamente apresentado um menu de escolha do processamento que se pretende alterar (Fig. 80). Ao seleccionar-se um dado processamento da tabela, o número identificador e a descrição deste são colocados na caixa de texto ao fundo do painel e o botão *OK* fica activo para que se possa proceder à alteração pretendida (Fig. 81). Ao ser pressionado é disponibilizado o menu mostrado na Fig. 82, que apresenta todas as variáveis em tabelas EAV associadas ao processamento seleccionado (neste caso, o processamento seleccionado não tem nenhuma) e a variável que contém a descrição do processamento efectuado. Os valores dessas variáveis podem ser alterados por meio da edição das células da tabela, como já foi descrito anteriormente para a alteração de dados de um paciente.

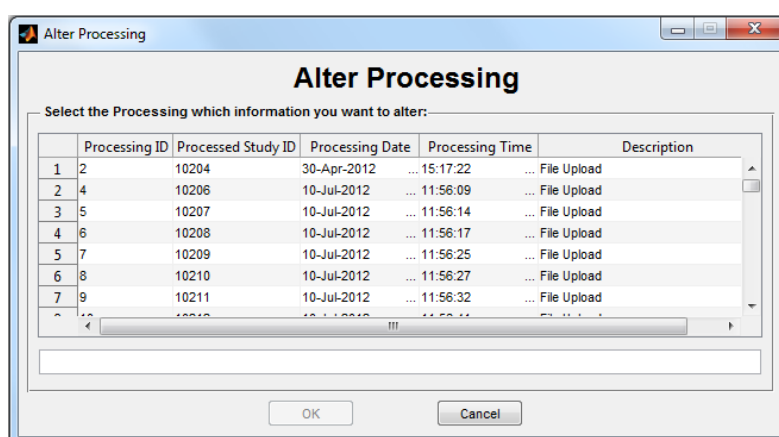


Figura 80 - Menu de selecção do processamento a alterar.

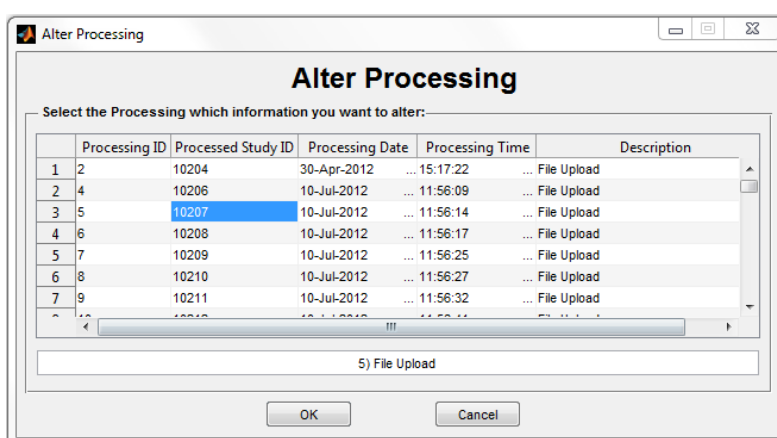


Figura 81 - Pormenores da colocação do identificador e descrição do processamento escolhido na caixa de texto, e do facto do botão *OK* se tornar activo.

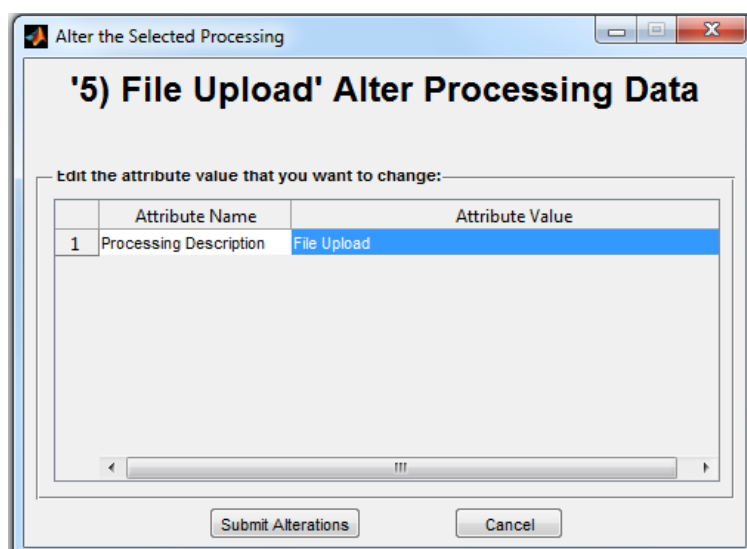


Figura 82 - Menu de alteração das variáveis associadas ao processamento seleccionado, por meio de edição das células da tabela onde são apresentadas.

No que respeita à alteração da tabela de descrição das variáveis, apenas se permite alterar o campo respeitante à definição da variável, uma vez que uma mudança no tipo de dados que armazena e na entidade em que se encontra implicaria uma alteração da estrutura do modelo da BD. Assim sendo apenas se criou um menu que mostra o nome da variável e a sua descrição numa tabela, onde apenas é permitida a edição das células respeitantes à descrição da variável (Fig. 83).

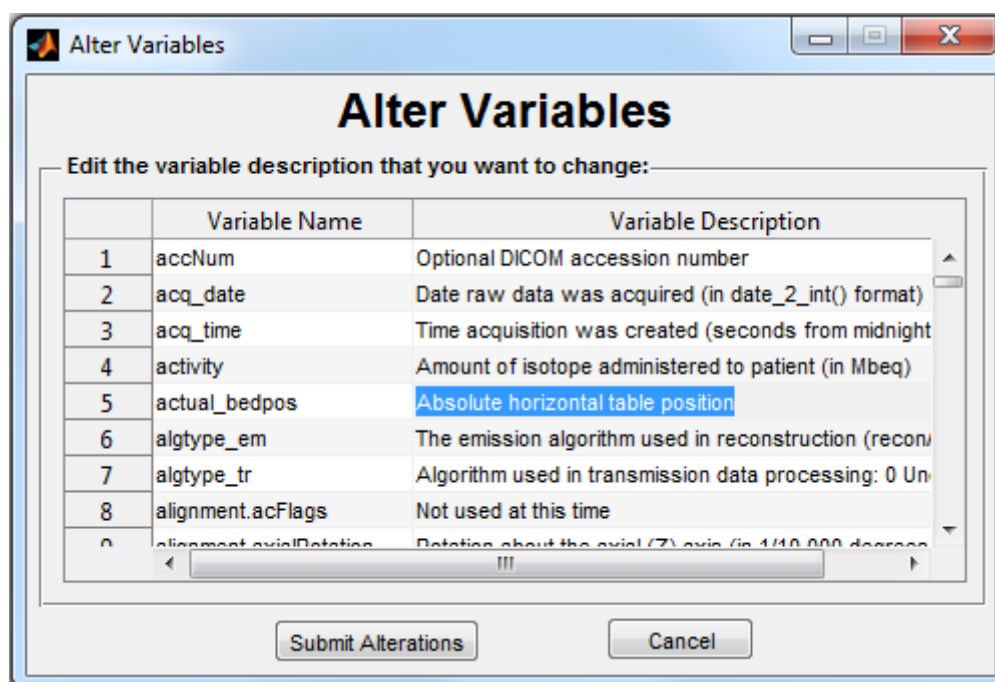


Figura 83 - Menu que permite a alteração da descrição das variáveis presentes na BD.

8.2.4. Processamento dos Dados da Base de Dados

Passemos agora à funcionalidade de processamento de dados (*Process Data*) disponibilizada pela plataforma. Ao escolher utilizá-la, aparece o menu da Fig. 84, onde podemos ver uma tabela com todos os exames constantes da BD e outra com todas as rotinas que lhe foram implementadas. O que o utilizador tem de fazer é seleccionar os exames que lhe interessam processar e as rotinas que deseja aplicar-lhes para as listas respectivas (só quando se seleccionar pelo menos um de cada um destes itens é que as opções de visualização dos resultados do processamento e de guardar esses resultados, bem como o botão *Process* ficam activos – ver figuras 85, 86 e 87), e depois pressionar o botão *Process*. A partir daqui, vai ser executada um programa que aplica cada uma das rotinas escolhidas a todos os exames seleccionados. Como este pode ser um processo muito demorado, foi implementada a barra de espera ilustrada na Fig. 88. É necessário escolher pelo menos uma das opções: ou visualizar os resultados do processamento ou guardar esses resultados na BD, caso contrário não teria sentido executar o processamento. Caso esta norma não seja respeitada, é mostrado um aviso ao utilizador (Fig. 89) e o processamento de dados não é executado. No final do processamento é mostrada uma mensagem ao utilizador a dar conta do sucesso da operação (podemos vê-la na Fig. 90).

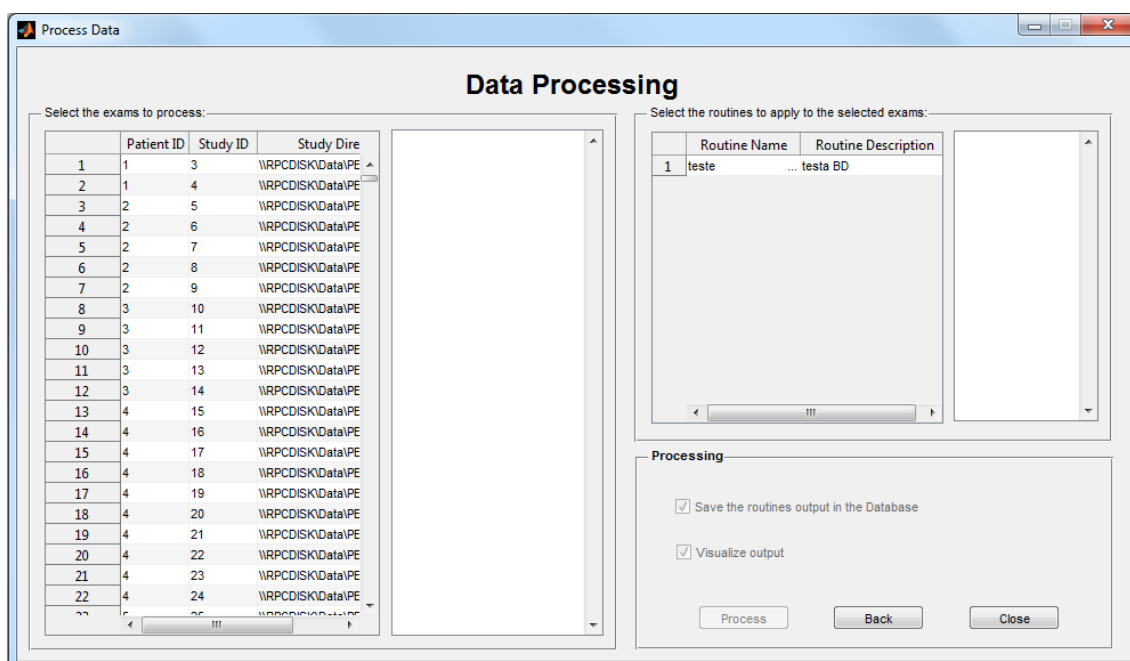


Figura 84 - Menu de processamento de dados.

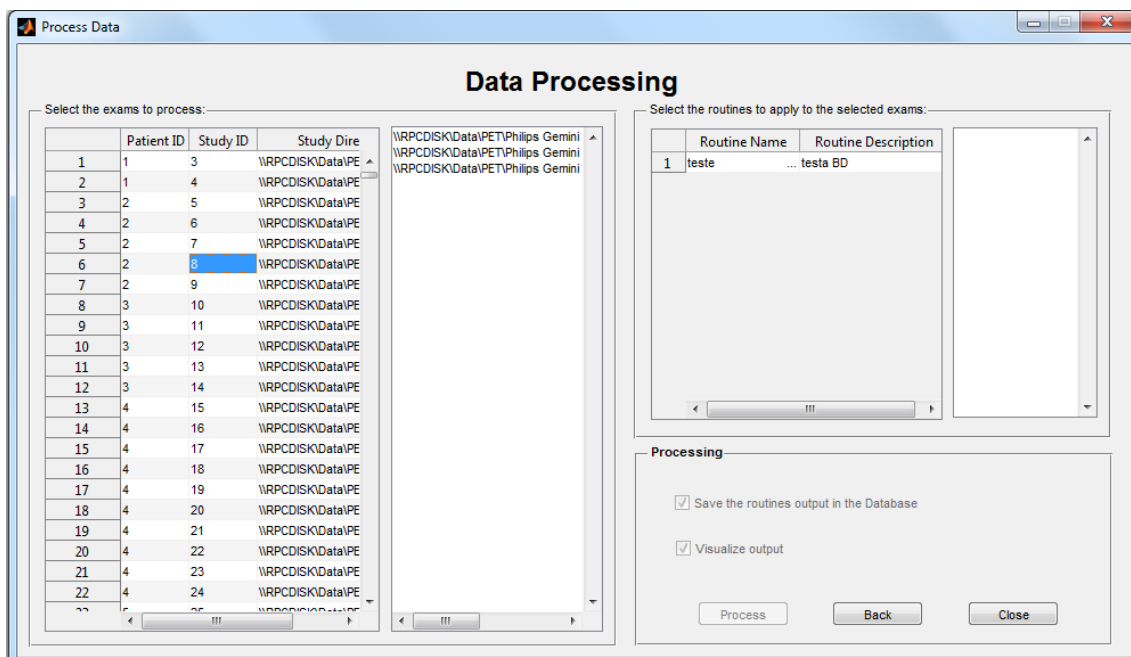


Figura 85 - Pode-se ver que, com apenas a selecção dos exames a processar, as funcionalidades de guardar e visualizar os resultados, bem como o botão que permite o processamento, permanecem inactivos.

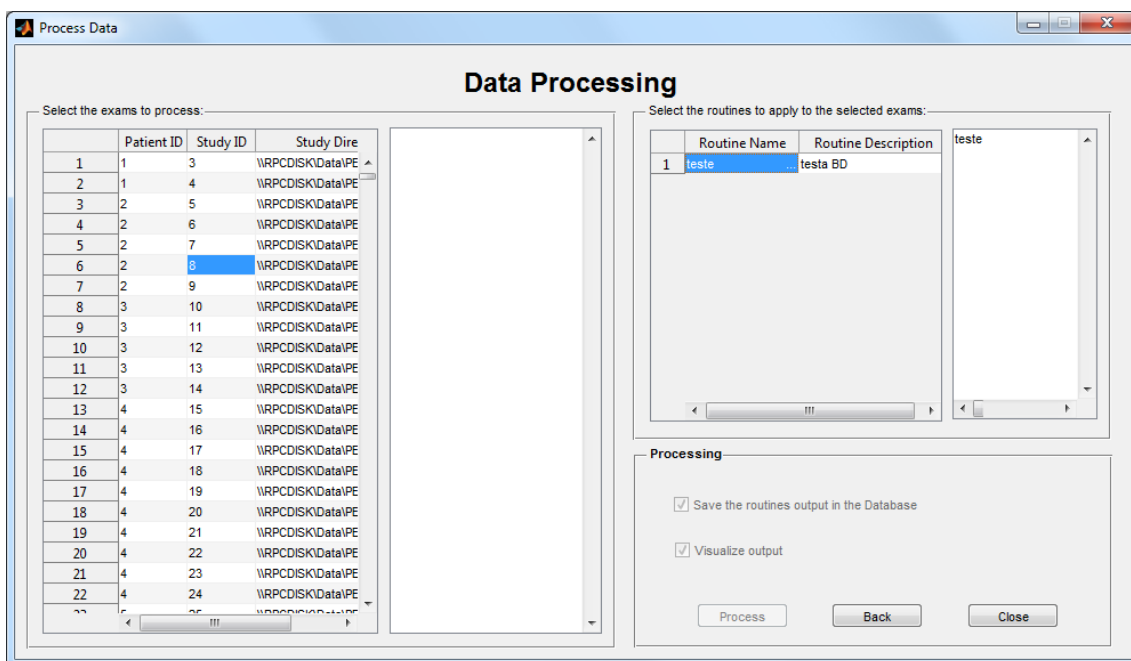


Figura 86 - Com selecção apenas das rotinas, e não dos exames, acontece o mesmo que descrito para a figura anterior.

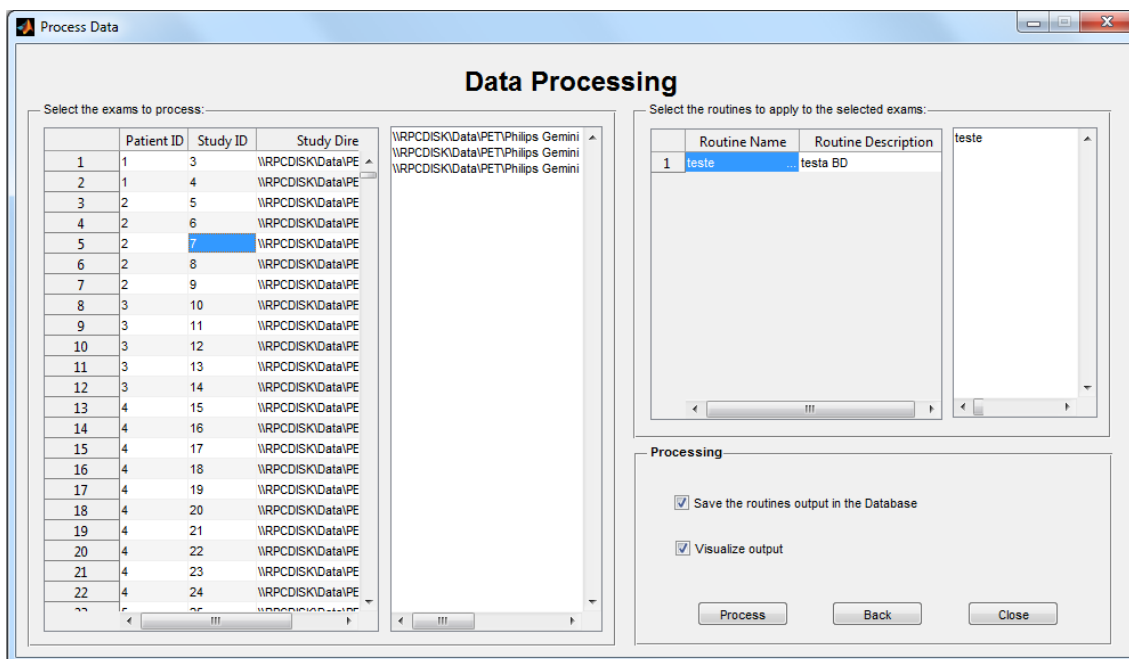


Figura 87 - Finalmente, com ambas as listas preenchidas, as funcionalidades referidas nas duas figuras anteriores ficam disponíveis.

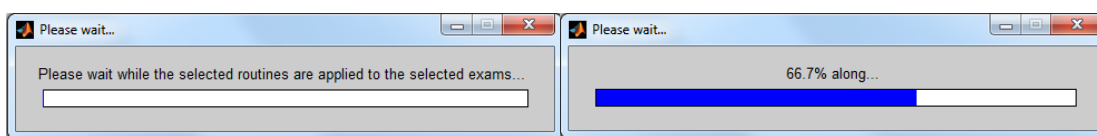


Figura 88 - Barra de espera implementada para a monitorização da evolução do processamento de dados.

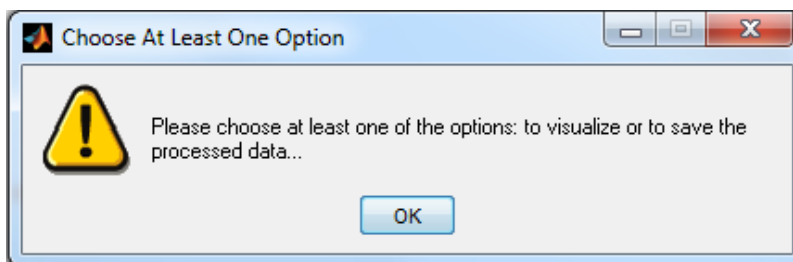


Figura 89 - Aviso mostrado ao utilizador a informá-lo da necessidade de ter de escolher pelo menos uma das opções disponíveis.

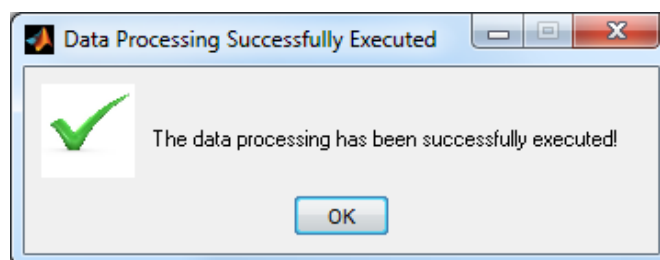


Figura 90 - Mensagem que notifica o utilizador do fim e do sucesso do processamento.

8.2.5. Ferramentas Administrativas da Base de Dados

Caso o Administrador decida aceder às ferramentas administrativas (*Administrative Tools*) que lhe são disponibilizadas, surgirá o menu da Fig. 91. Nele estão catalogadas 5 ferramentas de gestão da BD, que permitem, entre outras coisas, a adição e remoção de: utilizadores, pacientes, exames, processamentos e rotinas.



Figura 91 - Ferramentas de gestão e administração da BD.

Com a selecção da primeira ferramenta surge o menu representado na Fig. 92. Este menu permite adicionar um novo utilizador à BD atribuindo-lhe um perfil, um nome de utilizador e uma palavra-chave e pressionando o botão *Add User*. No entanto a adição só é levada a cabo se o nome de utilizador que se está a tentar introduzir ainda não existir nos registos da BD (caso exista é mostrada a mensagem de erro da Fig. 93, e o utilizador não é adicionado à BD). Se tentar introduzir-se o caracter ‘espaço’ em qualquer um dos campos mencionados, ou se se tentar adicionar um utilizador sem preencher previamente esses campos, são também mostradas as mensagens de erro representadas nas figuras 94, 95 e 96. A título de exemplo mostra-se o procedimento de introdução de um novo utilizador na Fig. 97, e a mensagem de sucesso mostrada após este ser inserido na BD (Fig. 98).

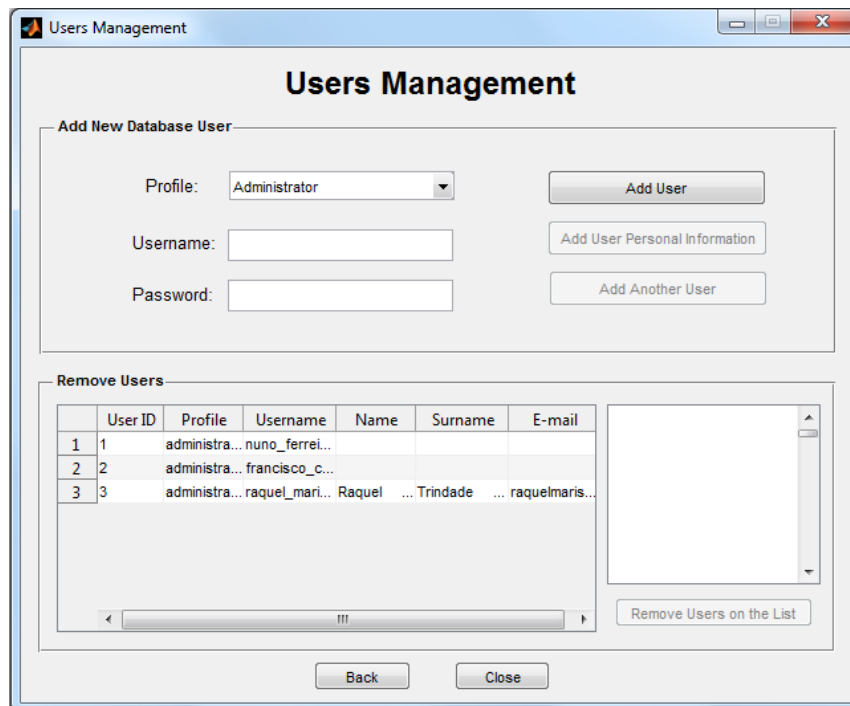


Figura 92 - Menu que permite a gestão dos utilizadores existentes na BD.



Figura 93 - Mensagem de erro apresentada ao Administrador quando este tenta inserir um nome de utilizador que já consta dos registos da BD.

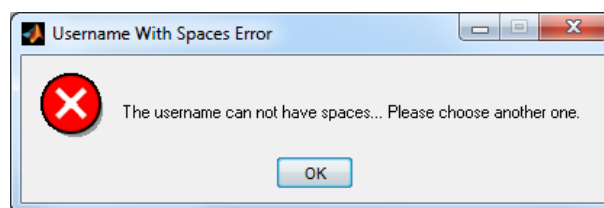


Figura 94 - Mensagem de erro que é apresentada quando o Administrador tenta inserir um caracter 'espaço' num novo nome de utilizador.

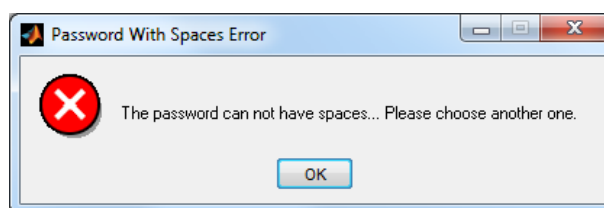


Figura 95 - Mensagem de erro que é apresentada quando o Administrador tenta inserir um caracter 'espaço' numa nova palavra-chave.

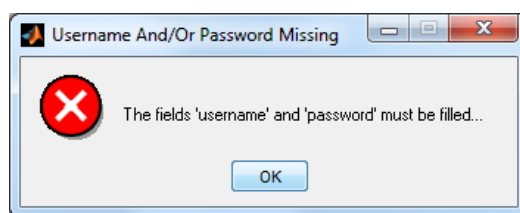


Figura 96 - Mensagem de erro mostrada ao Administrador quando este pressiona o botão *Add User* sem ter preenchido um dos campos (ou ambos) referentes ao nome de utilizador e à palavra-chave.

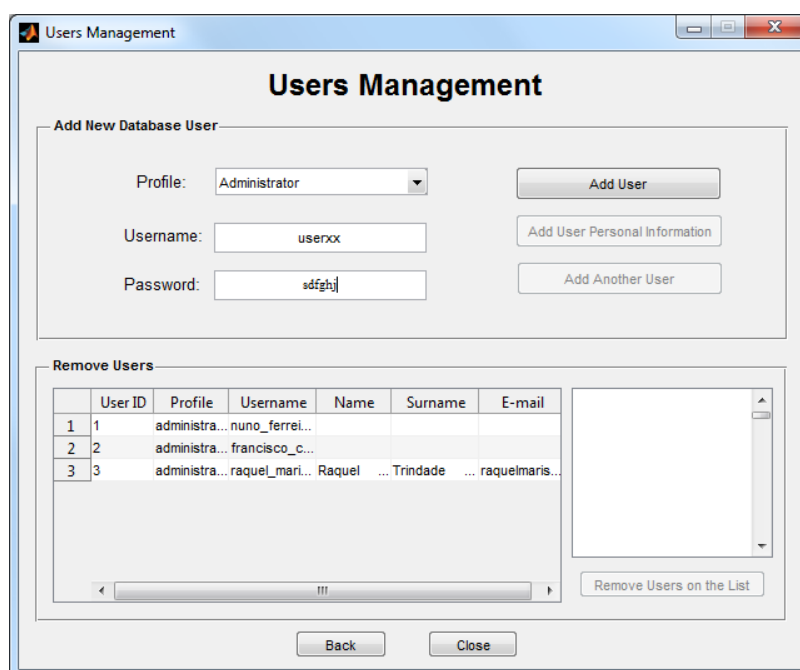


Figura 97 - Exemplificação da adição de um novo utilizador.

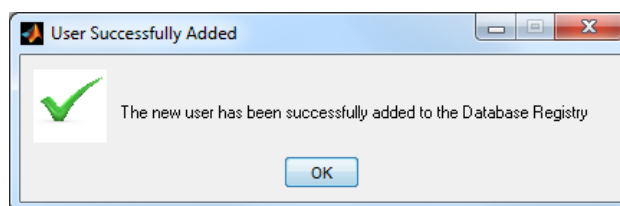


Figura 98 - Mensagem que informa que o novo utilizador foi adicionado com sucesso à BD.

Na Fig. 99 podemos ver que logo após a adição do utilizador, a tabela com todos os utilizadores, disponível no painel de remoção, é actualizada. Ficam também operacionais os botões que permitem a adição de mais informação relacionada com o utilizador que se acabou de inserir, ou então a adição de outro novo utilizador. Caso se opte por adicionar a informação pessoal do novo utilizador, surge o menu da Fig. 100 – (a). Um exemplo de possível preenchimento deste menu, é dado na Fig. 100 – (b). Basta depois pressionar o botão *Submit* e as informações introduzidas são guardadas na BD. De realçar que é guardada uma cópia da

imagem escolhida para foto de perfil na pasta ‘Medical Image Database’, para que a leitura desta depois seja directa na pasta corrente de trabalho doMATLAB®. No final desta operação é apresentada uma mensagem de sucesso ao utilizador (Fig. 101). Caso se tenha optado pela adição de outro utilizador à BD (*Add Another User*), a janela de inserção de dados volta ao estado inicial (Fig. 102).

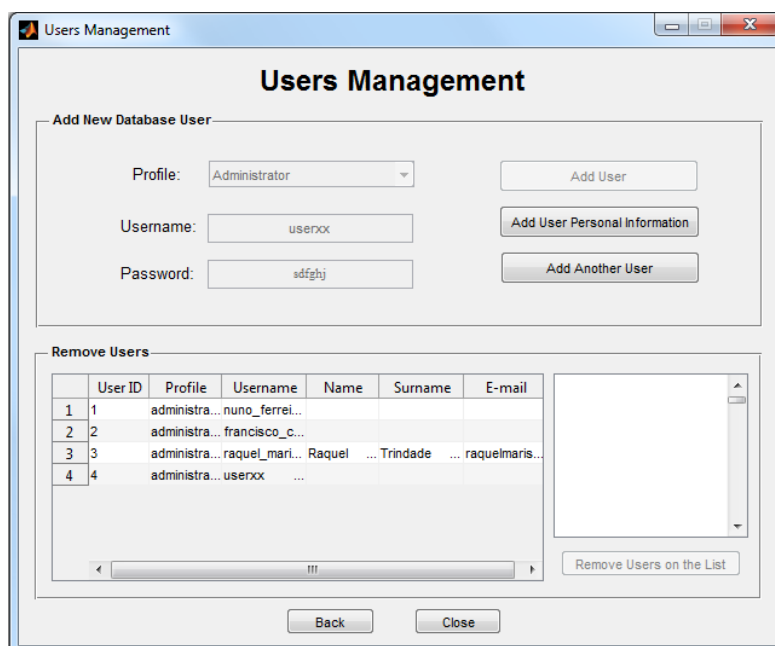


Figura 99 - Menu após a adição do novo utilizador à BD. Vê-se o pormenor da actualização automática dos elementos presentes na tabela onde se encontram todos os utilizadores existentes na BD.

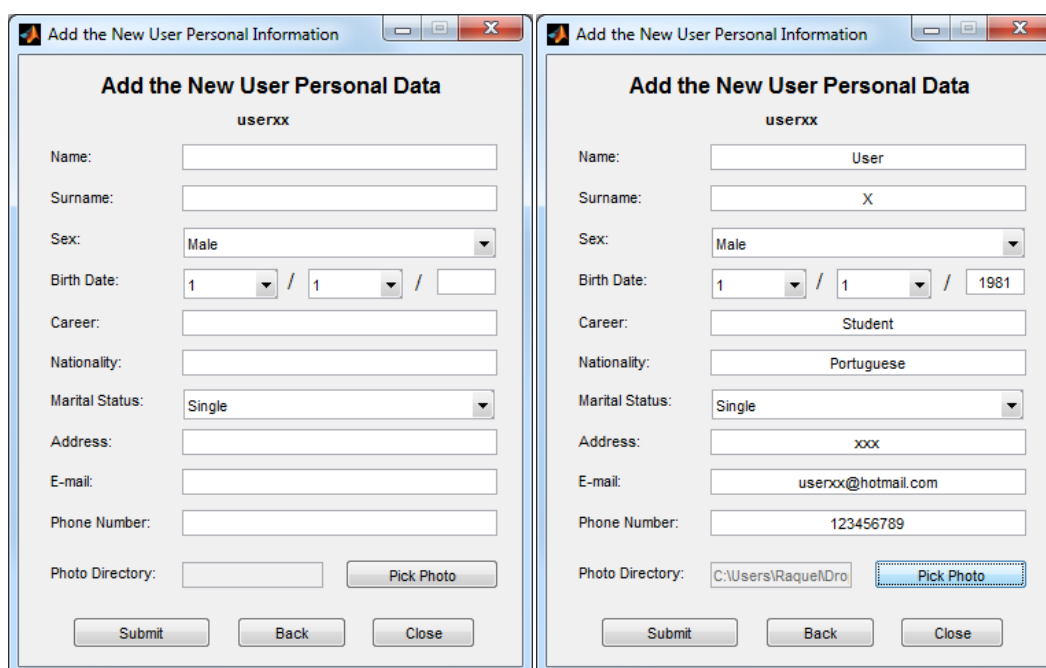


Figura 100 - Menu de adição da informação pessoal do utilizador. (a) - À esquerda podemos ver o menu por preencher. (b) – À direita vemos um exemplo de preenchimento do menu. Note-se que o botão *Pick Photo*

possibilita a escolha da imagem que o utilizador irá ter como foto de perfil (através de uma caixa de diálogo mostrando apenas as imagens constantes de uma dada directoria)

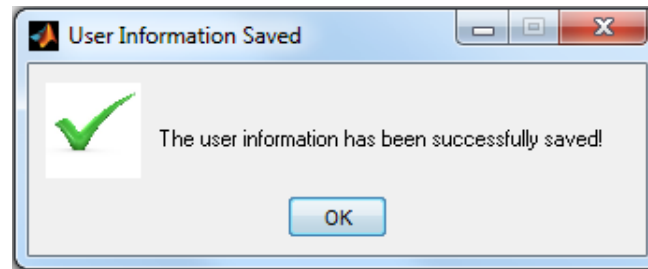


Figura 101 - Mensagem apresentada ao Administrador dando-lhe conta do sucesso da introdução dos dados pessoais do novo utilizador na BD.

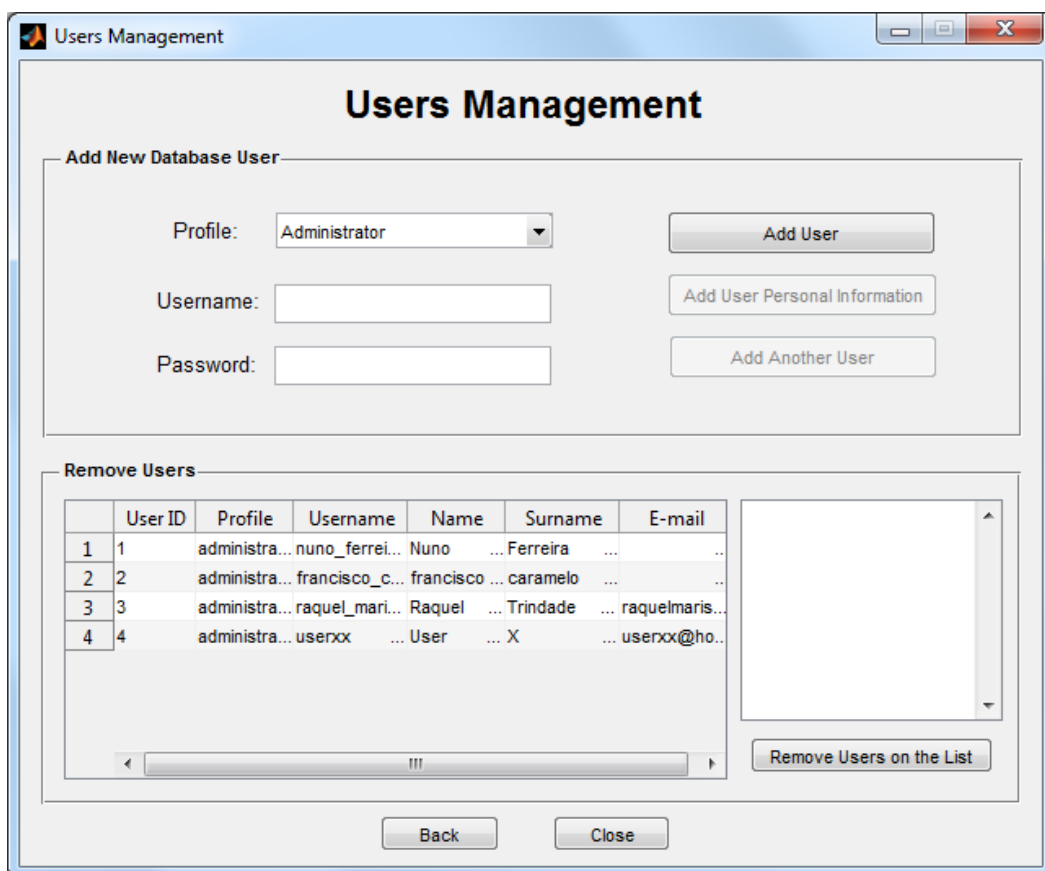


Figura 102 - Aspecto do menu após se seleccionar a opção *Add Another User*.

Caso contrário, se pretendam eliminar utilizadores da BD, basta seleccionar os nomes dos utilizadores a eliminar para a lista ao lado da tabela que contém as informações relativas a todos os utilizadores registados na BD, e pressionar o botão *Remove Users on the List*. Como podemos ver nas figuras 103 e 105 os utilizadores seleccionados são automaticamente removidos, e no final do processo é apresentada uma mensagem ao Administrador a dar conta do sucesso da operação de remoção (Fig. 104).

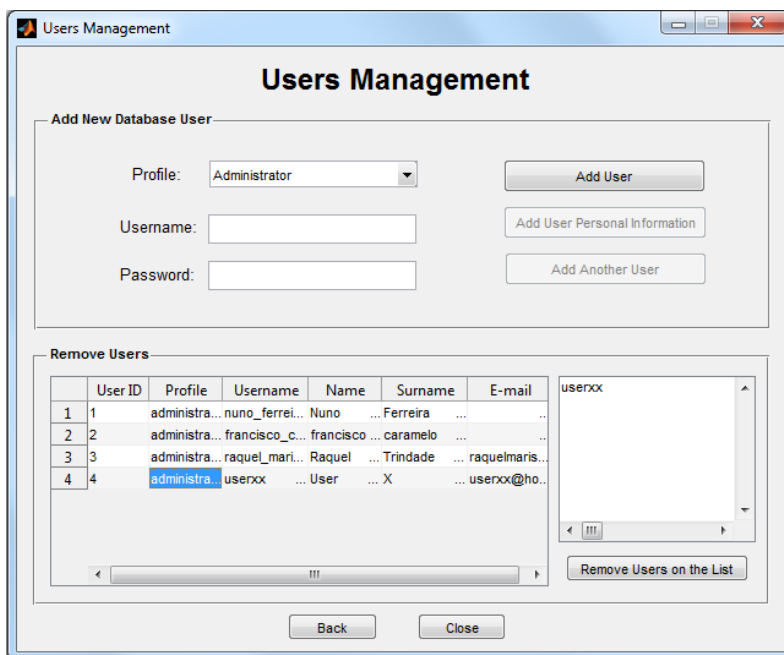


Figura 103 - Selecção dos utilizadores a serem removidos.

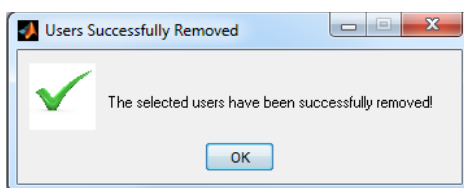


Figura 104 - Mensagem apresentada ao Administrador a dar conta do sucesso da operação de remoção.

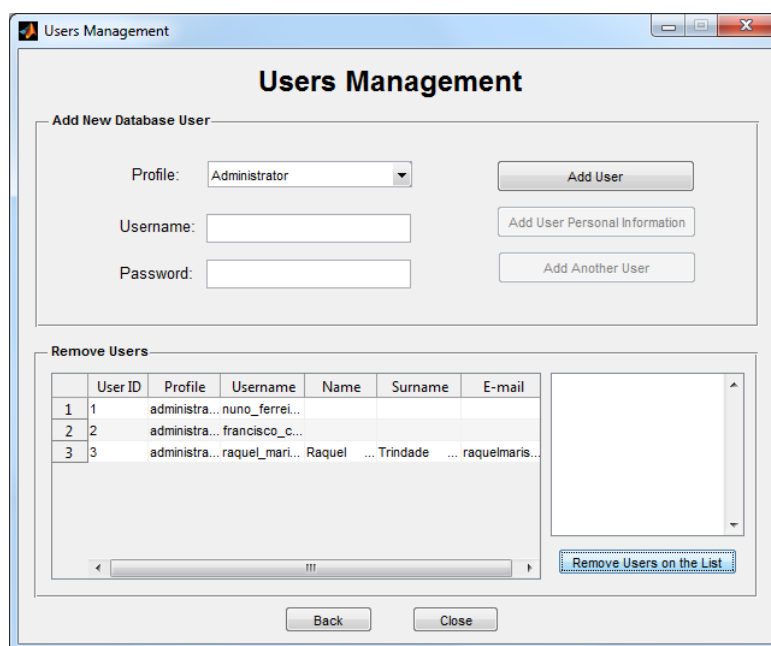


Figura 105 - No final do processo, a lista de utilizadores existentes na BD é automaticamente actualizada, tal como acontecia após a adição de um novo utilizador.

Passando agora à segunda ferramenta administrativa, a que permite a gestão de pacientes, ela é apresentada mediante o menu da Fig. 106. Nele é explicitado que a adição de novos pacientes apenas é permitida mediante a actualização da BD, com o carregamento de novos ficheiros (o mesmo acontece para a adição de exames). O menu permite sim a remoção de pacientes da BD, através da selecção das entradas da tabela apresentada, onde se encontram listados todos os pacientes da BD. Após o preenchimento da lista de remoção, deve premir-se o botão *Eliminate Patients on the List* para proceder à eliminação (atenção que este botão só fica activo quando a lista se encontra preenchida – Fig. 107). O sistema não procede automaticamente à remoção. Primeiro apresenta um aviso ao utilizador, que o informa que a sua acção apagará também todos os exames da BD que estejam ligados aos pacientes que se pretende remover, e todos os processamentos que, por sua vez, estejam ligados a esses exames (Fig. 108). Se o utilizador escolher continuar com a operação mesmo assim, todos os dados anteriormente referidos serão eliminados. Caso contrário a operação é abortada.

Neste menu é também disponibilizada uma rotina que procura elementos redundantes ao nível dos pacientes constantes da BD. Este algoritmo procura exames que tenham sido realizados pelo mesmo paciente, e então aglomera-os: pretende-se identificar possíveis casos de emparelhamento, mas pedir validação desse emparelhamento ao utilizador antes de o executar. Verifica se o paciente com determinada data de nascimento, determinado sexo e um nome com um determinado grau de similaridade se repete nas entradas da tabela ‘Informação do Paciente’ da nossa BD (nesta pesquisa excluem-se os pacientes que tenham cada um dos valores da sua data de nascimento a nulo, os quais, geralmente, estão identificados como desconhecidos, através da designação ‘Unknown Anonymous’). Se se encontrar tal redundância, e esta for confirmada pelo administrador que efectuou a pesquisa, então, na BD, os exames de todos os pacientes redundantes são atribuídos ao primeiro paciente a ser adicionado com tais características, e as entradas correspondentes aos pacientes redundantes relativamente ao primeiro são eliminadas da BD. Note-se que na memória física do sistema os ficheiros permanecem tal e qual como estavam. Assim sendo, a directoria para os dados processados para aquele exame continuará a ser a inicial, as pastas dos pacientes redundantes. Foi implementada uma barra de espera, porque o processo de procura de pacientes com características semelhantes costuma ser um pouco demorado (Fig. 109), e para cada conjunto de pacientes redundantes encontrado, é mostrado um menu como o da Fig. 110. O utilizador deve seleccionar os pacientes da tabela que confirma serem redundantes, adicioná-los à lista, e premir o botão *Confirmation* (este também só fica disponível enquanto a lista tem elementos a preenche-la – Fig. 111).

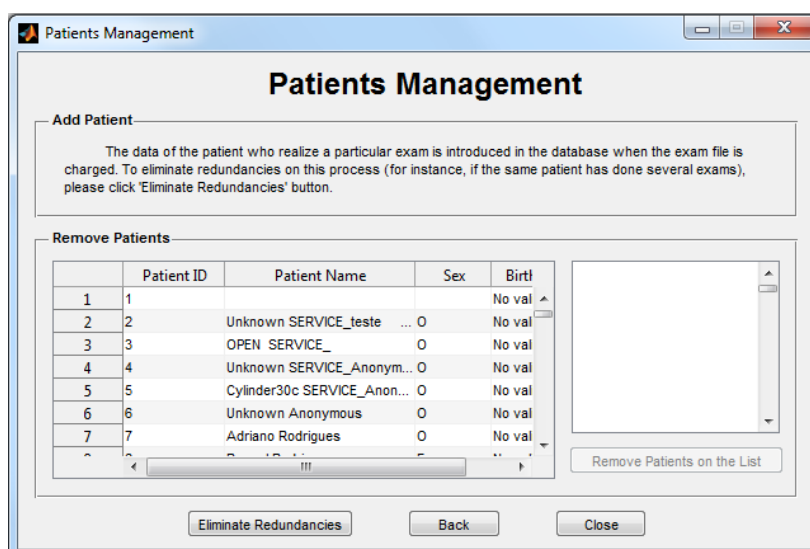


Figura 106 - Menu que permite a gestão dos pacientes da BD.

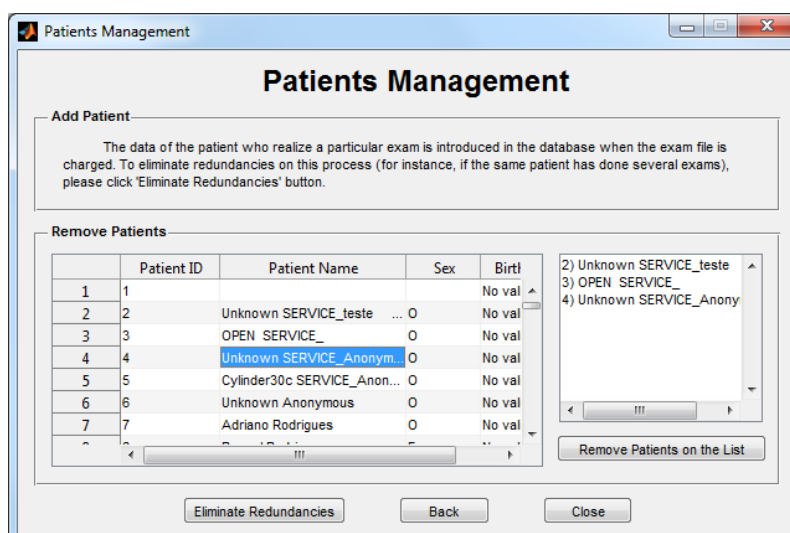


Figura 107 - Pormenor do botão de remoção só permanecer activo enquanto a lista contém itens.

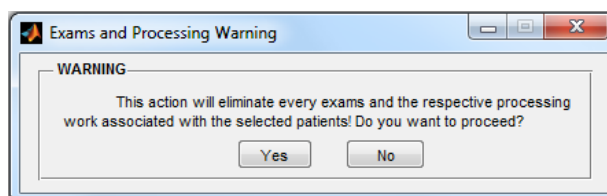
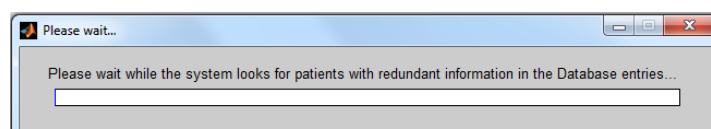


Figura 108 - Aviso que notifica o utilizador do facto da remoção de pacientes implicar também a remoção dos exames a eles associados, e dos processamentos efectuados a estes últimos. O utilizador deve escolher se pretende ou não prosseguir com a operação.



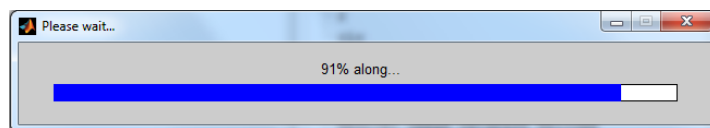


Figura 109 - Barra de espera implementada para monitorizar o processo de procura de pacientes redundantes nas entradas da BD.

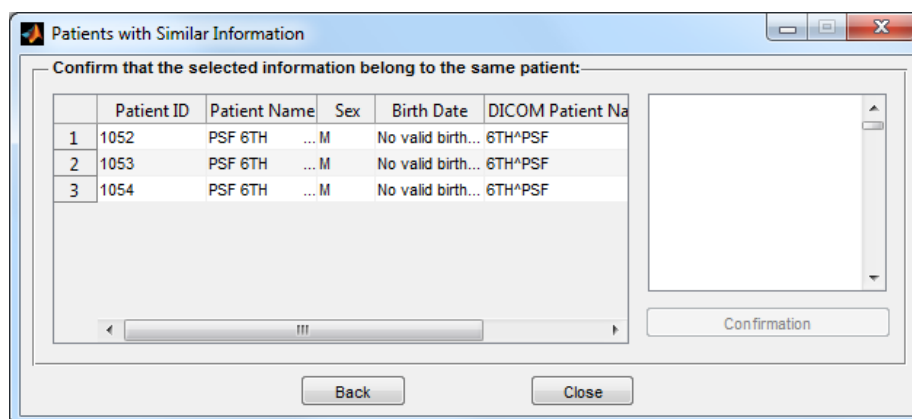


Figura 110 - Menu que mostra um dado conjunto de dados suspeitos de pertencerem a um mesmo paciente

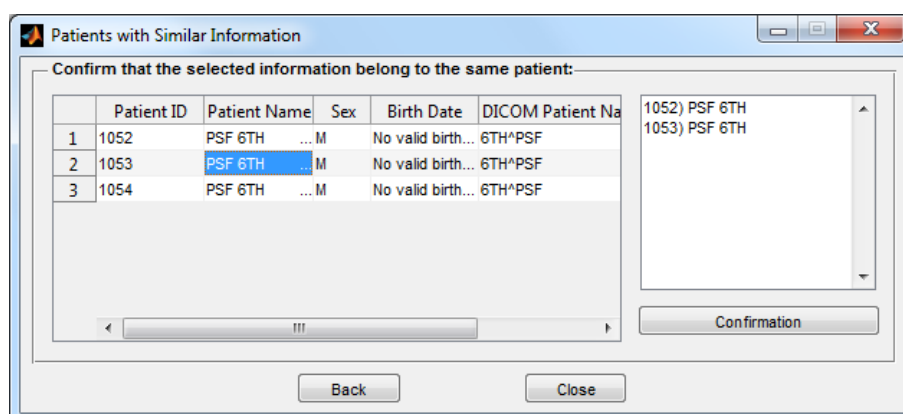


Figura 111 - Menu mostrando o pormenor do botão *Confirmation* só ficar activo aquando do preenchimento da lista de pacientes redundantes.

Passando agora à gestão dos exames existentes na BD, como já foi referido, apenas disponibiliza a eliminação de exames. A sua adição tem que ser feita através do carregamento de novos ficheiros de exames na BD. Após a selecção dos exames que se pretendem eliminar de entre todos os exames da BD que são listados na tabela do menu da Fig. 112, para a lista ao lado desta (Fig. 113), o utilizador deve premir o botão *Remove Exams on the List*. Mais uma vez é notificado que esta operação implica a remoção também de todos os processamentos associados aos exames a remover, e deve decidir se pretende continuar com a operação ou não (Fig. 114).

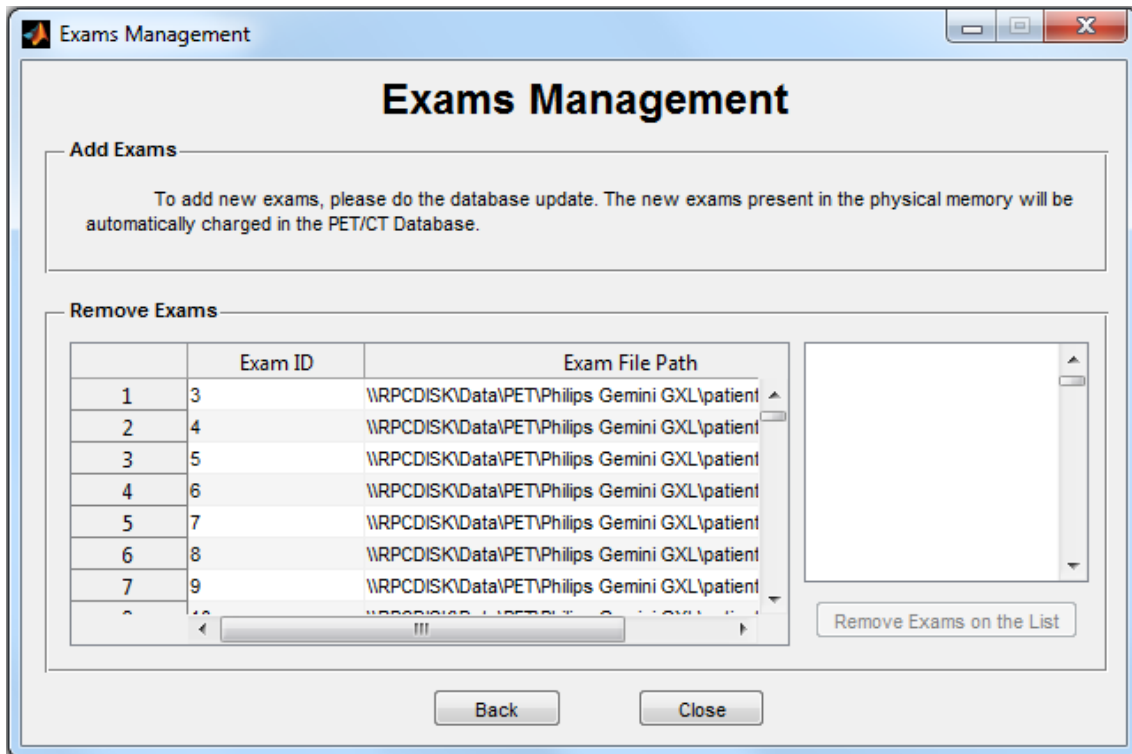


Figura 112 - Menu que permite a remoção de exames da BD.

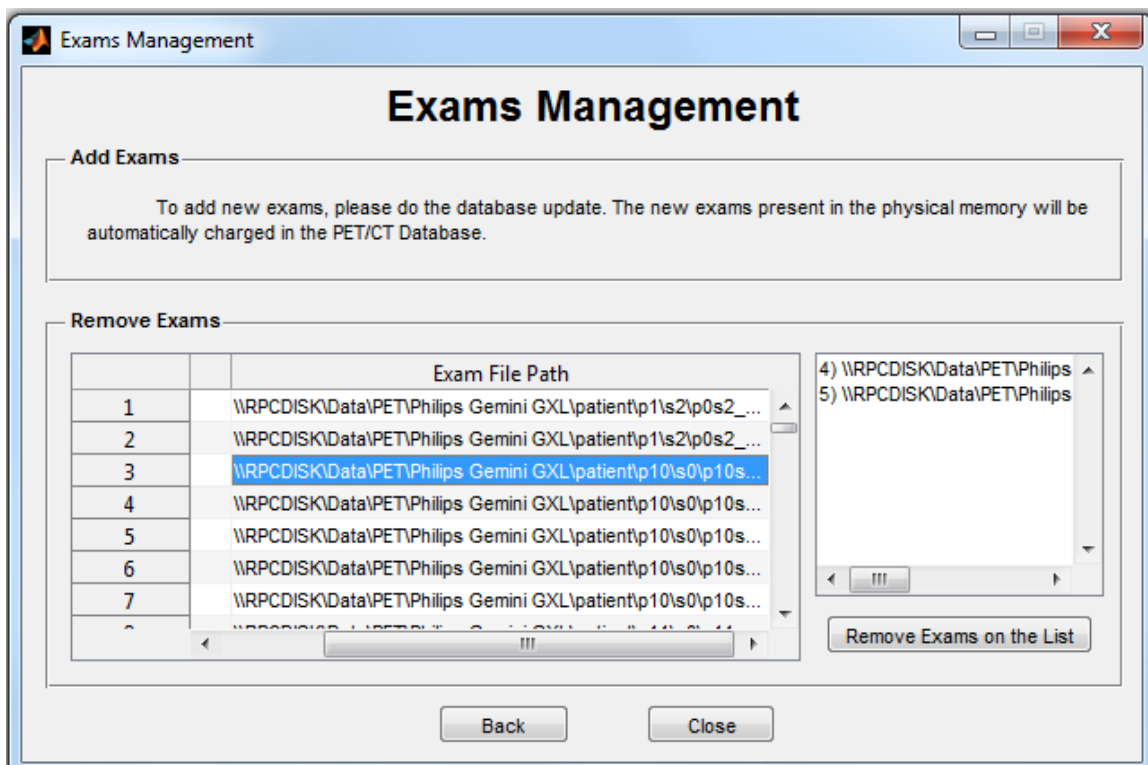


Figura 113 - Ilustração da forma de selecção dos exames a remover (note-se que o botão para proceder à remoção só fica activo quando a lista de itens a remover se encontra preenchida)

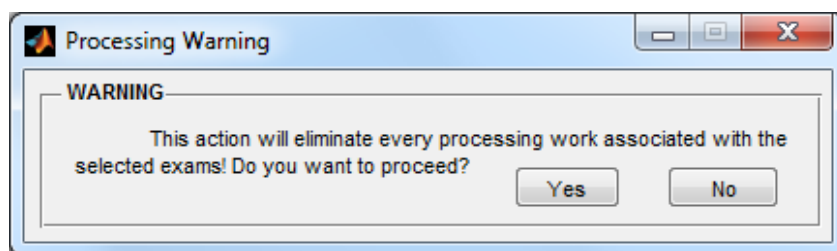


Figura 114 - Menu que informa o utilizador do facto de também serem removidos todos os processamentos associados ao exames seleccionados, e que lhe permite escolher continuar ou não com a operação de remoção.

No final das operações de remoção tanto de pacientes, como de redundâncias nos pacientes, como de exames, como de processamentos e rotinas (que serão descritos de seguida nesta secção) são apresentadas mensagens de sucesso semelhantes à que se pode ver na Fig. 104, mas adaptadas para as situações supra-mencionadas. Quando se cancela uma das operações de remoção, é apresentada uma mensagem semelhante à da Fig. 115.

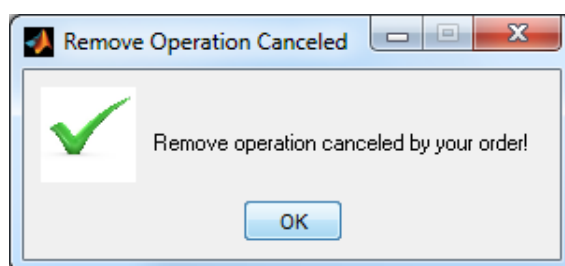


Figura 115 - Mensagem apresentada quando se cancela uma operação de remoção.

Na ferramenta de gestão dos processamentos, vemos a informação de que um processamento só é adicionado mediante a aplicação de rotinas aos dados guardados na BD, mediante a utilização do programa *Process Data*. O modo de remoção de processamentos é semelhante ao já descrito para pacientes e exames (ver Fig. 116), com a diferença que neste caso não são eliminadas quaisquer outro tipo de informações para além dos processamentos em si. Há também uma opção neste menu que permite a remoção de todos os processamentos efectuados a um dado exame. A selecção do botão *Remove All The Processing associated with an Exam* disponibiliza o menu da Fig. 117, que apresenta uma tabela com os identificadores e as directorias de todos os exames que têm a si associados algum tipo de processamento. Devem seleccionar-se para a lista vazia aqueles aos quais se deseja remover todos os processamentos que lhes foram efectuados, e depois pressionar o botão *Remove Processing associated with these Exams*.

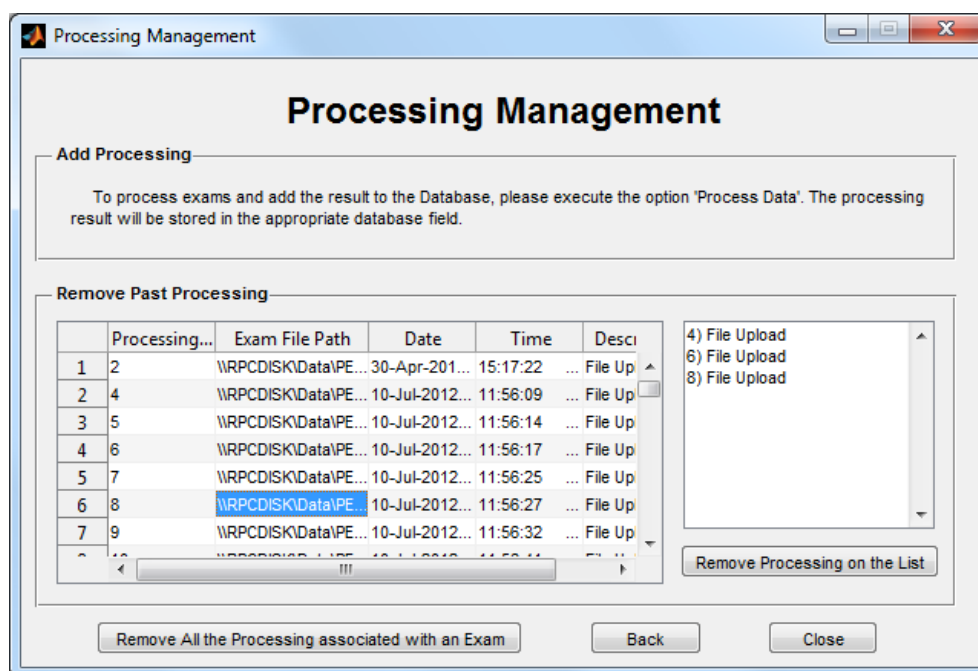


Figura 116 - Menu de gestão dos processamentos efectuados aos dados.

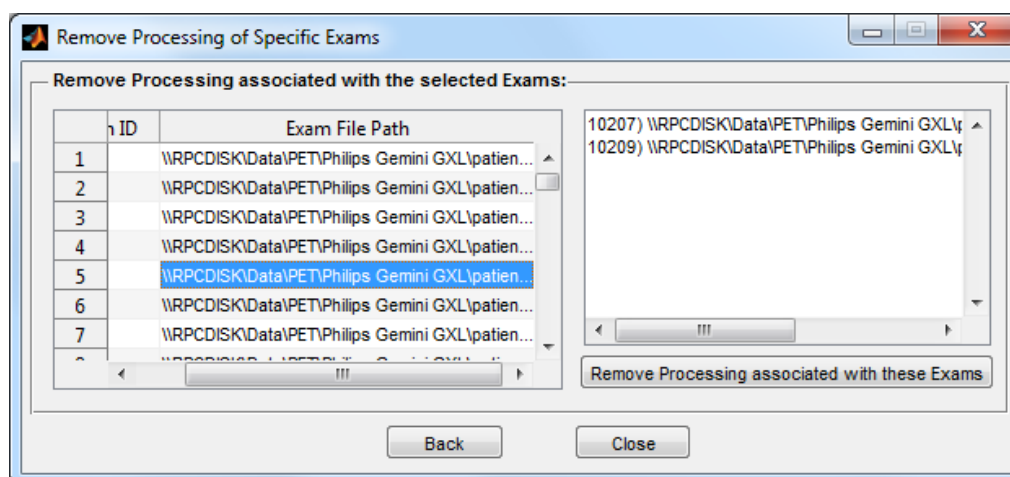


Figura 117 - Janela que permite a remoção de todos os processamentos realizados a um dado conjunto de exames.

Por fim, o Administrador dispõe da ferramenta de adição e remoção de rotinas que podem ser aplicadas aos dados da BD para processá-los. O procedimento de remoção de uma rotina funciona da mesma forma que o descrito para os casos anteriores, pelo que se vai focar o processo de adição daqui em diante. Quando o Administrador tenciona implementar uma nova rotina na BD, a primeira acção que deve realizar é colocar o ficheiro MATLAB[®] onde esta se encontra armazenada na pasta de trabalho da interface ('Medical Image Database'). Depois deve inserir o nome da rotina (sem a extensão '.m') no campo a ele destinado no menu apresentado na Fig. 118. Caso já exista uma rotina com esse nome na BD ela não poderá ser inserida, e será mostrada a mensagem de erro apresentada na Fig. 119. Caso se esteja a tentar inserir uma rotina

sem ter colocado previamente o ficheiro respectivo na pasta já mencionada, é também mostrada uma mensagem de erro ao utilizador, que podemos ver na Fig. 120.

Routines Management

Add New Routine

Routine Name:

Routine Description:

Input Variables Description

Number of Input Variables:

Output Variables Description

Number of Output Variables:

Change the variables order in the callback, if necessary:

Remove Routines

	Routine	Functionality
1		
2		
3		
4		

Figura 118 - Menu que permite a adição de novas rotinas e a remoção destas da BD.

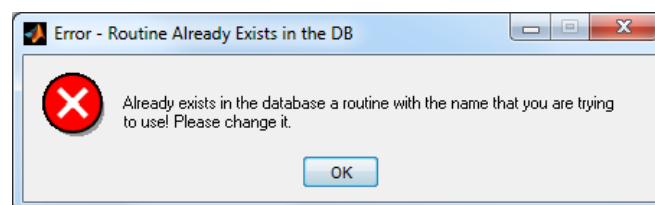


Figura 119 - Mensagem de erro mostrada ao utilizador quando ele tenta inserir uma rotina cujo nome já existe nos registos da BD.



Figura 120 - Mensagem de erro mostrada ao utilizador quando ele tenta inserir uma rotina na BD, sem ter colocado previamente o ficheiro '.m' que lhe corresponde na directoria da interface.

No caso de não haver qualquer problema com o nome inserido, então será disponibilizado um campo onde se pode inserir a descrição do tipo de processamento desempenhado por essa rotina, e o painel a partir do qual devem inserir as informações relativas às variáveis de entrada da rotina fica activo (Fig. 121). O Administrador deve inserir o número de variáveis de entrada, e pressionar o botão *Enter* do teclado. Caso esse seja um número válido (maior que zero), é activado o botão (*Add Input Variables* – Fig. 122) que irá accionar um menu de descrição das variáveis de entrada tantas vezes quantas o número de variáveis inserido (caso se trate de um número fraccionário será convertido automaticamente para um inteiro; no caso de não ser maior que zero será mostrada a mensagem de erro da Fig. 123). Esse menu de descrição das variáveis é mostrado na Fig. 124. Nele, o utilizador deve começar por declarar se a variável de entrada para a rotina, que está a descrever no momento, consta da BD ou é uma variável externa a esta. Deve notar-se que se definiu como obrigatória a utilização de pelo menos uma variável da BD como variável de entrada para uma qualquer rotina de processamento de dados que se adicione. Caso contrário não se estariam a processar os dados da BD, mas sim outros quaisquer, o que poderia corromper a natureza da nossa BD, que se associa à imagem médica. Imaginemos então que a opção escolhida para a primeira variável de entrada foi a de se tratar de uma variável da BD (Fig. 125). Assim sendo, fica activo um novo painel onde se pode escolher uma de entre todas as variáveis inseridas na BD, e para auxiliar essa escolha, é mostrada a definição da variável que foi seleccionada no momento num caixa de texto abaixo do menu *pop-up* que permite a escolha. Definindo-se a segunda variável como externa à BD, é activado o terceiro painel do menu e causa, podendo-se introduzir o nome e o tipo de dados (permite-se a escolha entre um número, uma *String*, um vector ou matriz de números ou caracteres e a directoria de um dado ficheiro) da variável em questão (Fig. 126).

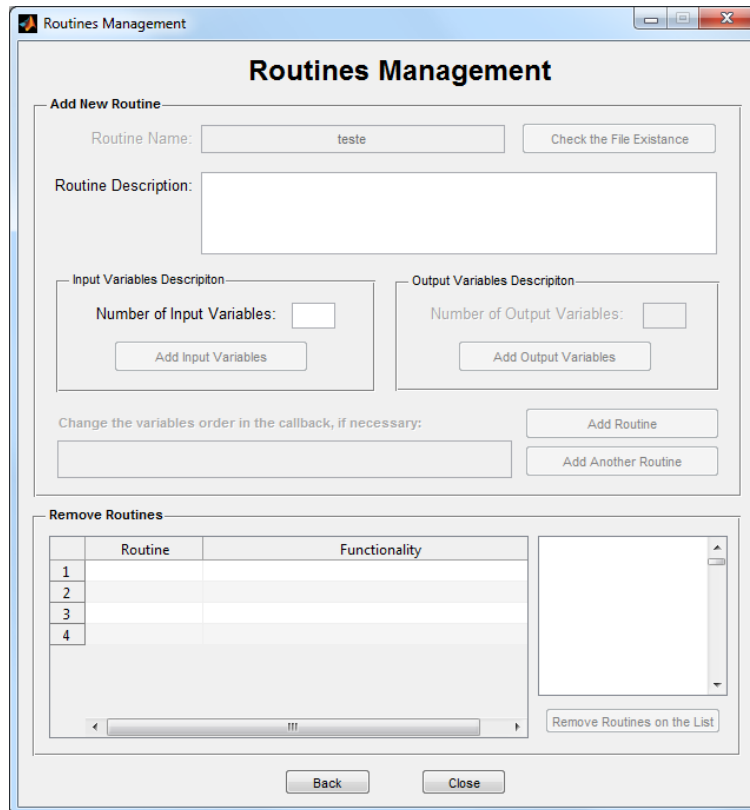


Figura 121 - Disponibilização do campo de descrição da rotina, e do campo para inserir o número de variáveis de entrada que esta utiliza.

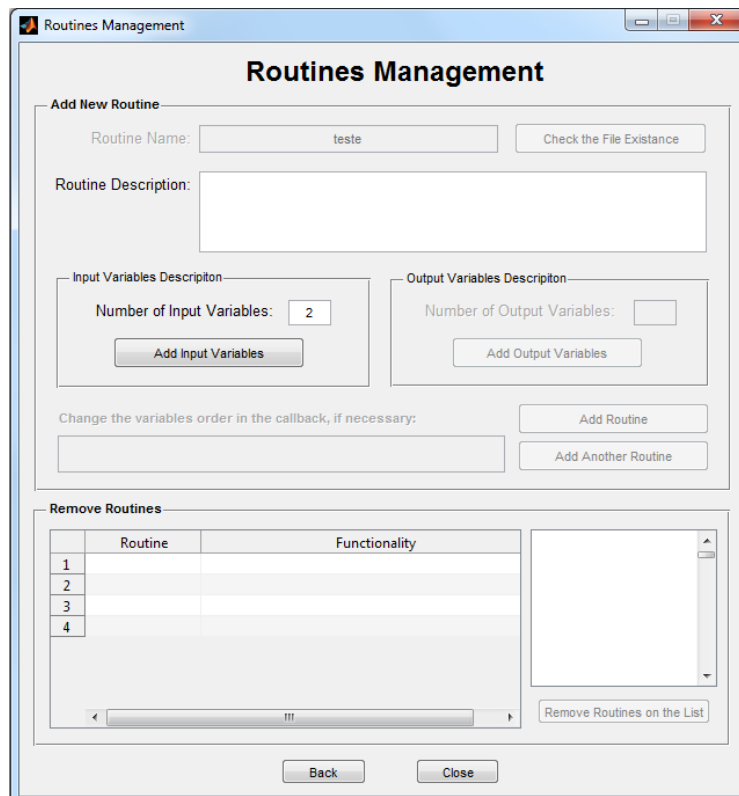


Figura 122 - Disponibilização do botão *Add Input Variables*, após a inserção de um número válido de variáveis de entrada.

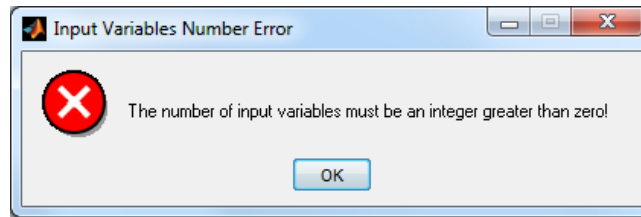


Figura 123 - Mensagem de erro mostrada quando o Administrador tenta inserir um número menor ou igual a zero de variáveis.

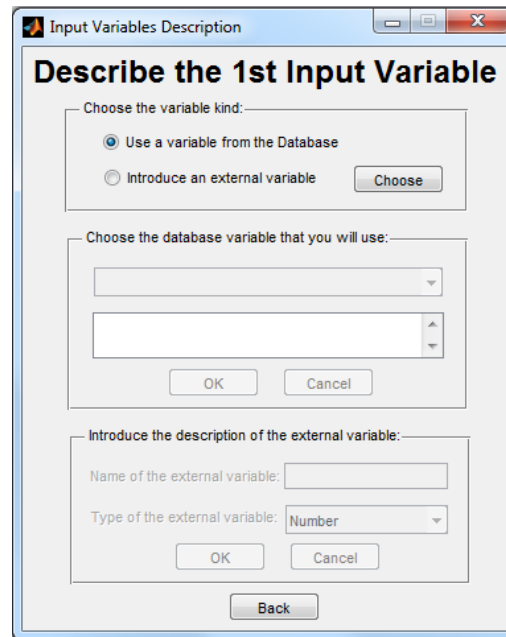


Figura 124 - Menu de descrição das variáveis de entrada.

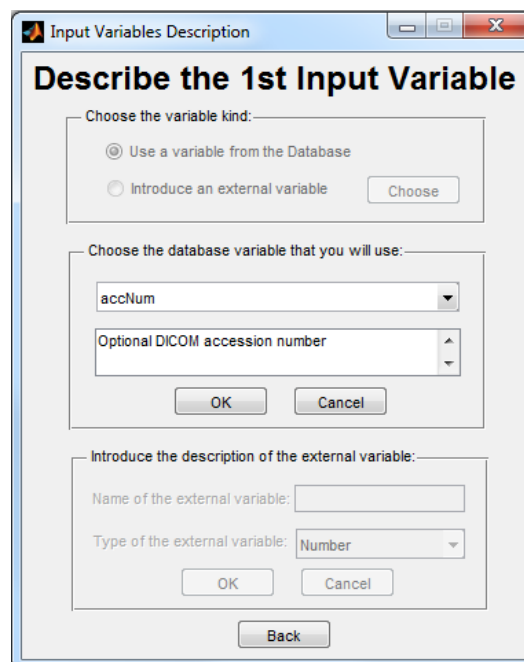


Figura 125 - Descrição da primeira variável de entrada como sendo uma das variáveis integrantes da BD.

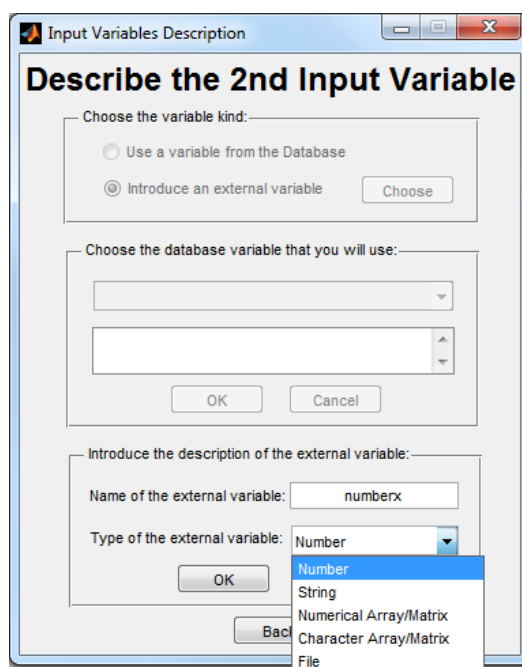


Figura 126 - Aspecto do menu quando se pretende descrever uma variável de entrada que não faz parte da BD. Mostra-se o pormenor do tipo de dados que se pode atribuir a essa variável.

Caso se interrompa a operação de descrição das variáveis (seleccionando o botão *Back* no menu anterior), sem descrever todas as variáveis de entrada, será apresentada a mensagem de erro da Fig. 127, a operação será cancelada, e o utilizador terá de inserir as descrições todas de novo.

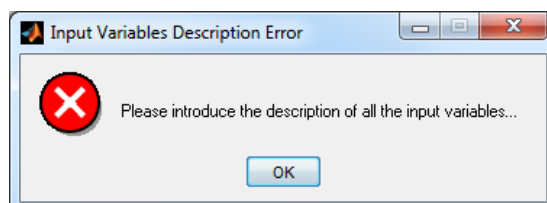


Figura 127 - Mensagem de erro mostrada ao Administrador no caso de ele interromper a operação de descrição de variáveis de entrada.

Se as variáveis de entrada forem todas descritas, então é disponibilizado no menu principal de gestão de rotinas o campo para inserir o número de variáveis de saída da rotina (Fig. 128). Esse número deve ser maior ou igual a zero, pois o *output* da rotina pode ser apenas uma interface de visualização de dados, ou alguma ferramenta do género (não haver uma variável em concreto a guardar os resultados da rotina). Caso esta restrição não seja respeitada é mostrada a mensagem de erro presente na Fig. 129. Depois o utilizador deve novamente pressionar *Enter* e ficará disponível o botão que permite a descrição de cada uma das variáveis de saída (Fig. 130). Ao ser pressionado, primeiro mostra um menu que permite a escolha da

entidade na qual os resultados da rotina serão armazenados (Fig. 131; caso se cancele esta operação sem selecionar uma entidade, é mostrada a mensagem de erro na Fig. 132), depois mostrará o menu da Fig. 133 tantas vezes quantas o número de variáveis de saída introduzido (caso seja zero, esta operação não ocorre). Se a operação de introdução das variáveis for cancelada, antes de todas as variáveis de saída serem descritas, é mostrada uma mensagem de erro semelhante à da Fig. 127 mas adaptada para o caso das variáveis de saída. No menu de descrição das variáveis de saída deve introduzir-se o nome, o tipo de dados e a descrição da variável (os tipos de dados disponibilizados são os mesmos que os apresentados na Fig. 126).

	Routine	Functionality
1		
2		
3		
4		

Figura 128 - Disponibilização do campo para inserir o número de variáveis de saída.

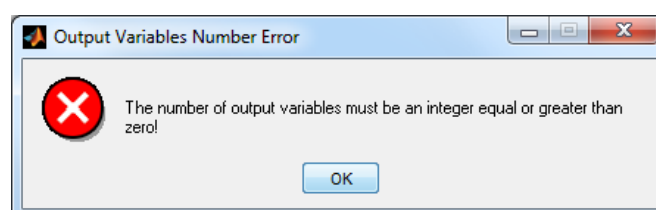


Figura 129 - Mensagem de erro mostrada ao utilizador, caso ele não insira um número válido de variáveis de saída.

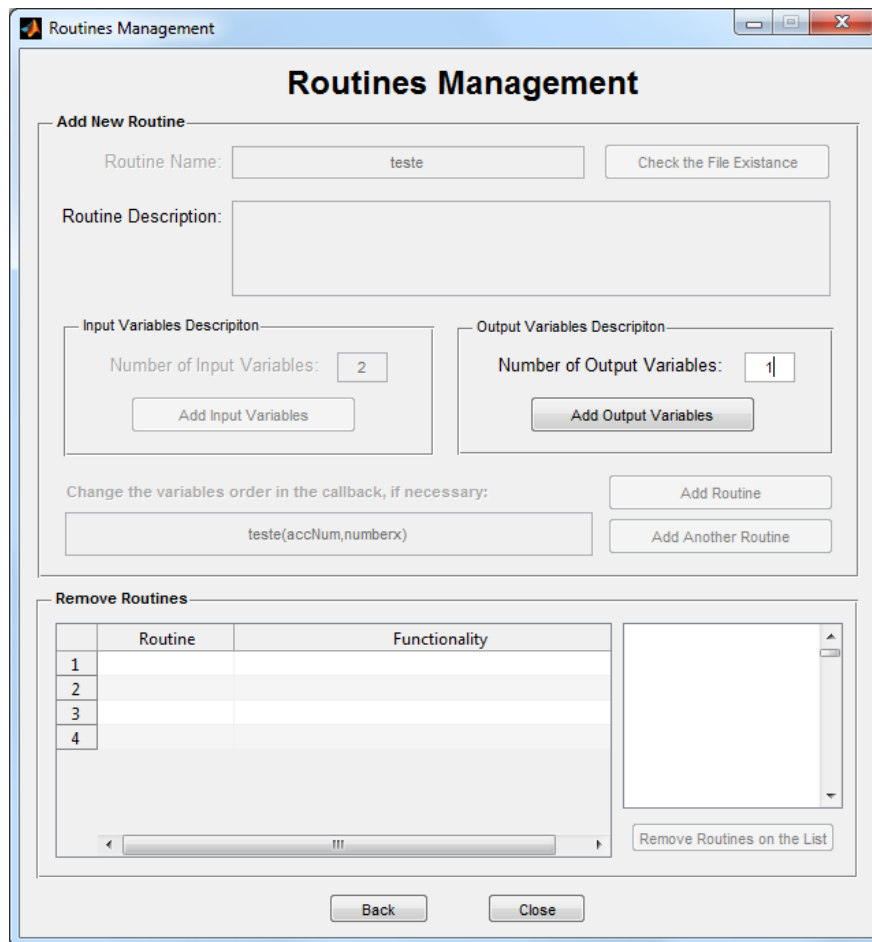


Figura 130 - Activação do botão de descrição das variáveis de saída.

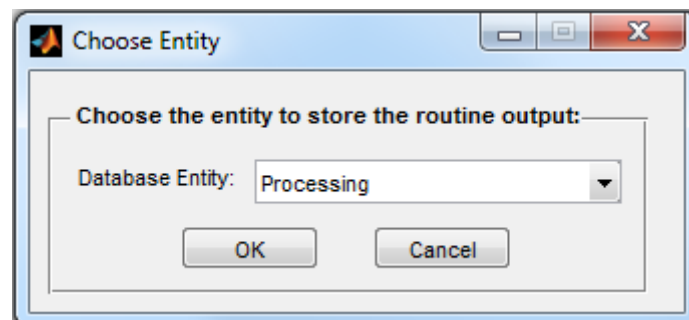


Figura 131 - Menu que permite a escolha da entidade à qual se associarão os resultados da rotina. Permite a associar esses dados às entidades: 'Processamento', 'Paciente', 'Estudo' e 'Corte'.

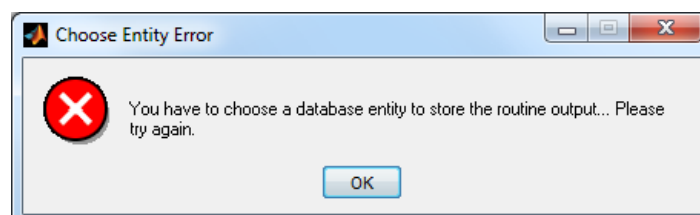


Figura 132 - Mensagem de erro mostrada ao utilizador quando a operação de escolha de uma entidade para guardar os dados que resultam da rotina é cancelada.

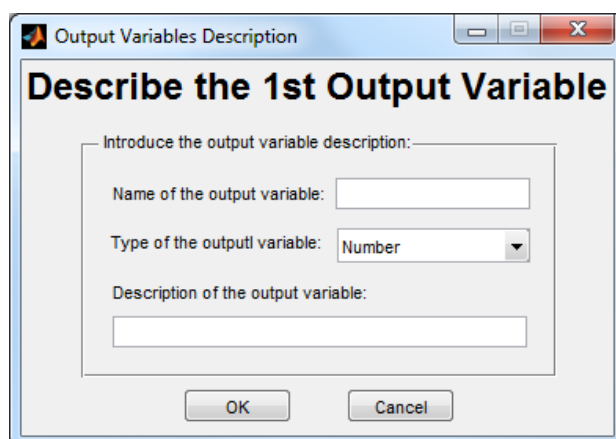


Figura 133 - Menu de descrição das variáveis de saída.

À medida que se foram descrevendo as variáveis de entrada e de saída da rotina que está a ser introduzida, foi-se formando um *callback* (um modo de chamar) essa função. Ao fim da introdução de todas as variáveis de saída, o campo onde esse *callback* foi criado fica activo (Fig. 134) para que o Administrador possa editá-lo (corrigir possíveis erros, ou alterar a ordem de chamamento das variáveis introduzidas anteriormente).

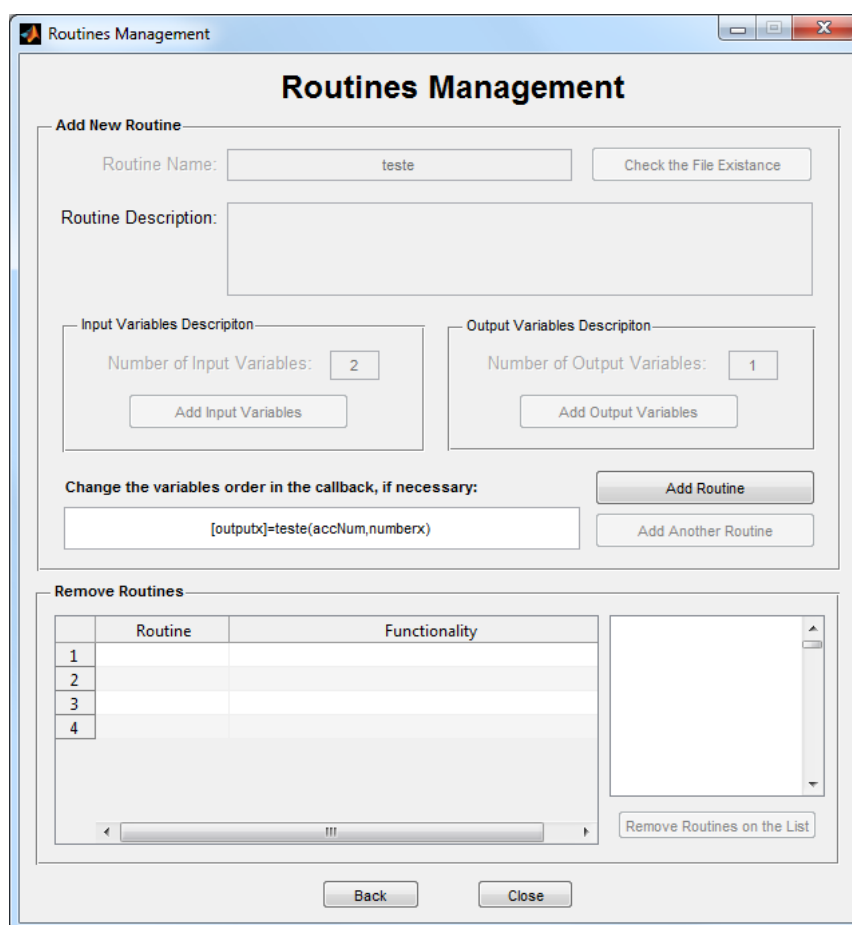


Figura 134 - Activação do campo onde o *callback* da função foi gerado, para que possa ser editado.

Por fim, o Administrador deve pressionar o botão *Add Routine* para adicionar a rotina à BD. Se não houver nenhum problema com o *callback* da função, é mostrada uma mensagem a dar conta do sucesso da operação (Fig. 135). Caso existam problemas ao nível do número de variáveis nele presentes, ou estrutura da sua sintaxe é mostrada a mensagem de erro da Fig. 136.

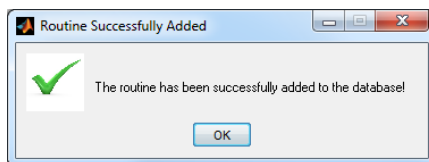


Figura 135 - Mensagem que informa o utilizador do sucesso da inserção da rotina na BD.

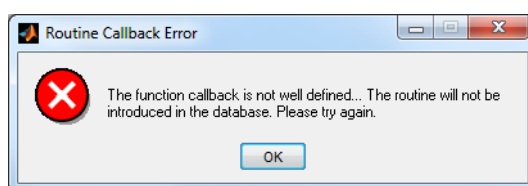


Figura 136 - Mensagem de erro mostrada quando se detecta algum problema ao nível da sintaxe e do número de variáveis presentes no *callback*.

Após a correcta introdução de uma rotina, fica activo o botão que permite adicionar outra nova rotina (Fig. 137). Se premido, limpa as informações anteriormente introduzidas no menu e prepara-o para receber novas.

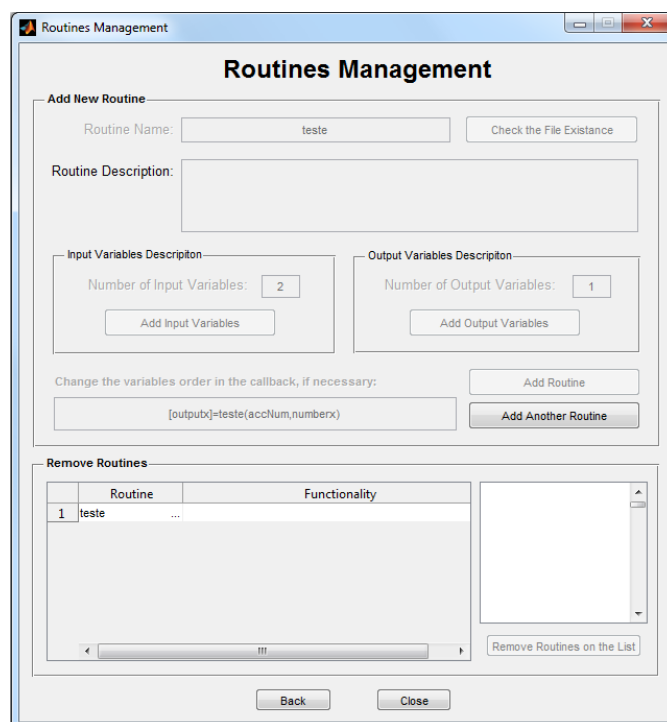
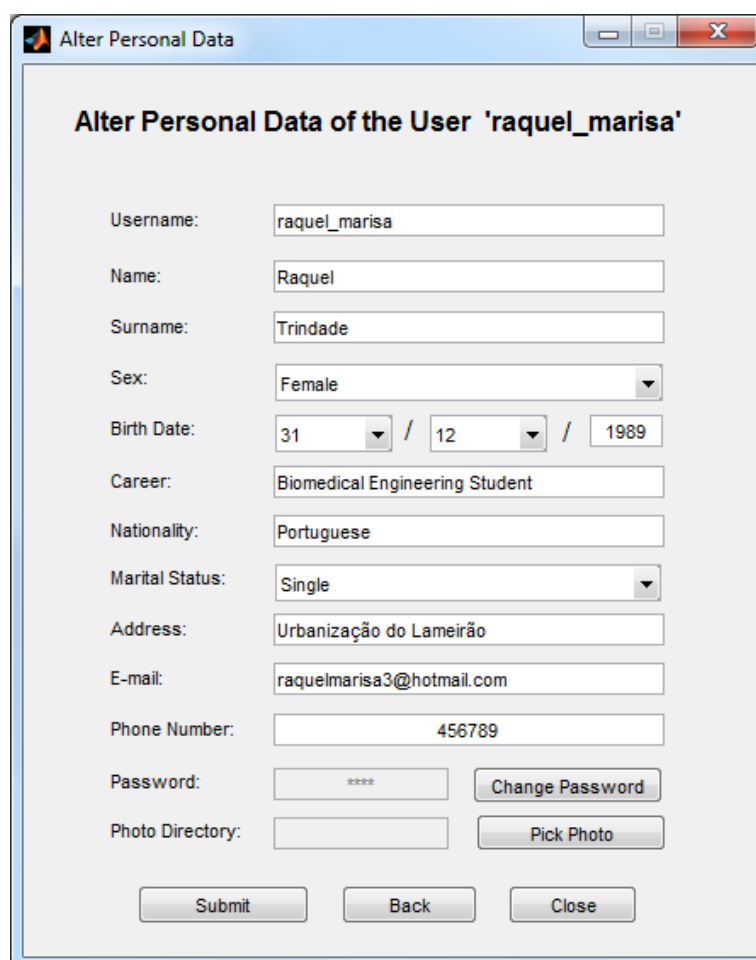


Figura 137 - Activação do botão que permite adicionar uma nova rotina (*Add Another Routine*).

8.2.6. Menu que Permite a Alteração da Informação Pessoal do Utilizador

Por fim, a última funcionalidade que a plataforma desenvolvida fornece, é um menu onde cada utilizador da BD, independentemente do seu perfil, pode alterar os dados pessoais que lhe estão associados (Fig. 138). Primeiro pode mudar o seu nome de utilizador, desde que escolha um que ainda não se encontre registado na BD, e que não contenha espaços (se estas regras não forem cumpridas, são mostradas as mensagens de erro apresentadas nas figuras 94 e 95, respectivamente). Definiram-se os seguintes aspectos: os géneros disponíveis são masculino e feminino; no dia do mês apenas é possível escolher entre 1 e 31, e nos meses do ano entre 1 e 12; os anos válidos para introdução de data de nascimento, dando já uma margem de erro muito grande, seriam entre 1900 e o ano corrente de alteração/introdução dos dados (caso se introduza um ano não incluído neste intervalo é mostrada a mensagem de erro da Fig. 139); os estados civis disponibilizados são solteiro(a), casado(a), divorciado(a) e viúvo(a); os campos que não sejam preenchidos, serão automaticamente gravados com o valor '0'.



The screenshot shows a web application window titled "Alter Personal Data" for the user "raquel_marisa". The window contains a form with the following fields and controls:

- Username: raquel_marisa
- Name: Raquel
- Surname: Trindade
- Sex: Female (dropdown menu)
- Birth Date: 31 / 12 / 1989 (date picker)
- Career: Biomedical Engineering Student
- Nationality: Portuguese
- Marital Status: Single (dropdown menu)
- Address: Urbanização do Lameirão
- E-mail: raquelmarisa3@hotmail.com
- Phone Number: 456789
- Password: **** (password field) with a "Change Password" button
- Photo Directory: (empty field) with a "Pick Photo" button

At the bottom of the form, there are three buttons: "Submit", "Back", and "Close".

Figura 138 - Menu onde o utilizador comum pode alterar (ou adicionar caso ainda não o tenha feito) os seus dados pessoais que se encontrarão registados na BD.

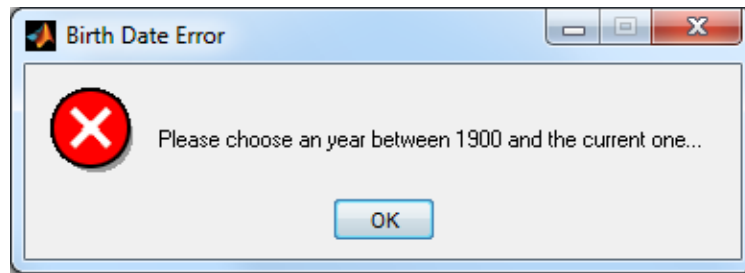


Figura 139 - Mensagem de erro mostrada quando se tenta introduzir uma data de nascimento descabida para um dos utilizadores da BD (note-se o ano de criação da BD: 2011).

Implementou-se um pequeno sub-menu que permite mudar a *password* (Fig. 140), em que o utilizador deve inserir a nova palavra-chave e depois confirmá-la. Quando essa confirmação não está correcta, é mostrada a mensagem de erro na Fig. 141.

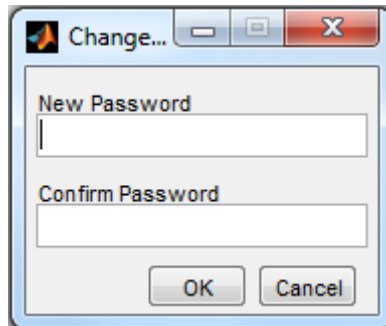


Figura 140 - Janela que permite a alteração da *password*.

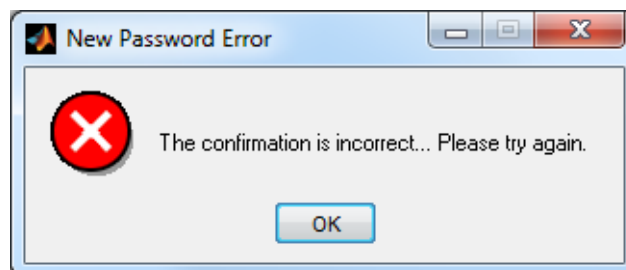


Figura 141 - Mensagem de erro apresentada ao utilizador quando a confirmação da nova palavra-chave não foi bem sucedida.

Por fim, para a selecção da foto de perfil do utilizador, este deve pressionar o botão *Pick Photo* e seleccionar uma imagem existente numa qualquer directoria do computador onde está a trabalhar (Fig. 142). Uma cópia dessa imagem será guardada na pasta de trabalho da interface, com o seu nome de utilizador e a extensão 'jpg'.

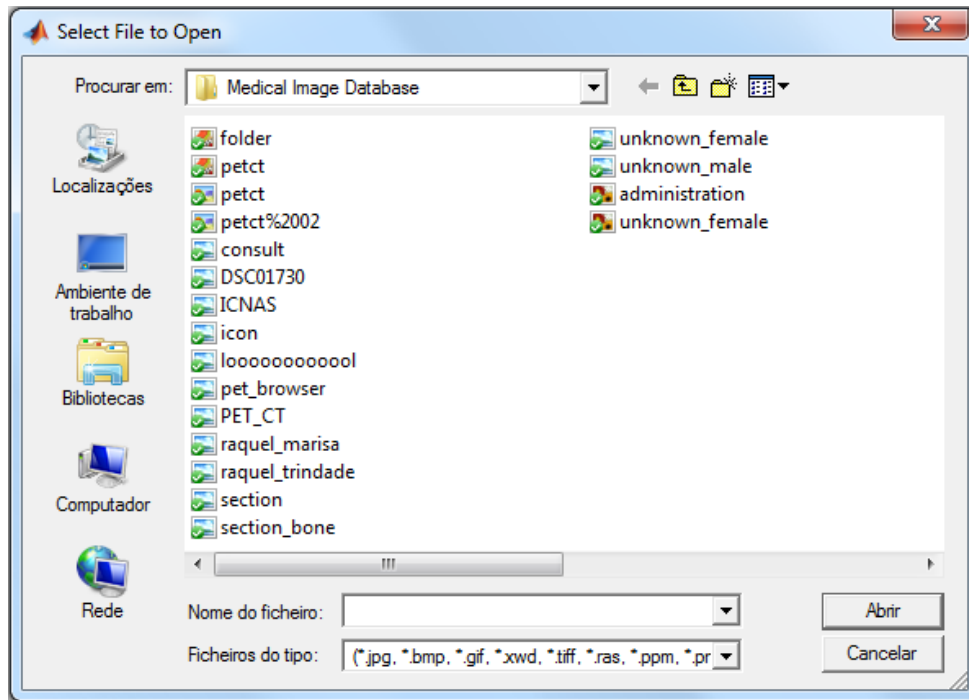


Figura 142 - Janela que permite a selecção da foto de perfil do utilizador.

No fim de alterar todos os dados que lhe interessam, o utilizador deve premir o botão *Submit* e, se não houver qualquer problemas com os dados introduzidos, eles serão inseridos ou actualizados na BD, e no fim é mostrada uma mensagem a dar conta do sucesso da operação (Fig. 143).

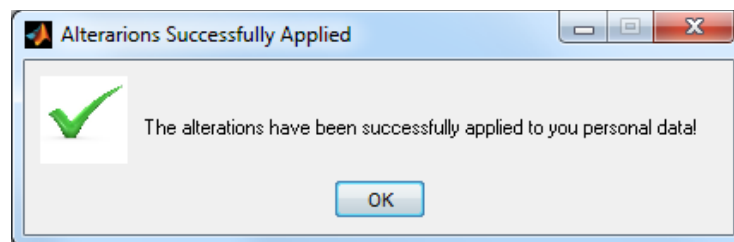


Figura 143 - Mensagem que é apresentada ao utilizador após se efectuarem as alterações por si definidas aos seus dados pessoais.

Capítulo 9

Testes Realizados e Resultados Obtidos

9. Testes Realizados e Resultados Obtidos

Como resultado deste projecto temos uma Base de Dados gerada e devidamente preenchida, associada a uma Interface que permite a visualização dos exames, realizar pesquisas condicionadas (e guardar os seus resultados em ficheiros de texto devidamente formatados e identificados, cujo conteúdo preserva a informação privada dos pacientes). A parte respectiva ao tratamento estatístico dos dados ficou por desenvolver, mas com um espaço reservado para ela na Interface desenvolvida. Esta ferramenta estará sempre aberta à implementação de novas funcionalidades mediante as necessidades da Instituição onde foi desenvolvida.

Uma das ferramentas mais poderosas disponibilizadas pela BD é a possibilidade de integração de outras rotinas desenvolvidas em MATLAB[®] (e futuramente em outras linguagens de programação), possivelmente criadas por outros programadores, para processar os dados armazenados na BD. Assim, por exemplo, poderão ser calculados os resultados de ferramentas de segmentação de imagem que tenham sido integradas no sistema, e os resultados obtidos a partir do seu chamamento destas rotinas serem armazenados no campo gerado para o efeito na BD. Isto permite que sejam efectuadas novas análises aos dados, e incentiva ao desenvolvimento de novas ferramentas de *software*, expandindo assim as funcionalidades/potencialidades da BD.

As funcionalidades de processamento de imagem estão ainda em desenvolvimento mas, neste estágio inicial, permitem ver uma ou várias imagens ao mesmo tempo e realizar algumas acções básicas como o *panning*, escolher a paleta de cores, redimensionar, *zoom in/out* e modificar o contraste e a luminosidade; de futuro ambiciona-se a realização de melhoramento de imagem (com diferentes tipos de filtros), a detecção de objectos (ou simplesmente de formas), a segmentação de imagem, classificação, o co-registo PET/CT e outras tarefas.

Faremos agora uma avaliação do desempenho da ferramenta desenvolvida mediante a comparação do tempo que demoraria proceder à pesquisa e carregamento de dados manualmente (isto é, utilizando somente as ferramentas disponibilizadas pela Philips[®], e não a organização e integração dessas na plataforma desenvolvida) relativamente ao tempo de trabalho do *software* desenvolvido.

Primeiro avaliar-se-á o tempo que leva o carregamento dos ficheiros na BD (processo que não é necessário para a análise manual dos dados). Atenda-se a que, com o processo manual de análise de dados, apenas se procede à abertura e visualização dos dados de um exame de cada vez (usando a rotina 'rdimg.m' definida pela Philips[®], sendo que a versão original desta nem sequer permite o acesso a todos os campos que se encontram parametrizado nos cabeçalhos, como foi explicado na secção 7.3.4), tornando a comparação de dados complicada e morosa. O

facto de se ter todos os dados armazenados num repositório único como uma BD já permite desempenhar as acções anteriormente referidas de forma simples e eficaz. Isto já constitui um motivo de peso para o tempo adicional que se gasta com o procedimento de carregamento de dados na BD. Estima-se que a actualização semanal da BD demore em média $20 \times 0 + 20 + 15 = 35$ a $20 \times 6 + 20 + 15 = 155$ minutos (tendo em conta uma média de realização de 20 exames por semana, estimando-se o tempo de carregamento de uma pasta de exames – que por vezes não contém nenhum dos tipos de ficheiros que são carregados na BD, mas, a grande maioria delas, contém 5 ficheiros ‘.img’ e um ‘.syn’ – entre 0 a 6 minutos, considerando que o processo de verificação de apagamento de ficheiros demore cerca de 20 minutos e que o utilizador perca no máximo 15 minutos no processo de reiniciação do MATLAB[®] após cada carregamento), mas note-se que a ferramenta apenas permite o carregamento de ficheiros descomprimidos, o que significa que toda a informação dela constante está disponível para ser explorada pelo utilizador, enquanto que pelo processo de análise manual, pode ocorrer que o ficheiro que a pessoa tente abrir se encontre comprimido, e este tenha que perder tempo a ir procurar outro que não esteja (e este processo repetir-se várias vezes, o que não é de todo produtivo) ou a utilizar uma forma de proceder à descompressão do ficheiro que escolheu (isto implica que se redireccione o ficheiro para a estação PET, o único sítio onde existe o *software* necessário para proceder à descompressão do ficheiro, e depois se volte a copiá-lo para o local/computador/aplicação de trabalho do utilizador).

Numa situação de pesquisa, no processo manual, o utilizador tem de primeiro procurar nas directorias adequadas o ficheiro que lhe interessa analisar, e depois de o abrir e extrair a informação dos seus cabeçalhos, terá de procurar entre as variáveis estruturadas, qual a informação que lhe interessa avaliar (para poder comparar a informação pesquisada com outra, terá de registar de alguma forma os dados recolhidos e proceder, do mesmo modo, à abertura e análise dos outros ficheiros com os quais desejar comparar a informação recolhida). A plataforma desenvolvida permite automaticamente a pesquisa em todos os ficheiros de exame, e o cruzamento de uma série de características para definir as restrições a aplicar aos dados. Dando um exemplo em concreto, para uma pesquisa de três variáveis em específico, pertencentes a diferentes tabelas da BD, e recolha dos valores dessas em todos os casos (entradas da BD) concordantes com as restrições definidas, conseguiu-se um tempo de pesquisa de menos de 15 minutos (pesquisando nos 10472 exames até agora carregados na BD). Estimou-se o tempo de procura manual dessas três variáveis, após a abertura do primeiro exame, dentro dos 261 campos extraídos dos cabeçalhos, entre 5 (para um utilizador experiente) a 25 minutos. Juntando-lhe o tempo inicial de leitura de um exame (não de uma pasta de exames, relativa a apenas um estudo), estimado entre 20 segundos (aproximadamente, 0,33 minutos) a 1 minuto, e multiplicando este último pelo número de exames K que se pretendam comparar, o

tempo total de análise de dados será no mínimo de $5+0,33 \times K$ minutos e no máximo de $25+1 \times K$ (para comparar com o número de exames existentes na BD, que são tidos em conta em cada pesquisa no caso da utilização da ferramenta desenvolvida, daria um total de $5+0,33 \times 10472=3460,76$ a $25+1 \times 10472=10497$ minutos). Note-se também, que na pesquisa manual, não dá para aplicar directamente restrições aos valores dos dados pesquisados.

Capítulo 10

Conclusão

10. Conclusão

Procedeu-se ao desenvolvimento de uma plataforma para armazenar, integrar e analisar informação de imagem médica, com a possibilidade de partilhá-la de forma segura (e também aos dados resultantes de processamentos adicionais) com outros utilizadores autorizados a aceder e manipular os dados, se for necessário. Estas operações envolvem sempre as restrições necessárias para proteger a informação pessoal dos pacientes.

A um nível institucional num centro de pesquisa imagiológica, a BD permite monitorizar procedimentos (realizando detecção de erros e optimização de procedimentos e protocolos) e ser usada como ferramenta de aprendizagem para novos investigadores, ajudando-os a ganhar experiência, analisando e interpretando muitos exames (porque o grande número de estudos na BD permite aceder a uma grande variabilidade de dados e analisar em detalhe muitas situações específicas diferentes).

À medida que a quantidade de informação de imagiologia médica aumenta, torna-se cada vez mais importante levar a cabo iniciativas para guardar e organizar a informação, de modo a ter sucesso em facilmente extrair conclusões a partir dela e em corrigir, modificar e melhorar procedimentos futuros, facilitando e optimizando o acesso a tais exames médicos assim como a rotina envolvida na sua realização.

Uma ferramenta de pesquisa (com base na linguagem de *queries* de uma BD) bem implementada associada a uma interface gráfica que potencie a sua fácil manipulação e compreensão e que estimule o utilizador a usar a ferramenta, pode abrir novos caminhos na exploração da informação médica, e assim permitir a análise e melhoria dos cuidados de saúde prestados, a evolução da pesquisa biomédica e a educação de novos profissionais de saúde.

10.1. Trabalho Futuro – Perspectivas de Expansão e Integração

De futuro, pretende-se a integração nesta BD não só de imagens de PET/CT mas também de imagens provenientes de outras modalidades praticadas no ICNAS, tais como a Mamografia por Emissão de Positrões (*Positron Emission Mammography* – PEM), a Tomografia Computorizada por Emissão de um Fóton Único (*Single Photon Emission Computed Tomography* – SPECT) e a Ressonância Magnética Nuclear (*Magnetic Resonance Imaging* – MRI). Acrescentando outros tipos de exames, a sua meta-informação e as imagens

propriamente ditas, de forma a se poder integrar a informação dos dois, pode obter-se assim uma melhor qualidade e fiabilidade da análise a ser feita. Considere-se nomeadamente a possibilidade de co-registo de imagens de PET/CT já em análise com imagens de RMN de um mesmo paciente.

Foram dados alguns passos neste sentido. Desenvolveu-se um protótipo do modo como se modificaria o modelo conceptual da BD de forma a acrescentar os exames deste tipo (Fig. 144). Pesquisou-se um pouco sobre o formato em que os ficheiros RMN se encontram, extensão ‘.IMA’ e chegou-se a encontrar uma função em MATLAB® para ler ficheiros nesse formato: ‘ima2mat.m’ [36]. Pesquisaram-se também outras formas de parametrizar o mapa de bytes dos ficheiros neste formato [37], e também mediante ao recurso ao *standard* DICOM [38].

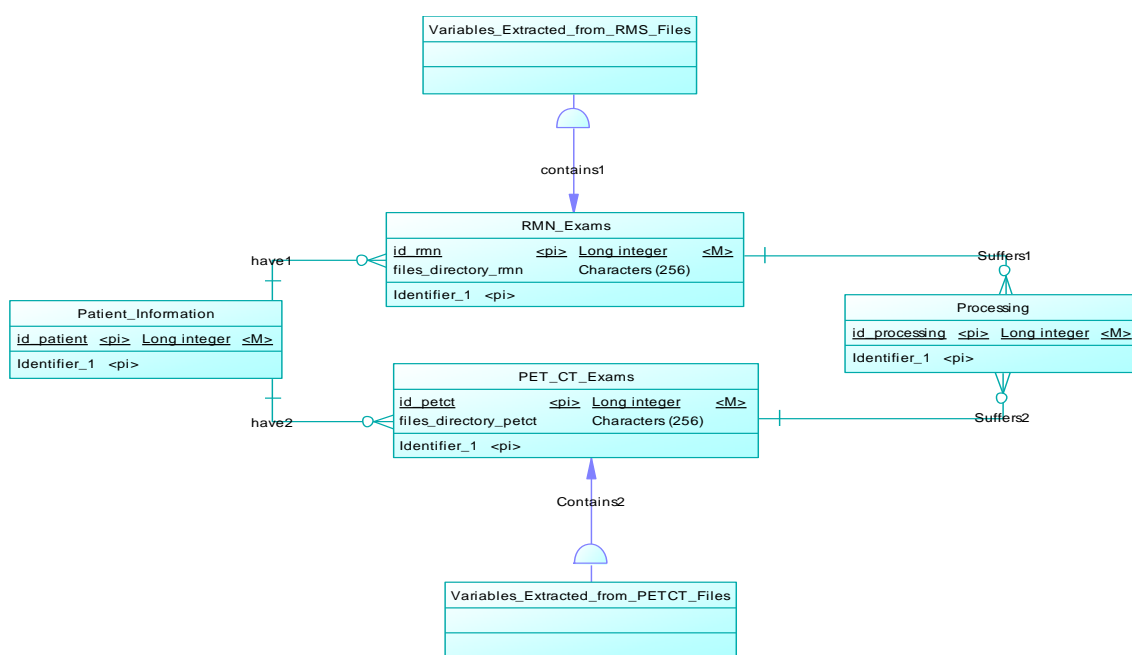


Figura 144 - Possível modelo conceptual relacional para a implementação de uma BD não só de imagens de PET/CT mas também de dados de RMN. Diagrama desenvolvido com auxílio da ferramenta Sybase Power Designer 15®.

Um objectivo ainda mais ambicioso é a integração de imagens de outras instituições, permitindo a comparação e eventual agrupamento de estudos de maneira a realizar análises mais significativas de grupos de pacientes.

Pretende-se adicionar cada vez mais funcionalidades à BD, mediante a adição de novos campos para guardar a informação extraída por rotinas desenvolvidas posteriormente.

Pensou-se em associar-se rotinas com os seguintes fins: sinalizar se os dados pertencentes a um dado exame e armazenados na BD foram analisados por humanos que atestaram a sua coerência e validade (o output desta rotina será guardado num campo que terá o valor 1 caso tenha havido confirmação, e o valor 0 em caso contrário); um programa que

permita declarar a existência de erros em determinadas variáveis (serão assim guardados os nomes das variáveis com erros e os identificadores das entradas em que os erros ocorrem); uma ferramenta de processamento de imagem que permita identificar se se trata de um exame realizado a um fantoma ou a um humano, distinguindo-se no último caso o facto de se tratar de um exame de corpo inteiro ou apenas de uma parte deste (valores de saída: 0 para o caso de um fantoma, 1 para um exame de corpo inteiro a um humano e 2 para apenas uma parte do corpo deste); diversas rotinas de processamento de imagem que permitam a extracção dos vóxeis associados a determinados órgãos do corpo (bexiga, pulmões, coração, cérebro).

Pretende-se também proceder à implementação de uma ferramenta capaz de ler executáveis (por exemplo, programas de extensão ‘.SAV’) desenvolvidos em outras linguagens de programação, através do apelo à execução de uma linha de comandos do Windows[®] internamente no MATLAB[®]. Como foi visto, este tipo de procedimento já foi utilizado com sucesso para o visualizador de ficheiros *imago* desenvolvido em IDL[®].

Futuramente, ir-se-á implementar uma forma de guardar um histórico da actividade do utilizador. A ideia passa pela implementação de ‘log files’ para gravar todas as acções que foram desempenhadas por cada utilizador, podendo assim voltar para trás, e desfazer cada uma delas. Pensa-se também fazer uso das rotinas MATLAB[®] de escrita de ficheiros *imago* (descritas no Anexo E) possivelmente para proceder à alteração dos dados imagiológicos guardados em aquisições passadas, e adicionar os novos ficheiros gerados ao leque de estudos da BD criada. Serão também adicionados novos tipos de visualizadores, com características e funcionalidades diferentes dos dois visualizadores de exames implementados na interface até ao momento (apresentados no Anexo I.2 deste documento).

Por fim, pretende-se implementar a ferramenta para a qual já se deixou um espaço alocado no menu de consulta de dados, que permite a Análise Estatística da Base de Dados. Ela deve realizar vários testes estatísticos, devolvendo vários tipos de gráficos (barras, dispersões, diagramas de bigodes, tridimensionais, etc.). Devido à dimensão da BD, há inúmeras relações que podem ser exploradas, inúmeras contagens realizadas, definição e análise de vários grupos, e só este estudo pode ser matéria-prima para a concepção de uma nova tese de mestrado. Daí se ter optado por deixá-lo em *standby* para implementação futura.



Bibliografia

Bibliografia

- [1] S. T. C. Wong, “Digital Image Management in Biomedicine,” in *Medical Image Databases*, Boston/Dordrecht/London: Kluwer Academic Publishers, 1998, pp. 1–22.
- [2] E. R. da Saúde, Ed., *Estudo sobre a Concorrência no Sector da Imagiologia*. Porto: , 2009, p. 63.
- [3] M. do R. Giraldez, “Avaliação de novas Tecnologias PET/CT Medicina Nuclear,” *Acta Medica Portuguesa*, vol. 23, pp. 291–310, 2010.
- [4] A. Natário, A. Nogueira, A. I. Santos, A. de Albuquerque, A. Ventosa, A. L. Bastos, A. Oliveira, D. Gazzo, F. Ponte, I. Henriksson, J. G. Correia, J. M. P. de Lima, J. G. Pereira, L. Oliveira, M. J. Proença, M. do R. Vieira, M. T. Rézio, M. Ramos, M. Seixas, P. Almeida, and P. Rosário, *Rede de Referência de Medicina Nuclear e Recomendações para a Instalação de Unidades de Ciclotrão*. Lisboa, Portugal: , 2011, p. 105.
- [5] I. P. Instituto Nacional de Estatística and I. P. Instituto Nacional de Saúde Doutor Ricardo Jorge, Eds., *Inquérito Nacional de Saúde 2005/2006*. Lisboa - Portugal: , 2009.
- [6] “Cornell University: Computer Vision and Image Analysis Group - Public Image Databases.” [Online]. Available: <http://www.via.cornell.edu/databases/>. [Accessed: 21-Nov-2011].
- [7] “Image Databases.” [Online]. Available: http://www.imageprocessingplace.com/root_files_V3/image_databases.htm. [Accessed: 22-Nov-2011].
- [8] “SIPI Image Database.” [Online]. Available: <http://sipi.usc.edu/database/>. [Accessed: 22-Nov-2011].
- [9] F. Petitcolas, “Photo Database.” [Online]. Available: http://www.petitcolas.net/fabien/watermarking/image_database/. [Accessed: 22-Nov-2011].
- [10] P. S. Conti and D. K. Cham, Eds., *PET-CT: a case-based approach*. New York, New York, USA: Springer, 2005, p. 304.
- [11] “PET/CT database helps daily clinical practice,” *Diagnostic Imaging*, p. 1187112, 2007.
- [12] “A Comprehensive PET-CT Database for Tracking Patient Variables.” [Online]. Available: <http://www.freeonlineresearchpapers.com/pet-ct-patient-tracking>. [Accessed: 10-Dec-2011].
- [13] P. Nunes, “Conceito de Gestão e de Gestor,” 2006. [Online]. Available: http://www.notapositiva.com/trab_professores/textos_apoio/gestao/01conc_gestao.htm. [Accessed: 20-Jun-2012].
- [14] “ICNAS - Instituto de Ciências Nucleares Aplicadas à Saúde.” [Online]. Available: <http://www.uc.pt/icnas>. [Accessed: 20-Jan-2012].

- [15] “Organization of IBILI.” [Online]. Available: <http://www.uc.pt/en/fmuc/ibili/OrganizationIBILI>. [Accessed: 20-Jun-2012].
- [16] “Recursos do ICNAS.” [Online]. Available: http://www.uc.pt/fctuc/dei/icnas/orgaos_do_icnas/contactos/mapa. [Accessed: 20-Jun-2012].
- [17] “IBILI - Instituto Biomédico de Investigação de Luz e Imagem.” [Online]. Available: <http://www.knownow.com.pt/instituto-biomedico-de-investigacao-de-luz-e-imagem-ibili-universidade-de-coimbra/>. [Accessed: 20-Jan-2012].
- [18] S. Wong and K. S. J. Hoo, “Medical Imagery,” in *Image Databases: Search and Retrieval of Digital Imagery*, vol. 8, L. D. Bergman and V. Castelli, Eds. California: John Wiley & Sons, Inc., 2002, pp. 83–103.
- [19] V. Castelli and L. D. Bergman, “Digital Imagery: Fundamentals,” in *Image Databases: Search and Retrieval of Digital Imagery*, vol. 8, L. D. Bergman and V. Castelli, Eds. New York, New York, USA: John Wiley & Sons, Inc., 2002, pp. 1–10.
- [20] S. D. W. Townsend, “Physical Principles and Technology of Clinical PET Imaging,” *Annals of the Academy of Medicine, Singapore*, vol. 33, no. 2, pp. 133–45, Mar. 2004.
- [21] “Guia PET-CT – Tomografia por Emissão de Positrões com Tomografia Computorizada,” *Hospital da Luz - Departamento de Medicina Molecular*. [Online]. Available: http://www.hospitaldaluz.pt/upload/5/fckeditor_files/file/Teste PDF/HLUZ_MEDMOL_Guia_PET_CT.pdf. [Accessed: 23-Nov-2011].
- [22] N. C. Ferreira, “PET/CT - Princípios Físicos,” in *Jornadas de Radiologia*, 2011, p. 54.
- [23] T. M. Blodgett, B. M. McCook, and M. P. Federle, “Positron emission tomography/computed tomography: protocol issues and options,” *Seminars in nuclear medicine*, vol. 36, no. 2, pp. 157–68, Apr. 2006.
- [24] F. Furtado, Pedro (Departamento de Engenharia Informática, “Notas da Cadeira de ‘Bases de Dados e Análise de Informação’,” Coimbra, 2010.
- [25] R. Elmasri and S. B. Navathe, *Fundamentals of database systems*, Third Edit., vol. 28, no. 5. 2000, p. 869.
- [26] H. Garcia-molina, J. D. Ullman, and J. Widom, *Database Systems: The Complete Book*, Second Edi. Upper Saddle River, New Jersey: Prentice Hall, 2008.
- [27] M. D. Costa, “Plano de Dissertação de Mestrado - Análise de estratégias de persistência de dados para SRES baseados na especificação openEHR, utilizando a modelagem de dados relacional,” Universidade Federal de Pernambuco, 2012.
- [28] R. S. Pressman, “User Interface Design,” in *Software Engineering: a practitioner’s approach*, 5th ed., New York, New York, USA: Mc-Graw Hill, 2001, pp. 401–418.
- [29] T. Mandel, *The Elements of User Interface Design*. Wiley, 1997.
- [30] D. A. Norman, “Cognitive Engineering,” in *User Centered Systems Design*, Lawrence Earlbaum Associates, 1986.

- [31] “GEMINI - Luminary Site Support Package,” no. September. Philips Medical Systems (Cleveland), Inc., Cleveland, Ohio, U.S.A., 2005.
- [32] “GEMINI GXL: Instructions for Use - System Operation.” Philips Medical Systems (Cleveland), Inc., Cleveland, Ohio, U.S.A., 2007.
- [33] A. M. R. da Silva and C. A. E. Videira, *UML, Metodologias e Ferramentas CASE*. Porto: , 2001.
- [34] “Postgresql Data Types and Operators.” [Online]. Available: <http://courses.linuxchix.org/mod/resource/view.php?id=26>. [Accessed: 03-Apr-2012].
- [35] G. H. A. A. Dias, “Capítulo 10 - Modelo Físico dos Dados,” in *Análise de Sistemas I*, 2002.
- [36] “ima2mat.” [Online]. Available: <http://perso.telecom-paristech.fr/~nicolas/MATLAB/ima2mat.m>. [Accessed: 15-Mar-2012].
- [37] Christo and R. Ray, “MATLAB Toolbox for .ima.” [Online]. Available: http://www.mathworks.com/matlabcentral/newsreader/view_thread/15017. [Accessed: 15-Mar-2012].
- [38] “Ima images in matlab.” [Online]. Available: <http://stackoverflow.com/questions/8658280/ima-images-in-matlab>. [Accessed: 15-Mar-2012].

Bibliografia Utilizada na Concepção da Parte Prática deste Trabalho

- [1] “MATLAB GUI (Graphical User Interface) Tutorial for Beginners,” 2010. [Online]. Available: <http://blinkdagger.com/matlab/matlab-gui-graphical-user-interface-tutorial-for-beginners/>.
- [2] A. A. Ameri and R. Y. A. Ashi, “Introduction to Graphical User Interface (GUI) MATLAB 6.5,” UAE UNIVERSITY.
- [3] V. A. Bose-Mullick, *Matlab 6.5: GUI Tutorial*. .
- [4] “SQL Case Insensitive Search.” [Online]. Available: <http://forums.devshed.com/database-management-46/sql-case-insensitive-search-845018.html>. [Accessed: 05-Jul-2012].
- [5] “MATLAB Central - String operation.” [Online]. Available: http://www.mathworks.com/matlabcentral/newsreader/view_thread/275664. [Accessed: 02-Jul-2012].
- [6] “SQL Tutorial.” [Online]. Available: <http://www.firstsql.com/tutor2.htm>. [Accessed: 26-Jun-2012].

- [7] "What is the SQL command to return the field names of a table?" [Online]. Available: <http://stackoverflow.com/questions/100504/what-is-the-sql-command-to-return-the-field-names-of-a-table>. [Accessed: 23-Apr-2012].
- [8] "PostgreSQL - Arrays." [Online]. Available: <http://www.postgresql.org/docs/8.0/interactive/arrays.html>. [Accessed: 15-Apr-2012].
- [9] "Selecting Social Media Monitoring Tools to Manage Your Reputation." [Online]. Available: <http://www.onlinerepmanagement.com/2011/08/social-media-monitoring.html>. [Accessed: 26-Jun-2012].
- [10] "IDL DataMiner Guide: Example." [Online]. Available: <http://star.pst.qub.ac.uk/idl/Example.html>. [Accessed: 13-Nov-2011].
- [11] "How can I access a postgresql database from matlab with without matlabs database toolbox?" [Online]. Available: <http://stackoverflow.com/questions/2698159/how-can-i-access-a-postgresql-database-from-matlab-with-without-matlabs-database>. [Accessed: 20-Jan-2012].
- [12] "MATLAB Central - Postgres JDBC Driver not found/Loaded." [Online]. Available: http://www.mathworks.com/matlabcentral/newsreader/view_thread/295491. [Accessed: 20-Jan-2012].
- [13] "ASCII Chart in Matlab." [Online]. Available: <http://www.matrixlab-examples.com/ascii-chart.html>. [Accessed: 20-Jan-2012].
- [14] "MATLAB Central - Masking typed passwords with asterisks." [Online]. Available: <http://blogs.mathworks.com/pick/2007/10/05/masking-typed-passwords-with-asterisks/>. [Accessed: 20-Mar-2012].
- [15] "MATLAB Central - logindlg.m." [Online]. Available: <http://www.mathworks.com/matlabcentral/fileexchange/28101-biomet-database-tool/content/logindlg.m>. [Accessed: 20-Mar-2012].
- [16] "MATLAB Central - Too many files open." [Online]. Available: http://www.mathworks.com/matlabcentral/newsreader/view_thread/164519. [Accessed: 05-May-2012].
- [17] "MathWorks - Why do I receive 'Too many files open' error messages after several hours when executing code which involves reading text files?" [Online]. Available: <http://www.mathworks.com/support/solutions/en/data/1-5Q6XYM/index.html?product=ML&solution=1-5Q6XYM>. [Accessed: 04-May-2012].
- [18] "Unknown Female Image." [Online]. Available: http://www.google.pt/imgres?q=unknown&hl=pt-PT&biw=1280&bih=657&gbv=2&tbn=isch&tbnid=ajELTdFLRvT3wM:&imgrefurl=http://www.antihairslave.com/2012/01/january-2012-length-check.html&docid=X-Mo78pof6_DyM&imgurl=http://4.bp.blogspot.com/-PHxUoopaCsU/Tw3xR7Gox4I/AAAAAAAAAEF0/G6_OtiYZILM/s1600/length%2Bunknown.png&w=582&h=597&ei=_Z1UT5XJO4Ox8gPcpI3wBQ&zoom=1&iact=rc&dur=625&sig=104846480102772555733&page=1&tbnh=134&tbnw=131&start=0&ndsp=20&ved=1t:429,r:8,s:0&tx=50&ty=55. [Accessed: 03-Feb-2012].

- [19] “Unknown Male Image.” [Online]. Available: http://www.google.pt/imgres?q=unknown&hl=pt-PT&biw=1280&bih=657&gbv=2&tbn=isch&tbnid=rYdkxXZQB5G5gM:&imgrefurl=http://mp3.com/artist/Unknown&docid=XvjoYRg6TYjiKM&imgurl=http://userserve-ak.last.fm/serve/_/58531987/Unknown%2B_user.jpg&w=350&h=340&ei=_Z1UT5XJO4Ox8gPcpI3wBQ&zoom=1&iact=hc&vpx=1005&vpy=170&dur=8578&hovh=221&hovw=228&tx=142&ty=127&sig=104846480102772555733&page=1&tbnh=134&tbnw=134&start=0&ndsp=20&ved=1t:429,r:6,s:0. [Accessed: 03-Apr-2012].
- [20] “MATLAB Central - problem with listbox.” [Online]. Available: http://www.mathworks.com/matlabcentral/newsreader/view_thread/146698. [Accessed: 04-Apr-2012].
- [21] “Creating Graphical User Interfaces - Loading the List Box.” [Online]. Available: http://www.weizmann.ac.il/matlab/techdoc/creating_guis/ch_appl7.html#17499. [Accessed: 04-Apr-2012].
- [22] “Administrative Tools Image.” [Online]. Available: <http://mintywhite.com/xp/xsecurity/enable-hidden-administrator-account-windows-xp/>. [Accessed: 05-Apr-2012].
- [23] B. R. Hunt, R. L. Lipsman, and J. M. Rosenberg, “A Guide to MATLAB for Beginners and Experienced Users.” [Online]. Available: <http://books.google.pt/books?id=NekrOjT8oSIC&printsec=frontcover&dq=A+Guide+to+MATLAB:+For+Beginners&hl=pt-PT&sa=X&ei=UFB4T8evKOMw0AWp4vW8DQ&ved=0CDEQ6AEwAA#v=onepage&q=A+Guide+to+MATLAB:+For+Beginners&f=false>. [Accessed: 23-Nov-2011].

Anexos Confidenciais