



Universidade de Brasília

Instituto de Ciências Exatas  
Departamento de Ciência da Computação

# Knowledge-Based Planning para automação de planejamentos de Radioterapia Hipofracionada de Próstata

José Eduardo Vaz Nascimento

Dissertação apresentada como requisito parcial para  
conclusão do Mestrado Profissional em Computação Aplicada

Orientador  
Prof. Dr. Marcelo Ladeira

Brasília  
2021

Ficha catalográfica elaborada automaticamente,  
com os dados fornecidos pelo(a) autor(a)

NN244k Nascimento, Jose Eduardo Vaz  
Knowledge Based Planning para automação de planejamentos  
de Radioterapia Hipofracionada de Próstata / Jose Eduardo  
Vaz Nascimento; orientador Marcelo Ladeira. -- Brasília,  
2021.  
74 p.

Dissertação (Mestrado - Mestrado Profissional em  
Computação Aplicada) -- Universidade de Brasília, 2021.

1. radioterapia. 2. knowledge-based planning. 3.  
aprendizado de máquina. 4. VMAT. 5. próstata. I. Ladeira,  
Marcelo, orient. II. Título.



Universidade de Brasília

Instituto de Ciências Exatas  
Departamento de Ciência da Computação

# Knowledge-Based Planning para automação de planejamentos de Radioterapia Hipofracionada de Próstata

José Eduardo Vaz Nascimento

Dissertação apresentada como requisito parcial para  
conclusão do Mestrado Profissional em Computação Aplicada

Prof. Dr. Marcelo Ladeira (Orientador)  
CIC/UnB

Prof. Dr. Fábio Ynoe de Moraes   Prof. Dr. Guilherme Souza Rodrigues  
Queen's University   EST/UnB

Prof. Dr. Gladston Luiz da Silva  
EST/UnB

Prof. Dr. Marcelo Ladeira  
Coordenador do Programa de Pós-graduação em

Brasília, 15 de dezembro de 2021

# Dedicatória

Dedico este trabalho à minha família:

Meu pai, Antonio Carlos; minha mãe, Tania... Exemplos GIGANTES de seres humanos!!

Minhas irmãs, companheiras de vida.

À minha esposa, Rosana, pelo incentivo e companheirismo de sempre...

*"O que faz andar a estrada? É o sonho. Enquanto a gente sonhar a estrada permanecerá viva. É para isso que servem os caminhos, para nos fazerem parentes do futuro"*

Mia Couto

# Agradecimentos

Em primeiro lugar agradeço a Deus. Por tudo.

À minha família, pelo apoio incondicional.

À Universidade de Brasília e ao PPCA, pelo oferecimento do programa de pós-graduação.

Ao Professor Marcelo Ladeira, que inicialmente abriu as portas do PPCA e depois me acolheu e incentivou para que eu concluísse o programa.

Ao Professor Donald Pianto, que me guiou e incentivou no meio do caminho.

A todos os colegas do PPCA.

À banca examinadora, que muito contribuiu com o trabalho.

Ao Hospital-Sírio Libanês, por incentivar e permitir o desenvolvimento do projeto.

A todos os amigos e colegas de trabalho, que permitiram a minha ausência em alguns períodos no trabalho e apoiaram o projeto.

A todos que participaram, direta ou indiretamente, do desenvolvimento do projeto e suas discussões: meu sincero agradecimento.

# Resumo

Radioterapia envolve o uso de radiação ionizante para tratamento de tumores. Uma etapa importante, prévia à entrega de dose de radiação, é o planejamento do tratamento, o qual varia em complexidade conforme a técnica utilizada e a região a ser irradiada. Uma abordagem que tem se consolidado na radioterapia de próstata em regime hipofracionado é a utilização de arcos volumétricos modulados, pois possibilita alta conformidade da dose de prescrição nos volumes alvos e um tempo de tratamento encurtado, quando comparado com outras técnicas que utilizam modulação do feixe. No entanto, a etapa de planejamento, em geral, é a que mais demanda tempo de profissionais não médicos e a sua qualidade depende da habilidade do planejador, de modo que pode haver variação da qualidade do tratamento entregue em função do grau de experiência do planejador. Neste trabalho criou-se uma aplicação acoplada ao sistema de planejamento que, tendo-se um conjunto de imagens de tomografia computadorizada com os contornos dos volumes alvos e os volumes de órgãos sadios, automatiza a etapa de planejamento de próstata com o uso de aprendizado de máquina. Baseados em *Knowledge-Based Planning*, os algoritmos, quando aplicados ao histogramas de dose-volume dos casos utilizados para treinamento, predizeram novos histogramas com correlação de Pearson de 0.999, tendo-se as áreas dos histogramas como parâmetro. O aplicativo foi testado em 10 novos casos e comparado com o planejamento de 4 físicos especialistas, sendo que em 9 dos 10 casos apresentou qualidade semelhante aos planejamentos dos especialistas e atingiu os objetivos definidos. Assim, este aplicativo torna o fluxo de um Departamento de Radioterapia mais eficiente, o que aumenta a capacidade de planejamento sem a necessidade de novos investimentos; e, ao padronizar e automatizar esta etapa, auxilia na manutenção de um padrão institucional de planejamentos.

**Palavras-chave:** radioterapia, knowledge-based planning, aprendizado de máquina, VMAT, prostata

# Abstract

Radiotherapy involves the use of ionizing radiation to treat malignant tumors. An important step prior to the delivery of radiation dose, is the treatment planning, which varies in complexity according to the technique used and the region to be irradiated. One approach that has been consolidated in prostate radiotherapy in a hypofractionated regimen is the use of modulated volumetric arcs, as it enables high conformity of the prescription dose in target volumes and a shortened treatment time, when compared to other techniques that use beam intensity modulation. However, the planning stage, in general, is time consuming and the one that demands more time from non-medical professionals. Also, its quality is planner-dependent, so that there may be variation in the quality of the treatment delivered depending on the planner's level of experience. In this work, an application integrated to the treatment planning system was created that, having a volume of computed tomography images with the contours of the target volumes and the volumes of adjacent healthy organs, automates the prostate planning stage with the use of machine learning. Made from *Knowledge-Based Planning*, the algorithm, when applied to the dose-volume histograms of the cases used for training, predicted new histograms with a Pearson correlation of 0.999, using the areas of the histograms as parameters. The application was tested in 10 new cases and compared with the plans of 4 specialist physicists, where in 9 of the 10 cases it had similar quality to the specialists' plans and achieved the given plan objectives. Thus, this application makes the workflow of a Radiation Oncology Department more efficient, which increases the planning capacity without the need for new investments; and, by standardizing and automating this step, it helps to maintain an institutional planning pattern.

**Keywords:** radiotherapy, knowledge-based planning, machine learning, VMAT, prostate

# Sumário

<b>1</b>	<b>Introdução e Objetivos</b>	<b>1</b>
1.1	Objetivos . . . . .	3
1.1.1	Objetivos Específicos . . . . .	3
1.2	Capítulos subsequentes . . . . .	3
<b>2</b>	<b>Fundamentação Teórica</b>	<b>4</b>
2.1	Radioterapia . . . . .	4
2.2	VMAT e Otimização . . . . .	6
2.3	Radioterapia Hipofracionada de Próstata . . . . .	7
2.4	Knowledge Based Planning . . . . .	9
2.5	Regressão Linear Múltipla . . . . .	12
2.6	<i>Random Forest</i> . . . . .	12
2.7	<i>Support Vector Machine - SVM</i> . . . . .	12
2.8	<i>Grid_Search</i> . . . . .	12
2.9	<i>PCA</i> . . . . .	14
<b>3</b>	<b>Revisão da Literatura</b>	<b>15</b>
<b>4</b>	<b>Metodologia</b>	<b>23</b>
4.1	Dados . . . . .	23
4.2	Metodologia . . . . .	24
4.2.1	Automação do Planejamento . . . . .	26
4.3	Ferramentas . . . . .	29
4.3.1	Formato de arquivo DICOM . . . . .	29
4.3.2	Formato de arquivo DICOM-RT . . . . .	29
4.3.3	Biblioteca Pydicom . . . . .	30
4.3.4	Biblioteca Scikit-Learn . . . . .	31
4.3.5	ESAPI . . . . .	31



<b>5 Experimentos</b>	<b>32</b>
5.1 Modelos . . . . .	32
5.2 Auto Planning . . . . .	32
<b>6 Resultados</b>	<b>34</b>
6.1 Modelos de Aprendizado de Máquina . . . . .	34
6.2 Auto-Planning . . . . .	38
<b>7 Conclusão e Trabalhos Futuros</b>	<b>43</b>
7.1 Conclusão . . . . .	43
7.2 Trabalhos futuros . . . . .	45
<b>Appendices</b>	<b>47</b>
.1 Análise Descritiva e Concordância de pontos do DVH . . . . .	48
<b>Referências</b>	<b>60</b>

# Lista de Figuras

2.1	Fluxo de um paciente na Radioterapia. . . . .	5
2.2	Acelerador linear (esquerda) e detalhe da abertura da saída de radiação (direita). . . . .	6
2.3	Representação esquemática da função custo em função dos parâmetros de entrada. . . . .	8
2.4	Representação dos contornos do volume alvo (em azul), reto (em marrom) e bexiga (em amarelo), e dos arcos utilizados utilizando a técnica de VMAT. . . . .	9
2.5	Tabela com os objetivos do planejamento de próstata hipofracionada. . . . .	10
2.6	Exemplo de um DVH utilizado para avaliar uma métrica do planejamento de próstata. A linha amarela representa o DVH da bexiga, a linha marrom o DVH do reto, e a linha vermelha representa o DVH do volume alvo. . . . .	11
2.7	A saída de uma divisão binária recursiva em duas dimensões. . . . .	13
2.8	SVR utilizando um kernel polinomial de segundo grau. . . . .	13
2.9	Representação esquemática da Análise de Componentes Principais. . . . .	14
3.1	Variáveis que se relacionam com as componentes principais de Reto e Bexiga [1]. . . . .	19
3.2	Resumo das Referências Bibliográficas. . . . .	20
3.3	Resumo das Referências Bibliográficas. Continuação. . . . .	21
3.4	Resumo das Referências Bibliográficas. Continuação. . . . .	22
4.1	Tomografia computadorizada co-registrada com os contornos de OAR e a distribuição de dose. No canto direito superior os respectivos DVHs. . . . .	24
4.2	Interface gráfica do programa KBP.dll. . . . .	27
4.3	Fluxo Esquemático dos passos que ocorrem uma vez que o usuário inicia o processo de auto-planejamento. . . . .	28
4.4	Ações realizadas pelo programa KBP.dll: criação dos volumes auxiliares, composição dos campos de tratamento e otimização. . . . .	28
4.5	Pontos de otimização utilizados em uma otimização para o OAR reto. . . . .	29

4.6	Representação esquemática do ambiente VARIAN/Eclipse e das interfaces de acesso ao banco de dados através do ESAPI . . . . .	31
6.1	Resumo da análise dos modelos de predição para D50 e V_10-50. . . . .	35
6.2	Resumo da análise dos modelos de predição para dos DVHs. . . . .	36
6.3	Resultados da análise de <i>cross-validation</i> para a predição dos DVHs com <i>Random Forest</i> . . . . .	36
6.4	Áreas dos DVHs predito vs real para o modelo RF com <i>grid_search</i> . . . . .	37
6.5	Áreas dos DVHs predito vs real para o modelo RF com <i>grid_search</i> . . . . .	37
6.6	Sobreposição dos DVHs dos cinco Planejamentos para um dos Casos. . . . .	38
6.7	Razão das áreas dos DVHs de reto dos planejamentos dos especialistas (numerador) em relação ao auto-planejamento (denominador). . . . .	39
6.8	Razão das áreas dos DVHs de bexiga dos planejamentos dos especialistas (numerador) em relação ao auto-planejamento (denominador). . . . .	40
6.9	Porcentagem das vezes em que um ponto do DVH de reto do auto-planejamento é superior em 5 % ou mais do mesmo ponto no DVH elaborado pelo especialista. . . . .	40
6.10	Porcentagem das vezes em que um ponto do DVH de bexiga do auto-planejamento é superior em 5 % ou mais do mesmo ponto no DVH elaborado pelo especialista. . . . .	41
6.11	Média da diferença percentual entre as áreas dos histogramas dos especialistas em comparação com as áreas dos histogramas gerados pelo autoplanejamento. Valores negativos indicam área do DVH do autoplanejamento menor que a área do DVH do especialista. . . . .	41
6.12	Variabilidade entre as áreas dos DVHs dos especialistas para os volumes de reto e bexiga, normalizado pelas áreas dos DVHs dos planejamentos gerados pelo físico I. . . . .	42
7.1	Distribuição de dose gerada pela solução de auto-planejamento. Em destaque o formato côncavo das isodoses para proteger o OAR reto. . . . .	44
2	Análise de Correlação (acima) e Estatística Descritiva para a Predição de D50 e V_10-50 do reto. . . . .	48
3	Análise de Correlação (acima) e Estatística Descritiva para a Predição de D50 e V_10-50 da bexiga. . . . .	49
4	Análise de Correlação (acima) e Estatística Descritiva para a Predição de D50 e V_10-50 do reto_opt. . . . .	50
5	Análise de Correlação (acima) e Estatística Descritiva para a Predição de D50 e V_10-50 da bexiga_opt. . . . .	51

6	Análise de Correlação (acima) e Estatística Descritiva para a Predição do DVH do reto. . . . .	52
7	Análise de Correlação (acima) e Estatística Descritiva para a Predição do DVH da bexiga. . . . .	53
8	Análise de Correlação (acima) e Estatística Descritiva para a predição do DVH do reto_opt. . . . .	54
9	Análise de Correlação (acima) e Estatística Descritiva para a Predição do DVH da bexiga_opt. . . . .	55
10	Concordância entre os pontos de DVH de reto predito vs observado. . . . .	56
11	Concordância entre os pontos de DVH de bexiga predito vs observado. . . . .	57
12	Concordância entre os pontos de DVH de reto_opt predito vs observado. . . . .	58
13	Concordância entre os pontos de DVH de bexiga_opt predito vs observado. . . . .	59

# Capítulo 1

## Introdução e Objetivos

A incorporação de sistemas computadorizados e o uso de imagens volumétricas na especialidade de radioterapia permitiu uma sensível evolução dos tratamentos do ponto de vista do ganho terapêutico, em que se tem o objetivo de maximizar a dose de radiação entregue ao tumor enquanto a dose entregue aos tecidos sadios é minimizada. Tal evolução permitiu ir dos planejamentos baseados em imagens bidimensionais ao uso de imagens tridimensionais; das técnicas de tratamento de campo uniforme (3DCRT - *3D Conformal Radiotherapy*) às técnicas com campo de intensidade modulada (IMRT - *Intensity Modulated Radiation Therapy*), evoluindo então para a radioterapia entregue de maneira mais eficiente: o arco volumétrico (VMAT - *Volumetric Modulated Arc Therapy*). Em contrapartida, a quantidade de etapas e a complexidade dos processos aumentou na mesma proporção, de modo que se tornou um desafio garantir a qualidade, a segurança e a homogeneidade do fluxo de um paciente na radioterapia.

A etapa de planejamento do tratamento é um destes processos, em que a experiência e o conhecimento do planejador (físico ou dosimetrista) afetam diretamente a qualidade do plano. Desta forma, um planejamento não-ótimo pode comprometer o ganho terapêutico proporcionado por técnicas complexas de tratamento, como IMRT e VMAT. Também, esta etapa é geralmente das que se leva mais tempo dentro do fluxo de radioterapia e varia de acordo com a experiência do planejador. Isso afeta a capacidade do número de casos planejados em um departamento de radioterapia, assim como o intervalo de tempo para o paciente iniciar o tratamento.

A automatização desta etapa permite melhorar a consistência e homogeneidade da qualidade dos planejamentos, ao mesmo tempo em que aumenta a capacidade do número de planejamento de novos casos para radioterapia sem a necessidade de mais profissionais para realizar esta tarefa. Isto é particularmente importante em um país de economia emergente como o Brasil, em que há em algumas regiões dificuldade ou demora para que o paciente oncológico tenha acesso a radioterapia. Ainda, como mostrou o relatório

RT2030 da Sociedade Brasileira de Radioterapia <sup>1</sup>, ao fazer a projeção para o ano de 2030 do número de novos casos de pacientes que irão precisar de radioterapia ao longo do curso de seu tratamento, haverá dentro de uma década falta de profissionais treinados para atender tal demanda. O relatório também demonstra a necessidade e os desafios de tornar a radioterapia economicamente mais viável, dados os altos custos para colocar em operação uma nova instalação e para mantê-la em funcionamento.

Uma abordagem para a automação do planejamento é o uso do *Knowledge-Based Planning* (KBP), a qual utiliza uma biblioteca de planejamentos prévios de pacientes para criar modelos que irão prever a distribuição de dose ou histogramas de DVH (*Dose-Volume Histogram*) de órgãos de risco (OAR, *Organs at Risk*) em novos pacientes, baseados, por exemplo, em métricas geométricas e volumétricas, como a intersecção entre o OAR e o volume alvo ou as distâncias entre voxels de órgãos sadios e de volumes alvo, dentre outros. Ser capaz de prever a dose que os OAR irão receber, dada uma configuração geométrica de contornos e dose prescrita, pode automatizar ou semi-automatizar a etapa de planejamento. Isto aumenta a capacidade de planejamento de novos casos em um departamento de radioterapia e melhora a consistência dos mesmos, pois diminui a intravariabilidade da qualidade de planejamentos dentro do departamento, possibilitando manter um padrão institucional.

Em paralelo, o Instituto Nacional do Cancer (INCA) <sup>2</sup> estima de 65 a 70 mil novos casos de câncer de próstata no ano, sendo este o tipo de câncer mais comum para o sexo masculino e representando 29% dos diagnósticos da doença no Brasil, se forem excluídos os tumores de pele não-melanoma. A radioterapia é uma opção de tratamento para estes pacientes e vem, nos últimos anos, buscando esquemas mais curtos de tratamento, ou seja, com menor número de frações (hipofracionados) [2]. Isto também é relevante para países de economia emergente como o Brasil, pois ao mesmo tempo em que ajuda a aliviar a fila de pacientes no sistema público de saúde, torna mais eficiente o tratamento também nas clínicas privadas.

Neste trabalho, foi desenvolvido um *plug-in* que utiliza *Knowledge-Based Planning* em pacientes com tumores de próstata localizados tratados com esquemas hipofracionados para padronizar, automatizar processos e aumentar a produtividade e a eficiência do processo do planejamento.

---

<sup>1</sup><https://sbradioterapia.com.br>

<sup>2</sup><https://www.inca.gov.br>

## 1.1 Objetivos

O objetivo deste trabalho é o de criar uma metodologia baseado no padrão institucional que, dado um conjunto de imagens de tomografia computadorizada, um conjunto de estruturas com os alvos e os OAR definidos para o tratamento de próstata em regime hipofracionado, automatize ou semiautomatize a etapa do planejamento com o uso de KBP.

### 1.1.1 Objetivos Específicos

Uma vez que se busca a automatização ou semiautomatização dos planejamentos, os objetivos específicos refletem o aumento de produtividade e padronização que a ferramenta possibilita. São eles:

- Demonstração de que a qualidade do autoplanejamento gerado não é inferior ao planejamento do especialista.
- Eliminação ou diminuição do tempo de interação entre o planejador e o sistema de planejamento para realizar o planejamento;
- Aumento da capacidade do número de planejamentos de um departamento de radioterapia, sem que haja a necessidade de novos investimentos, facilitando acesso de pacientes oncológicos à radioterapia;
- Através da criação de um modelo baseado no padrão institucional, aumento da consistência dos planejamentos entre planejadores.

Os dados utilizados neste trabalho são reais e foram obtidos no banco de dados da radioterapia do Hospital Sírio-Libanês.

## 1.2 Capítulos subsequentes

Este trabalho está organizado da seguinte forma: No Capítulo 1 é contextualizado o problema abordado neste trabalho e a solução proposta, nos tópicos de Introdução e Objetivos. No Capítulo 2 são tratados os principais conceitos teóricos necessários para o entendimento do trabalho. No Capítulo 3 é apresentada a revisão da literatura, em que o objetivo é descrever trabalhos que têm utilizado o conceito de *knowledge-based planning* e os algoritmos que têm sido utilizados. No Capítulo 4 é descrita a metodologia que será utilizada para o desenvolvimento do trabalho. No Capítulo 5 são descritos os experimentos que foram realizados. No Capítulo 6 são descritos os resultados encontrados; por fim, no Capítulo 7 são apresentados a conclusão e os trabalhos futuros. Por fim, tem o apêndice, composto por figuras que ilustram a análise descritiva dos dados e parte dos resultados.

# Capítulo 2

## Fundamentação Teórica

Neste capítulo são descritas as técnicas de planejamento e tratamento de Radioterapia utilizados neste estudo, o conceito de *Knowledge-Based Planning* e suas abordagens, e os algoritmos de aprendizado de máquina utilizados nos dados deste estudo.

### 2.1 Radioterapia

Radioterapia consiste no uso de radiação ionizante no tratamento do paciente oncológico. Nos últimos anos, houve aumento significativo na complexidade dos planejamentos e tratamentos com a introdução de novas técnicas, como, por exemplo, a radioterapia de intensidade modulada (*IMRT – Intensity Modulated Radiation Therapy*) e radioterapia de arco modulada (*VMAT – Volumetric Modulated Radiation Therapy*), de maneira que o desafio de se assegurar a qualidade e segurança dos processos e do fluxo de informações aumentou na mesma proporção. As etapas principais de radioterapia, vide Figura 2.1, são as seguintes:

- Consulta Inicial: consiste na avaliação do médico radioterapeuta se haverá ou não indicação de radioterapia para o paciente diagnosticado.

- Simulação: havendo indicação de tratamento, o paciente realiza a simulação, que consiste: (i): na definição do posicionamento que o paciente será tratado e confecção de acessórios de imobilização, a depender do sítio a ser tratado; (ii) aquisição de uma tomografia volumétrica da região a ser tratada. Este conjunto de imagens é utilizado em seguida para contorno e planejamento.

- Contorno: nesta etapa, o médico radioterapeuta, no *software* de planejamento, utilizando a imagem de tomografia adquirida em conjunto com modalidades complementares, PET ou Ressonância Magnética, por exemplo, faz o contorno volumétrico dos alvos a serem tratados e dos órgãos sadios adjacentes ao tumor.



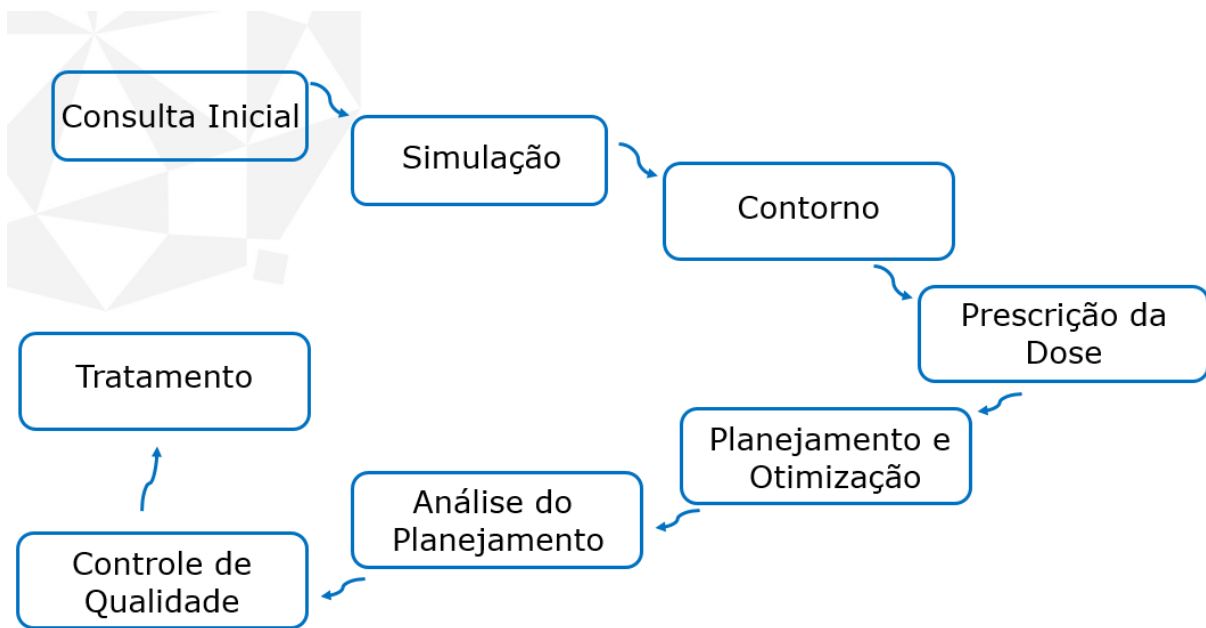


Figura 2.1: Fluxo de um paciente na Radioterapia.

- Prescrição da Dose de Radiação: finalizados os contornos, o médico radioterapeuta faz a prescrição da dose de radiação a ser recebida pelos tumores em tratamento.

- Planejamento e Otimização: nesta etapa, o planejador (físico ou dosimetrista) define como a dose prescrita será entregue aos tumores, preservando, ao mesmo tempo, órgãos adjacentes ao tumor. Aqui são definidos quantidade, formato, energia e angulação dos campos de tratamento. Também, é definida a técnica de tratamento que será utilizada, que pode variar muito em complexidade e acrescentar outros passos ao processo.

- Análise do planejamento: nesta etapa, o médico e o físico, em conjunto, analisam a adequação do planejamento realizado, em termos de dose nos tumores, e se os limites de dose nos órgãos, preconizados por protocolos internacionais, estão sendo respeitados.

- Controle de Qualidade: nesta etapa o físico faz medidas e avaliações para garantir que o acelerador linear (equipamento que realiza a entrega do tratamento) é capaz de entregar o plano aprovado com acurácia.

- Tratamento: então, com o paciente posicionado da mesma forma em que foi realizada a simulação, o tratamento baseado no planejamento realizado é executado pelo acelerador linear. Nesta etapa, imagens são realizadas pré-tratamento para correção do posicionamento e inúmeros parâmetros devem ser checados, como: ângulos da mesa de tratamento e do gantry (corpo do aparelho), parâmetros de abertura do campos e do feixe, etc. A Figura 2.2 ilustra à esquerda a foto de um acelerador linear, máquina utilizada para entregar a dose conforme o planejamento realizado. Durante a sessão o paciente fica deitado na mesa e o *gantry* é capaz de rotacionar 360 graus ao seu redor. À direita, deta-

lhe da abertura por onde sai o campo de radiação. Através de um dispositivo acoplado, é possível conformar o feixe de radiação em múltiplos formatos, o qual pode mudar à medida em que o gantry rotaciona.



Figura 2.2: Acelerador linear (esquerda) e detalhe da abertura da saída de radiação (direita).

## 2.2 VMAT e Otimização

A técnica de VMAT consiste na utilização de arcos de intensidade modulada na irradiação do volume alvo. Tal modulação do feixe - através da variação a cada grau da abertura do campo, da variação de velocidade do gantry e da taxa de dose - permite esculpir a isodose de prescrição no formato do tumor e criar distribuições de dose côncavas, minimizando a dose entregue em órgãos adjacentes saudáveis. Esta técnica tem demonstrado resultados de distribuição de dose altamente conformada, planejamentos com boa qualidade e tempos curtos de tratamento [3]. Este padrão de modulação é formatado durante o processo de otimização, que se dá após a definição da configuração dos arcos que serão utilizados.

O processo de otimização é interativo entre o *software* de planejamento e o planejador, e se inicia com o usuário definindo os objetivos de otimização com o uso dos histogramas dose-volume (DVH), o qual relaciona a distribuição de dose em um volume contornado. Os DVHs são utilizados para a interação com o planejador, pois eles se relacionam com as aberturas do feixe em cada incidência do feixe de radiação da seguinte maneira: as aberturas do campo de tratamento em cada ângulo de incidência contido

no arco definido, levam a uma certa distribuição de dose no paciente e nos contornos definidos, a qual é representada pelos respectivos DVHs. Tendo-se os objetivos iniciais de otimização definidos para os alvos e para os órgãos a serem protegidos (OAR, do inglês *Organs At Risk*) e a abertura inicial do campo em cada angulação do arco de tratamento, a qual é representada pelos DVHs dos volumes definidos, se inicia um processo iterativo, que busca minimizar a função custo definida por:

$$F_n = \left( \frac{1}{N} \right) \sum_{i=1}^N W(\vec{r}) (D_0(\vec{r}) - D_n \vec{r})^2 \quad (2.1),$$

em que  $F_n$  é o custo da  $n$ -ésima iteração,

$D_0(\vec{r})$  é a dose desejada em um certo volume de uma estrutura,

$D_n(\vec{r})$  é a dose calculada em um certo volume de uma estrutura,

$W(\vec{r})$  é o peso dado a um determinado objetivo de otimização,

$N$  é o número de objetivos de otimização[4].

Em outras palavras, busca-se uma distribuição de dose no paciente que se relacione com os objetivos pedidos pelo planejador. Para se alcançar isso, durante o processo de otimização ocorrem mudanças, para cada ângulo de incidência do feixe de radiação, na abertura do campo, na taxas de dose e na velocidade de rotação do *gantry*, em tentativas de minimizar a função custo por parte do sistema de otimização, até que haja convergência da função custo. Ao mesmo tempo, o planejador parametriza os objetivos da otimização para alcançar uma solução ótima, em que, dadas as situações de contorno, a dose prescrita pelo médico é dada ao volume alvo enquanto a mínima dose é entregue aos volumes adjacentes sadios.

É importante notar que existem limites de dose preconizados para os órgãos sadios - os quais podemos definir como objetivos do plano-, porém, o propósito do processo de otimização é encontrar não apenas uma solução que satisfaça os objetivos do plano, mas, sim, encontrar a solução ótima, o que pode ser entendido como um mínimo global da função custo. Na Figura 2.3 tem-se a representação em que a função custo convergiu para um mínimo local (ponto à esquerda), quando idealmente ela deveria convergir para um mínimo global (ponto à direita). Por este motivo e pelo fato da otimização ser um processo de interação entre o planejador e o sistema de planejamento, a qualidade do plano final pode variar muito entre planejadores para um mesmo paciente e o tempo para realizar esta tarefa em geral é a que mais consome tempo dos planejadores.

## 2.3 Radioterapia Hipofracionada de Próstata

Especificamente para a Radioterapia Hipofracionada de Próstata e no contexto em que os tratamentos deste estudo se deram, as etapas da radioterapia tiveram as seguintes configurações:

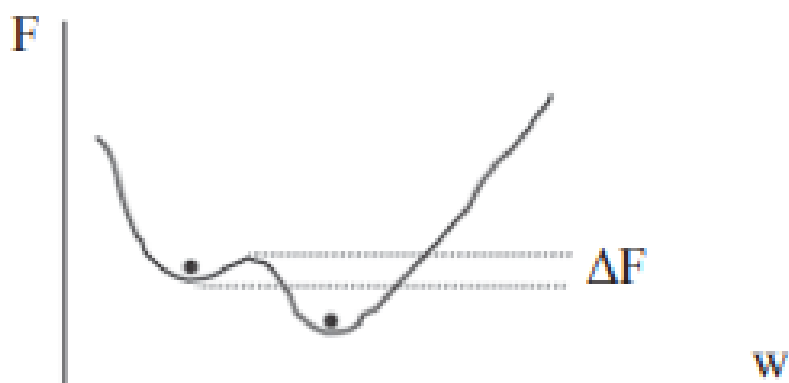


Figura 2.3: Representação esquemática da função custo em função dos parâmetros de entrada.

- Simulação. Posicionamento: Decúbito Dorsal, Cabeça para o *gantry*, Apoio de cabeça, indexador de pés e joelho (pró-step), mãos segurando argola. Imagens de CT: adquiridos em CT Simulador SOMATOM AS DEFINITION SIEMENS 20 canais, espessura do corte de 2.0 mm, da vértebra L2 até o terço posterior do fêmur.

- Contorno: o alvo definido consiste no contorno da próstata, adicionado de uma margem isotrópica de 7.0 mm. Os OAR definidos foram bexiga, reto, femur esquerdo, femur direito, bulbo peniano, genital, alças intestinais. Neste trabalho os OAR de interesse para geração de DVH predito foram o reto e a bexiga(Figura 2.4).

- Prescrição da Dose de Radiação: A dose prescrita foi de 20 frações de 3 Gy no volume alvo, total de 60 Gy.

- Planejamento e Otimização: Foi utilizada a técnica de VMAT, sendo que se utilizou 2 arcos completos, colimadores de 80 e 100 graus, feixe de fótons de 6 MV com taxa de dose de 600 UM/minuto, acelerador linear Varian True Beam STx 2.5, MLC HD120. Os OAR que possuíam intersecção com o volume alvo foram divididos em OAR\_opt(correspondente a fração do OAR sem intersecção, com margem de 2 mm) e OAR\_ovlp (correspondente a fração do OAR com intersecção). Assim, o volume de reto foi dividido em reto\_opt e reto\_ovlp; e o volume de bexiga foi dividido em bexiga\_opt e bexiga\_ovlp.

Os objetivos do planejamento hipofracionado de próstata hipofracionada, de acordo com o protocolo do Hospital Sírio-Libanês, são listados na Figura 2.5. O volume PTV (*Planning Target Volume*) se refere ao volume alvo, em que PTV\_eval é o volume alvo sem intersecção com o reto e PTV\_ovlp se trata do volume de intersecção entre o volume alvo e reto. No caso do objetivo  $D_{99\%} \geq 57$  Gy, significa que se espera que no mínimo 99% do volume PTV\_eval receba uma dose de radiação igual ou maior que 57 Gy. Analogamente,

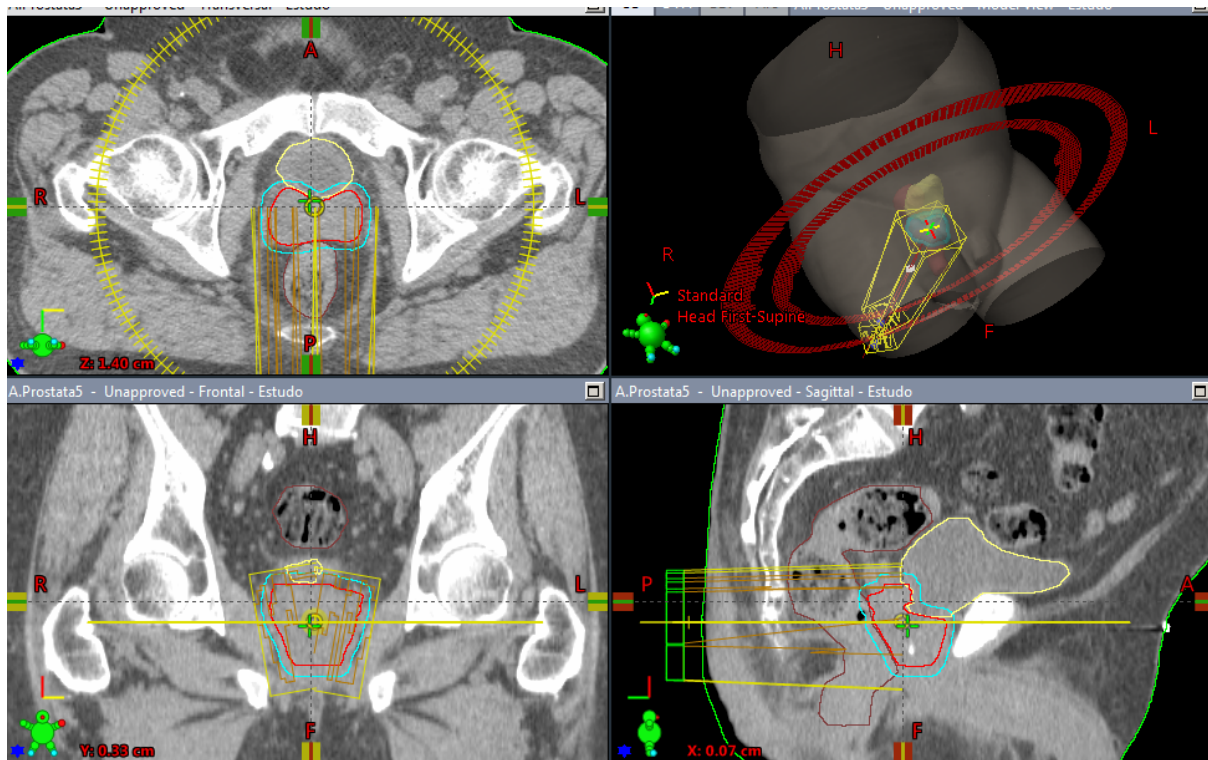


Figura 2.4: Representação dos contornos do volume alvo (em azul), reto (em marrom) e bexiga (em amarelo), e dos arcos utilizados utilizando a técnica de VMAT.

o objetivo relacionado ao reto  $V_{58\text{Gy}} \leq 5\%$ , significa que o volume do reto que recebe 58 Gy deve ser menor que 5%. Por fim, no objetivo para o intestino delgado  $D_{\text{máx}} < 50$  Gy, significa que a dose máxima no intestino deslgado deve ser menor que 50 Gy.

## 2.4 Knowledge Based Planning

Nos últimos anos, *Knowledge-Based Treatment Planning* (KBP) tem ganhado crescente interesse das comunidades acadêmicas e clínicas [5]. KBP explora um conjunto de dados de tratamentos realizados para prever métricas dosimétricas em novos planos de tratamento. Estas métricas previstas podem ser utilizadas como referências para a criação de novos planos, como uma ferramenta de controle de qualidade para planejamentos, ou como parte de um sistema de planejamento semi ou totalmente automatizado. Muitos trabalhos na literatura sobre KBP têm utilizado algoritmos de aprendizado de máquina em combinação com métricas geométricas ou parâmetros do plano.

Estas métricas dosimétricas são extraídas do histograma dose-volume, também chamado de DVH (*Dose-Volume Histogram*), que consiste em um histograma cumulativo que relaciona a dose de radiação entregue a um volume definido (alvo ou de tecido sadio).

<b>PTV_eval</b>	$D_{99\%} \geq 57\text{Gy}$
<b>PTV_ovl</b>	$D_{99\%} \geq 54\text{Gy}$
	$D_{\text{máx}} < 63\text{Gy}$
<b>Bexiga</b>	$V_{58\text{Gy}} < 5\%$
	$V_{47\text{Gy}} < 25\%$
	$V_{40\text{Gy}} < 50\%$
<b>Reto</b>	$V_{58\text{Gy}} < 3\%$
	$V_{55\text{Gy}} < 15\%$
	$V_{48\text{Gy}} < 50\%$
<b>Fêmur</b>	$V_{44\text{Gy}} < 1\%$
	$V_{40\text{Gy}} < 30\%$
<b>Delgado</b>	$D_{\text{máx}} < 50\text{Gy}$

Figura 2.5: Tabela com os objetivos do planejamento de próstata hipofracionada.

Estes histogramas são comumente utilizados como uma ferramenta na fase de avaliação do planejamento, pois condensa informações da distribuição tridimensional de dose em um gráfico bidimensional. Já a distribuição tridimensional de dose é gerada no sistema de planejamento com o uso de um volume de uma tomografia computadorizada. Na Figura 2.6 tem-se um histograma dose-volume para um planejamento de próstata, em que no eixo horizontal se representa a dose em cGy e no eixo vertical se representa o volume em porcentagem. A linha amarela representa o histograma de dose-volume para a bexiga, a linha marrom para o reto e a linha vermelha para o volume definido como alvo. Em destaque está a métrica de análise de dose entregue ao alvo D95, isto é, a dose mínima entregue em 95 por cento do volume.

A quantidade e a qualidade dos planejamentos utilizados para treinar os modelos podem afetar diretamente a acurácia das métricas preditas. De modo geral, um maior número de planejamentos aumenta as variações nas geometrias dos OAR, o que permite ao modelo explorar estas informações e realizar predições mais acuradas [6].

Os métodos de KBP têm sido utilizado principalmente nas seguintes abordagens [7]:

*i)* prever pontos específicos do histograma de dose-volume (DVH) para OAR. Em

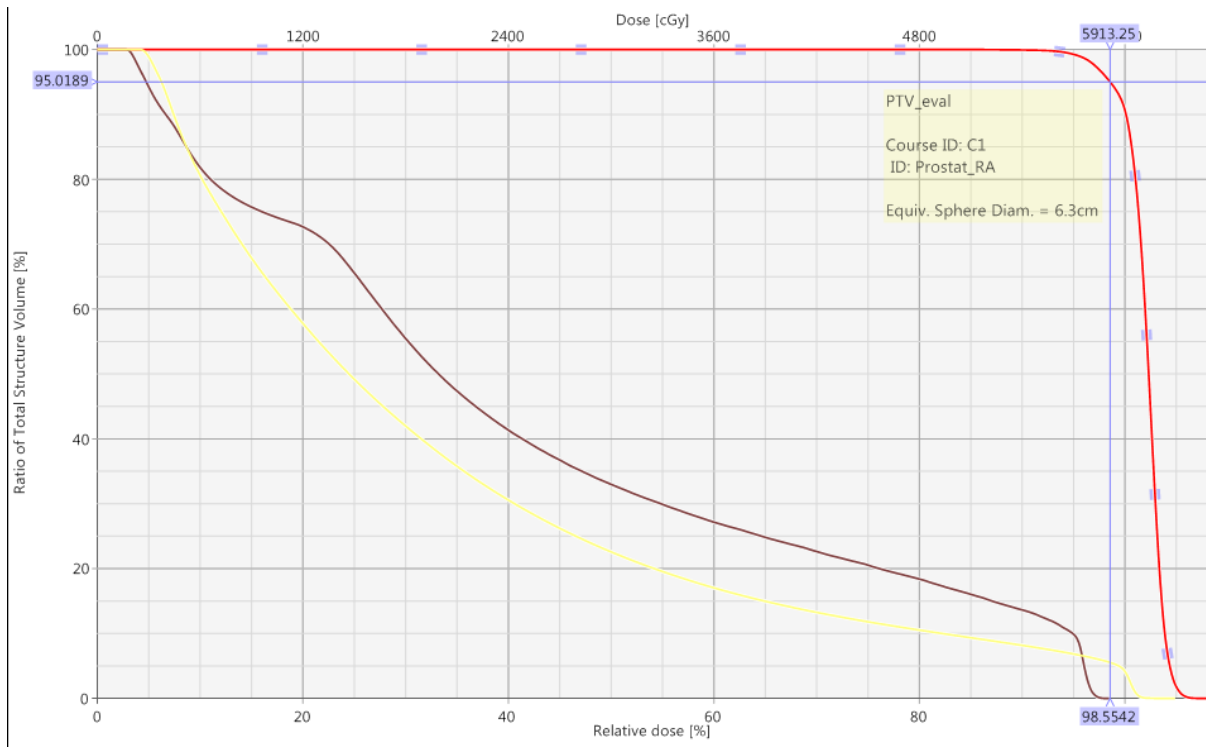


Figura 2.6: Exemplo de um DVH utilizado para avaliar uma métrica do planejamento de próstata. A linha amarela representa o DVH da bexiga, a linha marrom o DVH do reto, e a linha vermelha representa o DVH do volume alvo.

geral utilizado para controle de qualidade dos objetivos dosimétricos mais relevantes do planejamento;

*ii)* prever as curvas de histograma de dose-volume para cada OAR. Estas curvas, podem, então, ser utilizadas como um padrão de referência ou um objetivo final para o planejamento em curso.;

*iii)* encontrar em uma biblioteca de casos tratados o caso mais semelhante com o atual caso em planejamento, com o objetivo de semi automatização do planejamento;

*iv)* prever os pesos da função objetiva no processo de otimização, o que corresponde à importância relativa de cada OAR em relação ao volume alvo.

Neste trabalho optou-se pela abordagem descrita no item (*ii*), devido a praticidade em utilizar este dado previsto de modo direto no processo de otimização dos planejamentos.

Desta forma, soluções de KBP podem ser utilizadas para semi ou automatizar completamente o processo de otimização do planejamento, diminuindo a interação do planejador com o sistema. São também úteis para estabelecer modelos ótimos para a prática clínica, remover a dependência e a variabilidade entre planejadores na qualidade do plano e aumentar a eficiência do processo de planejamento.

## 2.5 Regressão Linear Múltipla

Regressão Linear Múltipla é uma abordagem quando há  $n$  preditores e uma variável resposta unidimensional. Nesta classe de modelos, assume-se que:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + \dots + \beta_n X_{ni} + \epsilon_i$$

em que  $X_j$  representa o valor do  $j$ -ésimo preditor da  $i$ -ésima observação (instância),  $\beta_j$  representa o efeito linear de  $X_{ji}$  em  $Y$ , e  $\epsilon$  é o termo de erro. [8].

## 2.6 *Random Forest*

Árvores de decisão podem ser aplicadas em problemas de regressão e de classificação. A árvore de regressão é formada pela divisão binária recursiva para formar uma árvore grande sobre os dados de treinamento, parando apenas quando cada nó terminal (folha) possuir o número mínimo de observações, como exemplificado na Figura 2.7 [8], em que no processo da divisão recursiva se deu origem a 5 regiões [8]. Ocorre então uma seleção das melhores sub-árvores, através da poda de custo de complexidade em função de um parâmetro  $\alpha$ . Utilizando *cross-validation*, os passos descritos anteriormente são repetidos para os  $(k-1)$  *folds*, em um processo que se varia o valor de  $\alpha$  para minimizar o erro quadrático médio no  $k$ -ésimo *fold*. Com esse valor de  $\alpha$  final se repete o processo de poda para escolher a melhor sub-árvore.

No caso de *Random Forest*, quando as árvores estão sendo criadas, uma amostra aleatória de  $m$  preditores é escolhida candidatos de divisão de todo o conjunto de preditores  $p$ .

## 2.7 *Support Vector Machine - SVM*

O algoritmo de SVM, em tarefas de classificação, busca encontrar um hiperplano que separe as classes existentes no conjunto de dados. Ele é capaz de lidar com classificações lineares e não lineares, ao utilizar *kernels* polinomial ou gaussiano. No caso de regressão por SVM, a ideia é que as instâncias do conjunto de dados fiquem contidas no hiperplano adicionado de uma margem  $\epsilon$ . Da mesma forma que ocorre em classificação, aqui se pode utilizar kernel distintos para trabalhar com regressões não lineares (Figura 2.8) ([9]).

## 2.8 *Grid\_Search*

Cada algoritmo de aprendizado de máquina tem os seus próprios hiperparâmetros, os quais são utilizados no momento de ajuste do modelo. O uso de *grid\_search* é uma abordagem



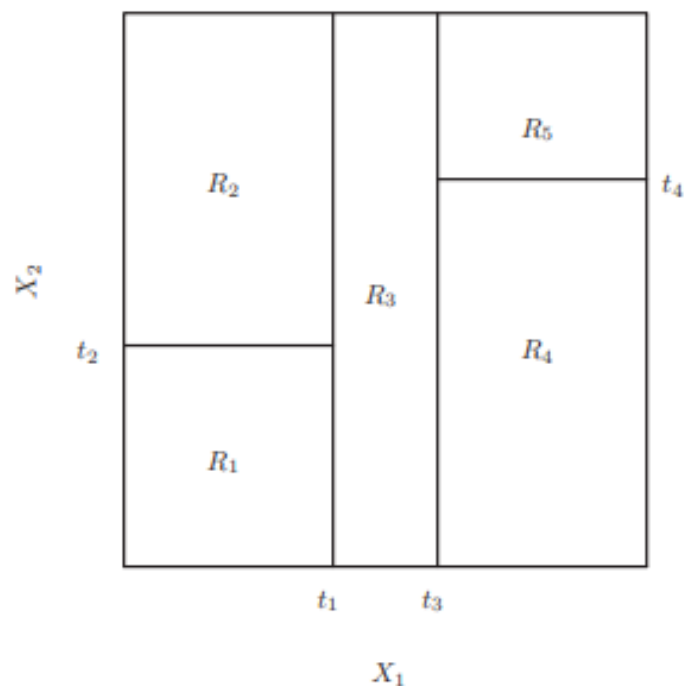


Figura 2.7: A saída de uma divisão binária recursiva em duas dimensões.

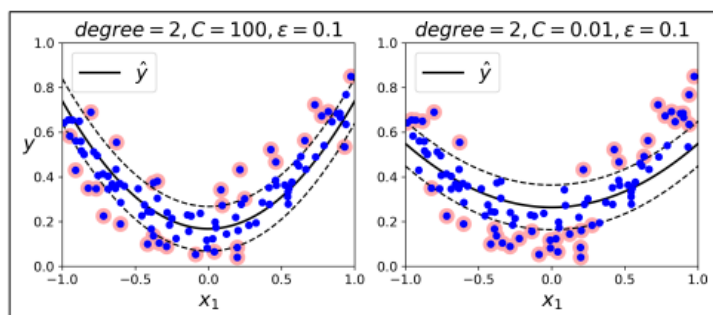


Figura 2.8: SVR utilizando um kernel polinomial de segundo grau..

para se encontrar a melhor combinação dos valores dos hiperparâmetros de um algoritmo dentro de uma faixa de valores especificada, à medida em que ele cria diversos modelos com cada combinação dos valores dos hiperparâmetros [9]. Neste trabalho foi utilizada esta ferramenta com o algoritmo de Regressão por *Random Forest*. Os hiperparâmetros selecionados para o *grid\_search* foram:

- *max\_depth*, a profundidade máxima da árvore.
- *max\_features*, o número máximo de *features* para considerar quando procurando pela melhor divisão.

- *min\_samples\_leaf*, o número mínimo de amostras requerida para estar em um nó de folha.
- *min\_samples\_split*, o número mínimo de amostras requerido para dividir um nó interno.
- *\_estimators*, o número de árvores na floresta.

## 2.9 PCA

*Principal Component Analysis* (PCA) é um conhecido método de redução de dimensionalidade, Figura 2.9 <sup>1</sup>. No caso de duas dimensões, inicialmente, ele identifica o eixo que possui maior variância no conjunto de dados de treinamento. Em seguida, encontra um segundo eixo, ortogonal ao primeiro, o qual tem maior contribuição para a variância remanescente. Para conjuntos de dados de maior dimensão, PCA irá encontrar um terceiro eixo, ortogonal aos dois primeiros, e o quarto em seguida, e assim por diante, a depender do número de dimensões que o conjunto de dados possui [9].

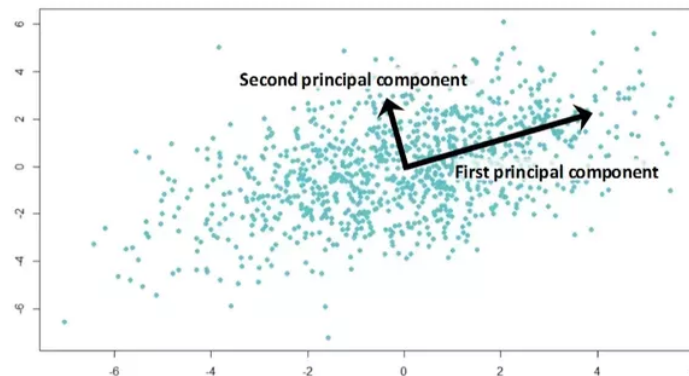


Figura 2.9: Representação esquemática da Análise de Componentes Principais.

---

<sup>1</sup><https://medium.com/@sadatnazrul/the-dos-and-donts-of-principal-component-analysis-7c2e9dc8cc48>

# Capítulo 3

## Revisão da Literatura

Nesta seção são revisados trabalhos que contribuíram para a literatura do *Knowledge-Based Planning*, seu uso e implicações em um departamento de radioterapia.

Chanyavanich et al. [10] mostrou a viabilidade de se usar o conceito de *Knowledge-Based Planning* em planejamentos de próstata. O objetivo foi o de encontrar, dado um caso novo para planejamento, um caso contido em uma biblioteca de 100 planos de próstata que mais se assemelhasse ao planejamento em questão. Uma vez encontrado, o planejamento para o caso novo era iniciado com os mesmos parâmetros de campo e de otimização que o utilizado no caso semelhante encontrado na biblioteca. Esta abordagem permitiu semiautomatizar os planejamentos de próstata, aumentando a eficiência dos processos.

De forma análoga, Schimidt et al. [11] criou um modelo que, para planejamentos de cabeça e pescoço, se encontrava o caso mais similar em uma base de dados de 103 planos de cabeça e pescoço. Em seguida, para se evitar planos sub-ótimos devido às particularidades da geometria encontrada em contornos deste sítio de tratamento, a matriz de dose era deformada de acordo com a sobreposição entre os volumes alvos e os órgãos sadios. Então, os parâmetros do planejamento eram copiados do caso similar encontrado na base de dados e o histograma dose-volume gerado para o novo caso era utilizado para realizar a otimização do planejamento sem intervenção humana.

Campbell e Miften [12] utilizou redes neurais para desenvolver um modelo de predição de distribuições de dose para planejamento em arco (VMAT) de radiocirurgia extracraniana de pâncreas. Utilizando um banco de dados de 43 pacientes, divididos em grupos para treinamento e validação, foram criados modelos para realizar a predição. Os modelos tiveram melhores resultados quando tratados separadamente de acordo com o médico do paciente, pois os padrões de definição de contornos e objetivos do planejamento eram distintos.

Ziemer et al. [13] propôs um método para automatizar o planejamento de radiocirurgia intracraniana de múltiplas metástases. Utilizando um modelo previamente desenvolvido que previa a distribuição de dose no planejamento de radiocirurgia para uma única lesão, fazia-se a predição de dose para as múltiplas lesões tratando cada uma de maneira independente e, em seguida, combinando as predições individuais de dose em uma única matriz de dose. Finalmente, criava-se um arquivo com parâmetros da isodoses resultantes, o qual alimentava os objetivos da otimização do planejamento.

Babier et al. [5], utilizando *Bagging Query* (BQ) e Análise de Componentes Principais (PCA), criaram modelos para prever histogramas de dose-volume (DVH) em tumores de orofaringe. A partir daí, parâmetros gerados do DVH eram inseridos na otimização do plano, levando à criação de planos de alta qualidade, sem intervenção humana.

Du et al. [14], para casos de radiocirurgia de nevralgia do trigêmeo, teve como objetivo prever a dose máxima no tronco, o volume do tronco que recebeu pelo menos 10 Gy e o volume de cérebro sadio que recebeu pelo menos 12 Gy. Foram utilizados 40 casos clínicos e modelos de regressão multivariada, utilizando como parâmetros de entrada o comprimento do nervo, as largura do nervo axial e sagital, a distância entre o isocentro e superfície do tronco, além do tamanho do cone utilizado para tratamento.

Shiraishi et al. [6] usando a metodologia de KBP, desenvolveu um modelo para prever histogramas dose-volume e métricas de qualidade do plano em casos de radiocirurgia intracraniana, com o objetivo de se identificar planos sub-ótimos.

Boutlier et al. [7] buscou determinar como a acurácia de modelos de *Knowledge-Based Planning* de casos de próstata é afetada pelo tamanho do conjunto de dados disponível. O parâmetro utilizado para comparação dos resultados entre os tamanhos de amostras foi o erro médio. Foi encontrado que: 1) para predição de pontos do histograma dose-volume, 200 amostras foram necessárias para prever os pontos de DVH escolhidos; 2) para predição de histogramas dose-volume, utilizando PCA e Regressão Linear Múltipla, ao menos 75 amostras são necessárias para predição do DVH de bexiga e 25 amostras para predição do DVH de reto; 3) para predição de pesos da função objetivo, 10 amostras foram necessárias para treinar o modelo de Regressão Logística, enquanto 150 amostras foram necessárias para treinar o modelo utilizando *k*-vizinhos mais próximos (KNN).

Younge et al. [15] utilizou planos gerados com a metodologia KBP para radiocirurgia extracraniana de vértebra, segundo o protocolo RTOG (*Radiation Therapy Oncology Group*) 0631. Os planos gerados via KBP foram comparados com os planejamentos originais em termos de conformidade da dose de prescrição no alvo, do índice de gradiente, dose em estruturas críticas, como medula, e dose em tecidos sadios adjacentes. Os planos gerados via KBP mostraram maior consistência entre si e melhora em algumas métricas, em comparação aos planos originais.

Nwankwo et al. [16] propõe validar o uso de um software comercial de KBP em planejamentos de próstata. Pela seleção randômica de 30 pacientes planejados com VMAT por pessoas com experiência, estes casos foram planejados novamente utilizando KBP por pessoas sem experiência em planejamentos. Os resultados mostraram que não existiram diferenças estatisticamente significantes entre as métricas analisadas nos planejamentos resultantes.

Chatterjee et al. [17] analisou o desempenho de um modelo de KBP para casos de sistema nervoso central, em que foram utilizados inicialmente 82 casos de glioblastomas tratados com VMAT ou IMRT. Em um passo intermediário, estes casos foram replanejados manualmente, buscando-se o melhor plano possível para ser utilizado no modelamento. Então, o modelo final foi validado em um conjunto de dados independente de 45 pacientes de glioblastoma, astrocitoma, oligodendroglioma e meningioma. Os planos gerados pelo método KBP resultaram em métricas de avaliação de alvo superiores aos planos feitos manualmente e as métricas de OAR, semelhantes. O tempo de planejamento utilizando KBP diminuiu sensivelmente, de cerca de 4 horas de planejamento manual para 13 minutos utilizando KBP.

Wang et al. [18] avaliou se o desempenho de uma solução comercial de KBP para planejamentos de mama esquerda, em que se tem o coração próximo e frequentemente impõe desafios à etapa de planejamento, pode ser superior ao cenário em que se tem planejadores com diferentes níveis de experiência. O modelo de KBP foi gerado através do uso de 80 planejamentos ótimos feitos previamente. 6 planejadores, sendo 2 de nível iniciante, 2 de nível intermediário, 2 de nível avançado, planejaram 10 outros casos que não foram utilizados para a criação do modelo, sendo que cada planejador gerou para cada caso um plano manual e outro utilizando KBP. Os resultados mostraram que o uso de KBP nos planejamentos aumenta a homogeneidade da qualidade dos planos gerados e melhora as métricas de qualidade para os planejadores menos experientes.

Tol et al. [19] avaliou o desempenho de uma solução comercial de KBP utilizando casos de cabeça e pescoço, utilizando um total de 60 pacientes e dividindo-os em grupos, de acordo com particularidades de planejamento. Resultados comparáveis aos planos originais foram alcançados com o KBP nas situações em que havia um caso parecido nos modelos. No entanto, quando se tratava de um caso que possuía configurações de órgão de risco ou volume alvo diferentes dos casos utilizados para modelamento, o planejamento utilizando o método KBP foi de qualidade abaixo do esperado.

Delaney et al. [20] estudou o desempenho de modelos de KBP utilizando e não utilizando planos sub-ótimos, chamados de *outliers*, para a modelagem, mudando, nesta etapa, a porcentagem de planos sub-ótimos. Utilizando um total de 70 pacientes de cabeça e pescoço, alguns planejamentos foram reotimizados sem tentativa de minimizar a dose

nas glândulas salivares, em que se criaram modelos contendo de 5 a 40 planos sub-ótimos. Foram analisados índices de homogeneidade dos alvos, e as doses entregues na cavidade oral, nas glândulas salivares e em estruturas de deglutição. Os resultados mostraram que a remoção de planos sub-ótimos dos casos utilizados na modelagem tiveram pequeno impacto nos índices de conformidade e na diminuição de dose entregue aos órgãos sadios nos planejamentos gerados pela metodologia KBP.

Amit et al. [21] teve como objetivo prever as incidências ótimas dos feixes de radiação para tratamentos convencionais de tórax, evitando a dependência da experiência do planejador para obter bons resultados de planejamento. Utilizando um total de 149 casos tratados de tórax, *random forest* e parâmetros anatômicos (quadrante do tórax em que se encontra o alvo, distância do PTV à superfície do paciente, projeção do overlap entre o alvo e o tecido sadio adjacente em função ângulo de incidência, dentre outros), foi criado um ambiente que seleciona as melhores incidências para um dado caso.

Yuan et al. [1] apresentou um método para identificação das principais *features* geométricas e volumétricas que afetam a variação de dose entre pacientes de um mesmo sítio anatômico (neste trabalho foram analisados pacientes de próstata e cabeça e pescoço), e, conseqüentemente, as variações de DVHs. Este método consistia em utilizar PCA e, em seguida, uma regressão *step-wise*. Foram analisadas as seguintes *features*: o *Distance to Target Histogram*(DTH), o qual representa um histograma da distância mínima entre cada voxel de um OAR e o volume alvo, o DVH e as relações de volume entre o OAR e o volume alvo. O trabalho encontra que para os volumes reto e bexiga, o DTH é principalmente representado por duas variáveis que se relacionam com as componentes principais, o  $d\_metric$  e o  $V_{d1-d2}$ , em que  $d\_metric$  corresponde à distância em que está uma certa porcentagem (*metric*) do volume do OAR; e o  $V_{d1-d2}$  corresponde ao volume do OAR que está entre as distâncias  $d1$  e  $d2$ , com coeficientes de determinação de 0.88 para a bexiga e de 0.68 para o reto. Analogamente, os DVHs de reto e bexiga são majoritariamente explicados por duas variáveis que se relacionam com as componentes principais:  $D\_metric$  e  $V\_D1D2$ .  $D\_metric$  representa a dose entregue em uma porcentagem do volume do OAR, representada por *metric*; e  $V\_D1D2$  representa a fração do volume do OAR que recebe entre as porcentagens de dose  $D1$  e a dose  $D2$ . Na Figura 3.1 aparecem listados valores encontrados para as principais variáveis de reto e bexiga:

Assim, por exemplo, tem-se que a primeira variável do reto para o DTH é representado por  $d60$ , ou seja, pela distância correspondente a 60% do volume do reto. Enquanto que a primeira variável do seu DVH é  $D50$ , correspondente à dose que recebe 50% do seu volume.

Posteriormente, as *features* identificadas foram utilizadas para recriar, aplicando uma regressão múltipla, os DVHs e se achou boa concordância entre os DVHs preditos e os

		<b>Reto</b>	<b>Bexiga</b>
<b>DTH</b>	d_metric	d60	d60
	V_d1d2	$V_{4.0-0.8}$	$V_{7.2-0.8}$
<b>DVH</b>	D_metric	D50	D50
	V_D1D2	$V_{10-50}$	$V_{10-50}$

Figura 3.1: Variáveis que se relacionam com as componentes principais de Reto e Bexiga [1].

reais.

Da Figura 3.2 à Figura 3.4, tem-se o resumo dos trabalhos citados neste capítulo.

Artigo	Localização do Tumor	Técnica de tratamento	Objetivo do Estudo	Tamanho da Amostra	Algoritmo de KBP	Resultado
10	Próstata	IMRT	Mostrar a viabilidade do uso do KBP	100	Procura por caso similar ("match") na base de dados	Planos gerados pelo método proposto mostraram qualidade semelhante aos planos feitos manualmente
11	Cabeça e Pescoço	IMRT	Semiautomatizar planejamentos de Cabeça e Pescoço	103	Procura por caso similar ("match") na base de dados. Deformação das isodoses na interface AOR-alvo	Planos gerados pelo método proposto mostraram qualidade igual ou superior aos planos feitos manualmente
1	Próstata; Cabeça e Pescoço	IMRT	Quantificar o efeito das características dos alvos e OAR, relacionado às suas relações espaciais, na distribuição de dose em planos de IMRT	64 casos de próstata 82 casos de cabeça e pescoço	PCA, Regressão Múltipla	Se identificou os fatores anatômicos em próstata e Cabeça e Pescoço que contribuem para a distribuição de dose. O modelo foi capaz de prever em boa parte dos casos o DVH resultante.
12	Pâncreas	SBRT	Predizer matriz de dose	43	Redes Neurais	Boa concordância entre a distribuição de dose predita e a calculada: erro médio de dose abaixo de 5 por cento. Piores resultados para doses baixas.
13	Crânio	SRS	Predizer matriz de dose. Em seguida, utilizar estes dados na otimização do plano	39 casos para treinamento da rede neural. 41 casos para determinar parâmetros da função de combinação de dose	Redes Neurais para prever matriz de dose de cada lesão de forma independente. Em seguida, aplicada uma função de combinação de dose	Em revisão cega por especialistas, os planos gerados pelo método KBP foram equivalentes ou superior em 78.1 por cento dos 41 casos do estudo.
5	Orofaringe	VMAT	Predizer DVH. Em seguida, utilizar estes dados na otimização do plano	217	Bagging Query (BQ) e PCA	Na etapa de predição, PCA mostrou melhores resultados que o uso de BQ. Por fim, os planos gerados a partir de PCA foram superiores aos planos originais e aos planos gerados por BQ.

Figura 3.2: Resumo das Referências Bibliográficas.



Artigo	Localização do Tumor	Técnica de tratamento	Objetivo do Estudo	Tamanho da Amostra	Algoritmo de KBP	Resultado
14	Nervo Trigêmio	SRS	Predizer métricas relevantes para a aprovação do planejamento	40	Regressão Multivariada	Foram determinadas correlação entre as variáveis de entrada e as métricas de interesse
6	Crânio	SRS / SRT	Predizer DVH e métricas de qualidade do planejamento a fim de se identificar planos sub ótimos	223	Skew normal distribution probability density function	Melhores resultados foram alcançados quando os modelos foram estratificados por volume do alvo e proximidade de OAR. Planos sub ótimos foram identificados e replanejados.
7	Próstata	VMAT	Determinar como o tamanho da amostra afeta a acurácia do modelo KBP	n = 10, 20, 30, 50, 75, 100, 150 e 200	PCA, regressão linear múltipla, regressão logística, k-vizinhos mais próximos	Em função do tipo de predição (pontos do DVH, DVH, pesos da função objetiva), foi determinado o n de amostras por algoritmo
15	Vértebra	SBRT	Usar a metodologia KBP para produzir planos com de SBRT de Vértebra de alta qualidade	22	PCA, Regressão Múltipla	O uso de KBP para SBRT de vértebra resultou planos que atendia aos requisitos do protocolo RTOG 0631 e mostrou maior consistência de qualidade de planejamento
16	Próstata	VMAT	Validar o uso de um sistema comercial KBP em planejamentos de próstata	30	PCA, Regressão Múltipla	Os resultados mostraram que não existiram diferenças estatisticamente significantes entre as métricas analisadas entre planejamentos resultantes e aqueles realizados manualmente.

Figura 3.3: Resumo das Referências Bibliográficas. Continuação.

Artigo	Localização do Tumor	Técnica de tratamento	Objetivo do Estudo	Tamanho da Amostra	Algoritmo de KBP	Resultado
17	Sistema Nervoso Central	MRT / VMAT	Analisar o desempenho de um sistema de KBP para casos de sistema nervoso central	82 para treinamento e validação, 45 para teste	PCA, Regressão Múltipla	Os planos gerados pelo método KBP resultaram em métricas de avaliação de alvo superiores aos planos feitos manualmente e as métricas de OAR, semelhantes. O tempo gasto pelo planejador diminuiu sensivelmente quando se fez uso de KBP.
18	Mama Esquerda	IMRT	Investigar se o método KBP aumenta a qualidade de planejamentos em relação à planos feitos manualmente por profissionais com diferentes níveis de experiência	80 casos para treinamento do modelo, 10 casos para comparação com o planejamento manual	PCA, Regressão Múltipla	Os resultados mostraram que o uso de KBP melhora a consistência e aumenta a qualidade das métricas dos planos originados de profissionais menos experientes
19	Cabeça e Pescoço	VMAT	Analisar o desempenho de um sistema de KBP para casos de cabeça e pescoço	60	PCA, Regressão Múltipla	Se observou resultados aquém do esperado o caso a ser planejado não possuía um similar nos dados que compunham o modelo
20	Cabeça e Pescoço	VMAT	Avaliar o impacto do uso de planos sub-ótimos na etapa de modelamento	70	PCA, Regressão Múltipla	Os resultados mostraram que a remoção de planos sub-ótimos dos casos utilizados na modelagem tiveram pequeno impacto na qualidade dos planos gerados pelo método KBP
21	Tórax	IMRT	Predizer as incidências ótimas para casos de IMRT de tórax	149	Random Forest	O método se mostrou viável e capaz de produzir planos de alta qualidade, sendo que 93 por cento dos planos foram considerados aptos para tratamentos por 3 diferentes médicos

Figura 3.4: Resumo das Referências Bibliográficas. Continuação.

# Capítulo 4

## Metodologia

Este capítulo descreve os dados disponíveis, a metodologia proposta e as ferramentas que são utilizadas no desenvolvimento do trabalho.

### 4.1 Dados

Inicialmente, foi realizada uma varredura no banco de dados do sistema de planejamento para selecionar os pacientes de câncer de próstata em estagio inicial tratados no esquema hipofracionado. O número de casos selecionados foi de 63, cujos arquivos de cada paciente consistem em:

- uma imagem tridimensional de tomografia computadorizada (CT) em formato DICOM;
- um conjunto de estruturas (*structure set*) no formato DICOM-RT STRUCT contendo os contornos desenhados;
- o planejamento no formato DICOM-RT PLAN;
- e uma matriz de dose no formato DICOM-RT DOSE.

As imagens de tomografia foram adquiridas em um CT Somatom Siemens, matriz de 512 x 512, espessura de corte de 2.00 mm, com cerca de 150 cortes. Em cada volume de tomografia os contornos do alvo e dos órgãos adjacentes, e o planejamento foram realizados no sistema Varian Eclipse versão 13.6/15.6, sendo a otimização realizada com o algoritmo PRO (*Progressive Resolution Optimization*) ou PO (*Photon Optimizer*), e o cálculo de dose tridimensional do volume com o algoritmo AAA (*Anisotropic Analytical Algorithm*) com resolução de 1.0 mm. Os tratamentos foram realizados no acelerador linear Varian True Beam 2.5, MLC HD 120, energia de 6 MV. O volume alvo consistiu

no volume delineado pelo radio-oncologista na imagem tridimensional de tomografia computadorizada adicionado de uma margem isotrópica de 7.0 mm. Na mesma tomografia, foram contornados os volumes sadios adjacentes: reto, bexiga, intestino delgado, femur esquerdo, femur direito e bulbo, de acordo com os respectivos guidelines. A partir dos volumes contornados e da matriz de dose, o sistema de planejamento fornece o histograma dose-volume (DVH), como representado na Figura 4.1.

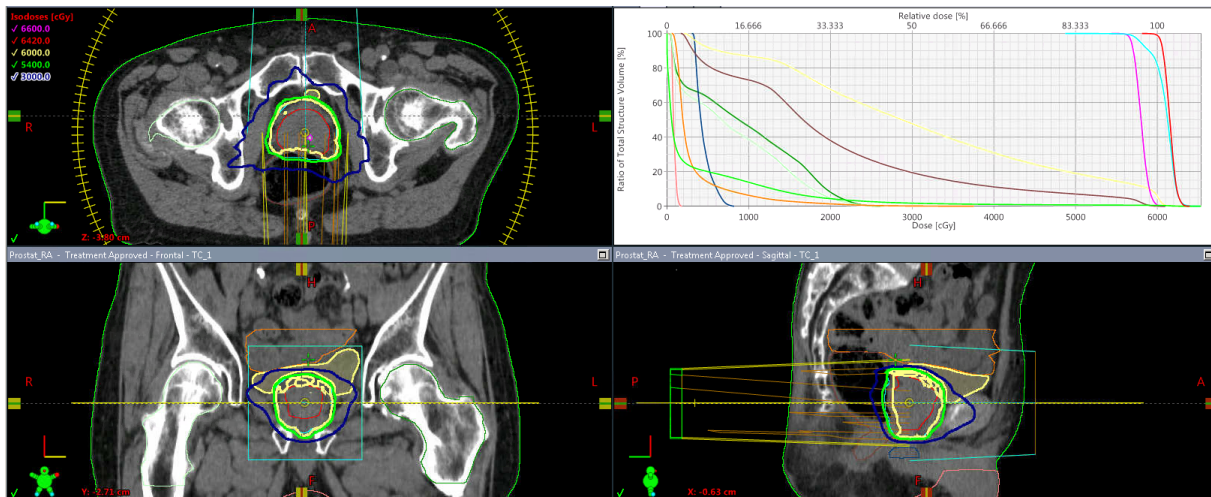


Figura 4.1: Tomografia computadorizada co-registrada com os contornos de OAR e a distribuição de dose. No canto direito superior os respectivos DVHs.

Neste trabalho, as predições de DVH se detiveram aos OAR reto, bexiga, reto\_opt e bexiga\_opt, estes dois últimos sendo a fração desses órgãos sem intersecção com o volume alvo(próstata). Isto porque reto e bexiga são os OAR que apresentam maior desafio no processo de otimização por apresentarem área de intersecção com o volume alvo, exigindo formato côncavo das isodoses. Os volumes reto\_opt e bexiga\_opt são utilizados como volumes auxiliares para potencializar a diminuição de dose na região do OAR que não possui intersecção com o volume alvo, já que nesta região não há objetivos conflitantes para a função custo em termos de entregar dose no volume alvo e proteger o OAR.

## 4.2 Metodologia

Tendo-se, então, a lista dos casos selecionados, foi feito e executado um programa para extrair os dados de interesse dos arquivos de cada paciente. Para cada caso, este programa percorre o volume tridimensional do paciente e extrai os seguintes dados para os OARs de interesse, no caso reto, bexiga, reto\_opt e bexiga\_opt:

1. volume absoluto do alvo (em cc)

2. volume absoluto do OAR (em cc)
3. volume-In do OAR (em %) - fração do volume do OAR contido nos mesmos cortes da tomografia em que está o volume alvo.
4. volume-Off do OAR (em %) - fração do volume do OAR que não está contido contido nos mesmos cortes da tomografia em que está o volume alvo.
5. volume de intersecção entre o alvo e o OAR (em %)
6. volume do OAR sem intersecção com o alvo (em %)
7. DVH do OAR e do alvo, resolução de 1.0%
8. DTH de cada OAR, resolução de 1.0 mm

Estes dados foram salvos em arquivos .csv para posteriormente serem trabalhados a fim de se gerarem os modelos de predição de DVH de OARs. Utilizando o ambiente Jupyter Notebook, os dados foram importados e explorados. Inicialmente, criou-se os modelos para predição das duas variáveis principais dos DVHs (as variáveis que se relacionam com as componentes principais do DVH). Assim, para cada OAR, foram gerados os modelos para predição de  $D\_metric$  e de  $V\_D1D2$ , as variáveis independentes listadas de 1-6, os parâmetros  $d\_metric$  e  $Vd1d2$  (variáveis que se relacionam com as componentes principais do DTH). Neste caso, o conjunto de dados foi formado por 63 instâncias. Testou-se os algoritmos de Regressão Linear Múltipla e Regressão por *Random Forest (Random Forest Regression)*. Devido ao fato do baixo número de amostras, o conjunto de treino e teste ficaram os mesmos.

Para geração do modelo de predição do DVH as variáveis independentes foram as listadas de 1-6, as duas variáveis principais do DVH preditas no passo anterior e as duas variáveis principais do DTH ( $d\_metric$  e  $V\_d1d2$ ). Os parâmetros de  $D\_metric$ ,  $V\_D1D2$ ,  $d\_metric$  e  $V\_d1d2$  foram processados conforme os valores encontrados por [1] para as variáveis principais do DVH e do DTH, respectivamente.

Para este fim o conjunto de dados foi formado por 63 casos, cada caso com 101 instâncias, totalizando de 6363 instâncias. Estes foram divididos aleatoriamente em conjuntos de treino e de teste na proporção de 2/3 e 1/3, respectivamente. Testou-se os algoritmos de Regressão Linear Múltipla, Regressão por Random Forest (*Random Forest Regression*), Regressão por Support Vector (*Support Vector Regression*), utilizando a biblioteca *scikit-learn*. Para o modelo Regressão por *Random Forest*, foi aplicado também a busca pelos parâmetros ótimos da árvore de regressão através do `grid_search`, com os parâmetros definidos por:

```
param_grid = 'bootstrap': [True],
```

```
'max_depth': [10, 20],  
'max_features': [2, 3, 4, 5, 6],  
'min_samples_leaf': [1, 2, 3],  
'min_samples_split': [3, 4, 5, 6, 7, 8, 9, 10],  
'n_estimators': [10, 20, 30]
```

Para análise de *overfitting* para o algoritmo de *Random Forest* foi feita a análise de *cross-validation*

Os modelos foram analisados pelos parâmetros de  $R^2$ , erro médio absoluto e raiz quadrada do erro médio quadrático. Os modelos de predição de DVH foram analisados também com a avaliação das áreas entre o DVH observado e o DVH predito.

Feita a avaliação, se escolheu os modelos que apresentaram melhores resultados de acordo com os parâmetros citados.

Os modelos escolhidos para D\_metric, V\_D1D2 e DVH, para cada OAR, foram salvos em formato .pkl. Criou-se, então, um programa em linguagem Python chamado ApplyModelo.py com a função de aplicar os modelos de aprendizado de máquina nos parâmetros calculados em um paciente e devolver os DVHs preditos das estruturas de OAR de interesse. Para não haver dependências de biblioteca Python nas máquinas em que este programa será executado, este programa foi convertido em um executável, passando a ser chamado de ApplyModel.exe.

### 4.2.1 Automação do Planejamento

A fim de se automatizar a etapa de planejamento, os modelos gerados foram utilizados em conjunto com o programa desenvolvido KBP.dll (Figura 4.2), desenvolvido em C#.

Este programa tem como funções configurar os campos de tratamento, gerar os volumes auxiliares de otimização, calcular e salvar os parâmetros que os modelos de predição irão utilizar, executar o programa ApplyModel.exe e importar o DVH predito para cada OAR como objetivos no processo de otimização (Figura 4.3).

Este último carrega os parâmetros calculados, aplica os modelos de aprendizado de máquina gerados para calcular os DVHs preditos e salva os histogramas em formato .csv. Então, os pontos previstos do DVH são carregados de volta pelo programa KBP.dll e utilizados como objetivos para o algoritmo de otimização. Para o volume reto, foi também utilizado a dose média prevista como parâmetro para a função de otimização. Na Figura 4.4, à esquerda se tem as ações iniciais do programa KBP.dll, em que os arcos de tratamentos são configurados e os volumes auxiliares de otimização gerados. À direita, se tem os objetivos de otimização, gerados a partir da predição de DVHs, utilizados na etapa de otimização.



## Selecione a Técnica:

Próstata

Iniciar

## Selecione o(s) Alvo(s)

 PTVp20x3Gy

6000

## Alvos Selecionados

PTVp20x3Gy 6000

## Selecione o Curso

C1

## Dados do Planejamento

Número de Frações:

20

Acelerador

TRUE-BEAM

Gantry Avoidance

Continuar

Entre com os dados

Figura 4.2: Interface gráfica do programa KBP.dll.

Realizada a otimização, é realizado o cálculo 3D de dose utilizando o algoritmo AAA\_13623 [22].

Como a função custo de otimização lida com objetivos conflitantes (por exemplo, entregar a dose de prescrição no volume tumoral com o mínimo de dose em órgãos que estão próximos ou em intersecção com o volume alvo), é uma prática posicionar os objetivos de otimização abaixo do que se realmente se deseja com o DVH real. Na Figura 4.5 se tem um exemplo de objetivos de otimização para o reto utilizados por um especialista em um caso de próstata. Note que os pontos que representam os objetivos da otimização estão abaixo do DVH calculado.

Por isso, findo o primeiro ciclo de otimização e cálculo, o programa KBP.dll repete a etapa de otimização e aplica a função *offset* descrita no Algoritmo 1. Dado um objetivo

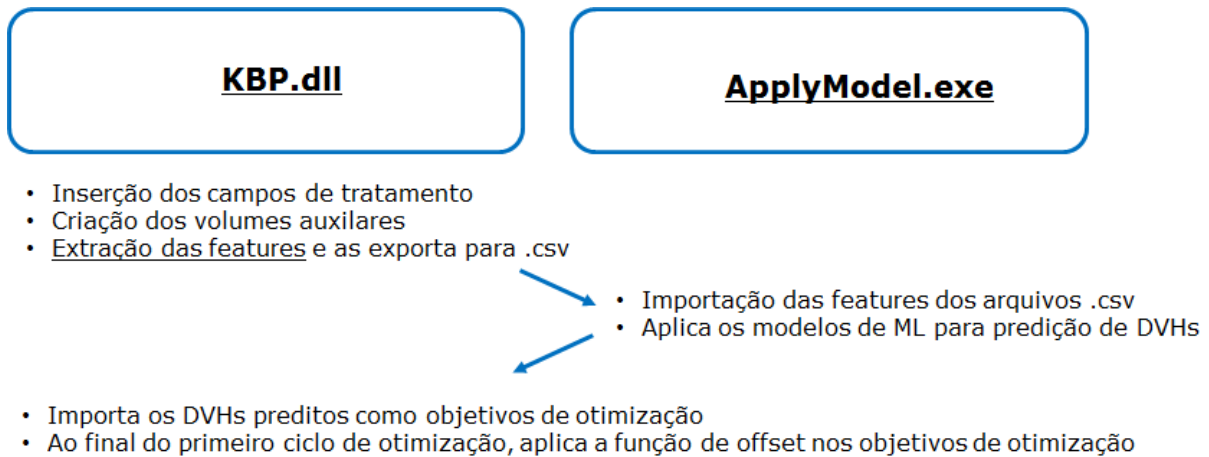


Figura 4.3: Fluxo Esquemático dos passos que ocorrem uma vez que o usuário inicia o processo de auto-planejamento.

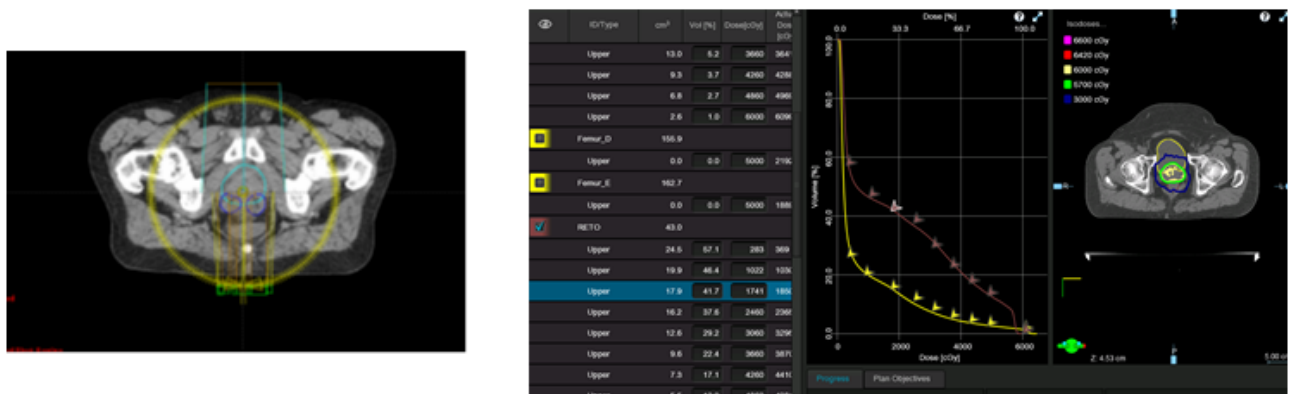


Figura 4.4: Ações realizadas pelo programa KBP.dll: criação dos volumes auxiliares, composição dos campos de tratamento e otimização.

de DVH, ela minimiza o valor do objetivo de acordo com o seu respectivo valor de volume, simulando o offset aplicado pelo especialista, através de um fator de correção (*corrFactor*).

Os objetivos de otimização são então substituídos pelos valores de auxPointVolumeDVH e a otimização, reiniciada. Ao fim desta segunda otimização, o cálculo de dose utilizando o algoritmo AAA\_13623 é realizado novamente. Neste ponto espera-se que o planejamento esteja pronto, em que o próximo passo seria a sua avaliação.

Assim, o usuário interage apenas com a interface mostrada na Figura 4.2 e tem um planejamento de próstata ao final da execução do programa.



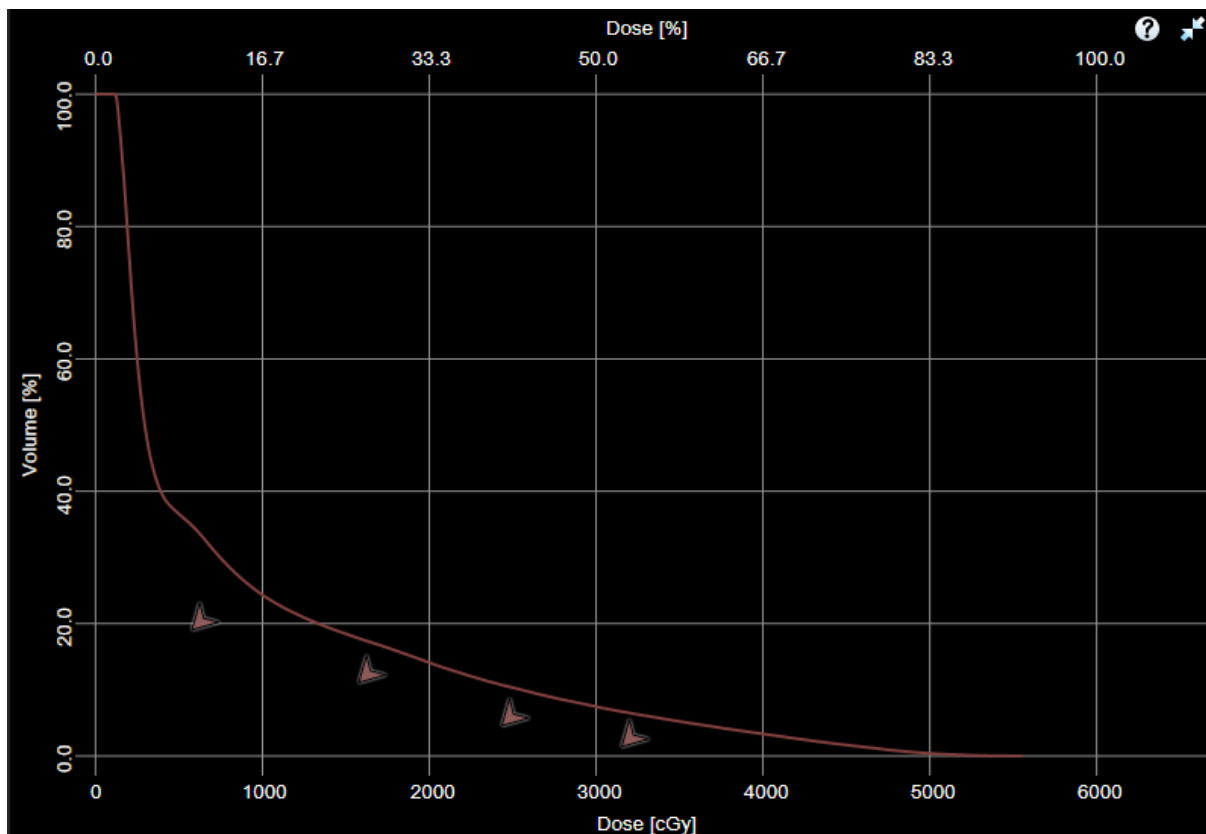


Figura 4.5: Pontos de otimização utilizados em uma otimização para o OAR reto.

## 4.3 Ferramentas

A seguir a descrição das ferramentas mais importantes utilizadas neste trabalho.

### 4.3.1 Formato de arquivo DICOM

DICOM® (*Digital Imaging and Communications in Medicine*)<sup>1</sup> é o padrão internacional para transmitir, armazenar, recuperar, imprimir, processar e exibir informações de imagens médicas. É a base da implementação do PACS (*Picture Archiving and Communication Systems*) em radiologia, uma vez que permite a comunicação de arquivos originados de máquinas de distintos fabricantes.

### 4.3.2 Formato de arquivo DICOM-RT

O uso de imagens é intensivo em radioterapia, o que justificou ser a primeira especialidade após a radiologia a incorporar o padrão DICOM com a criação de objetos DICOM-RT [23]. Para fins de comunicação de dados de radioterapia, o protocolo de transferência seguido é o

<sup>1</sup><http://www.dicomstandard.org>

---

**Algorithm 1** Algoritmo da função offset

---

**Entrada:** Objetivos do DVH - *pointVolumeDVH***Saída:** Objetivos do DVH editados pela função *offset* - *auxPointVolumeDVH***início**

```
if Nome da estrutura contém "reto" then  
    | downObj = 15;  
    | offset = 5;  
else  
    | downObj = 10;  
    | offset = 5;  
end  
if pointVolumeDVH  $\geq 40$  then  
    | corrFactor = downObj + offset;  
else  
    | corrFactor = downObj * pointVolumeDVH / 40;  
end  
if pointVolumeDVH  $\leq 20$  then  
    | auxPointVolumeDVH =  $\text{Min}(0.50 * \text{pointVolumeDVH}, \text{Abs}(\text{pointVolumeDVH} -$   
    | corrFactor));  
else  
    | auxPointVolumeDVH = pointVolumeDVH - corrFactor;  
end
```

**fim**

---

padrão DICOM para imagens médicas. As informações específicas da especialidade como, por exemplo, texto, gráficos, linhas de isodose e registro de imagens médicas, dão espaço à criação e ao uso de objetos específicos de radioterapia. São estes objetos: *RT Structure Set* (objeto que contém os contornos dos alvos e órgãos de risco, coordenadas de pontos de referência), *RT Plan* (objeto que contém parâmetros do plano de tratamento, tabela de tolerâncias, fracionamento de dose, dados de posicionamento do paciente), *RT Dose* (objeto que contém a matriz de dose, pontos de dose, isodoses, DVH), *RT Image* (objeto que contém DRR - *Digitally Reconstructed Radiograph*), e *RT Treatment Record*, o qual é dividido posteriormente em *RT Beams Treatment Record*, e *RT Treatment Summary Record*.

### 4.3.3 Biblioteca Pydicom

Pydicom <sup>2</sup> é uma biblioteca DICOM para manipular arquivos DICOM, como imagens médicas, relatórios e objetos de radioterapia. Esta biblioteca possibilita a leitura destes arquivos complexos em estruturas na linguagem de programação Python, facilitando sua

---

<sup>2</sup><https://pydicom.github.io/>

manipulação. Também, estes arquivos modificados podem posteriormente ser escritos em formato DICOM.

### 4.3.4 Biblioteca Scikit-Learn

*Scikit-learn*<sup>3</sup> é uma biblioteca de aprendizado de máquina de código aberto para a linguagem de programação Python. Compreende uso de algoritmos de classificação, regressão, agrupamento, redução de dimensionalidade, seleção de modelo e pré-processamento.

### 4.3.5 ESAPI

*Eclipse Scripting API (ESAPI)* é uma API (*Application Programming Interface*) que é executada no sistema de planejamento Eclipse. Essa API permite que programadores criem rotinas em C#.NET, dlls e programas que podem ler dados e realizar operações em arquivo de um paciente específico aberto no sistema de planejamento ou em uma série de pacientes existentes no banco de dados [24], como esquematizado na Figura 4.6[24]. Através desta ferramenta se pode criar relatórios, criar programas com interface gráfica em que o usuário pode interagir e automatizar tarefas, como inserção de campos de tratamento, manipulação de contornos, dentre outras.

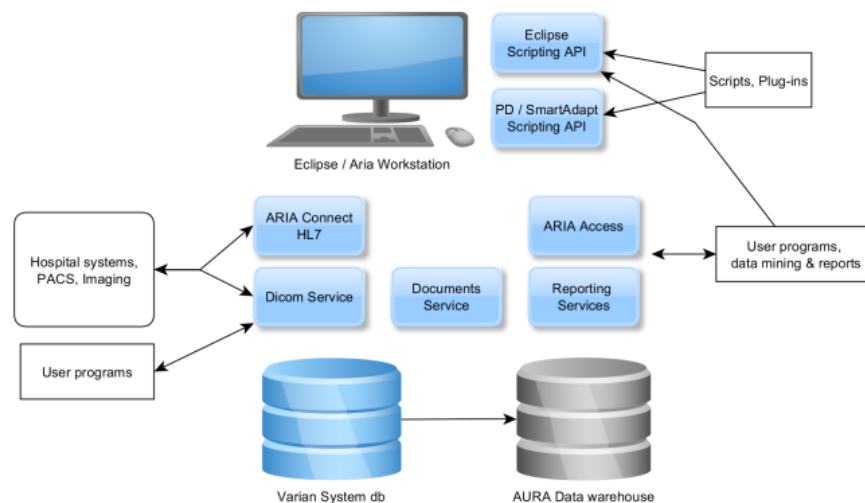


Figura 4.6: Representação esquemática do ambiente VARIAN/Eclipse e das interfaces de acesso ao banco de dados através do ESAPI .

<sup>3</sup><http://scikit-learn.org/>

# Capítulo 5

## Experimentos

Este capítulo descreve os experimentos realizados de acordo com a metodologia apresentada.

### 5.1 Modelos

Os algoritmos utilizados para predição do DVH (Regressão Linear Múltipla, Regressão por *Random Forest* e Regressão por *Support Vector*) das estruturas reto, bexiga, reto\_opt e bexiga\_opt foram testados para avaliar o seu desempenho. Para o sub-conjunto de dados de teste, foram plotados um gráfico dos pontos de DVH observado vs predito. Aqui, idealmente, se teria um gráfico linear com coeficiente angular igual a 1. Também se analisou, para este experimento, os valores de  $R^2$ , o erro médio absoluto e a raiz quadrada do erro médio quadrático.

Posteriormente, cada modelo foi aplicado aos parâmetros dos 63 casos disponíveis. Analisou-se, então, as áreas dos histogramas predito e real, através de teste estatístico para avaliar se as áreas eram iguais ou diferentes e através da Correlação de Pearson.

Através destes experimentos, se escolheu o algoritmo que mostrava os melhores resultados.

### 5.2 Auto Planning

Para avaliação do processo completo de automação do planejamento, foram selecionados 10 novos casos, contendo os arquivos de tomografia computadorizada, contornos de OAR e contorno de volume alvo. Estes casos não foram utilizados na etapa de treinamento do modelo. Os contornos de OAR e dos alvos foram realizados por médicos especialistas em radio-oncologia. Solicitou-se, então, que quatro físicos especialistas executassem a etapa de planejamento destes casos, desde a composição dos campos, passando pela criação dos

volumes auxiliares de otimização, otimização e avaliação dos objetivos do planejamento. Estes físicos são especialistas em física da radioterapia, sendo que 2 deles possuem mais de 10 anos de experiência e os outros 2 possuem entre 5 e 10 anos de experiência.

Em seguida, utilizando o programa KBP.dll, gerou-se o auto-planejamento para cada um dos 10 casos.

Tendo-se, então, estes cinco planejamentos por caso, foi:

1. avaliado se todos os planejamentos atendiam aos objetivos do planejamento, quanto a dose entregue no volume alvo e se eram respeitados os limites de dose nos OARs;
2. feita comparação das áreas entre os DVHs de reto e bexiga entre os planejamentos;
3. realizada comparação ponto a ponto do DVH de reto e bexiga entre os planejamentos;
4. avaliado o tempo salvo com a utilização do programa KBP.dll;
5. avaliada a variabilidade entre os DVHs dos planejamentos gerados pelos especialistas.

Os itens acima avaliados se relacionam com os objetivos específicos da seguinte forma: os itens 1-3 se relaciona ao objetivo de demonstrar que o autoplanejamento não é inferior ao planejamento feito pelo especialista. O item 4 reflete o tempo salvo pelo planejador e, conseqüentemente, o aumento de capacidade operacional do departamento de radioterapia. Já o item 5 mostra a variabilidade que ocorre entre planejamentos, o que não ocorreria com o um modelo automático gerando planejamentos, levando então à maior padronização e consistência entre planejamentos de distintos pacientes.

# Capítulo 6

## Resultados

Este capítulo descreve os resultados obtidos nos experimentos. Em um primeiro momento são descritos os resultados dos modelos de aprendizado de máquina criados e, em seguida, a comparação do auto-planejamento com os planejamentos dos especialistas.

### 6.1 Modelos de Aprendizado de Máquina

Os resultados desta seção visam demonstrar o processamento dos dados e os algoritmos de aprendizado de máquina testados, seguido da descrição da escolha dos algoritmos com melhor desempenho.

No seção do Apêndice, da Figura 2 à Figura 5 mostra-se um resumo da análise exploratória dos dados utilizados para predição de D50 (dose em porcentagem correspondente a 50% do volume no DVH) e V\_10-50 (fração do volume do OAR que estão entre 10% e 50% no DVH) para os OAR reto, bexiga, reto\_opt e bexiga\_opt. Tem-se, para cada OAR, um gráfico de correlação das variáveis independentes e uma descrição estatística dos dados.

Na Figura 6.1 tem-se, para cada OAR, as métricas para análise dos modelos utilizados para prever D50 e V\_10-50. Estas métricas foram calculadas aplicando-se o modelo de aprendizagem de máquina no conjunto de variáveis independentes e comparando os resultados com os valores esperados de D50 e V\_10-50.

Pode-se notar que em todas as situações a predição com o algoritmo de Regressão por *Random Forest* teve melhor desempenho que a Regressão Linear Múltipla. Para o algoritmo de Regressão por *Random Forest*, através da análise do erro médio absoluto, o qual ficou abaixo de 5%, nota-se que o desempenho está satisfatório levando-se em conta o contexto da aplicação, pois esta diferença entre pontos não leva a DVHs discrepantes.

Contidas no Apêndice, da Figura 6 à Figura 9 mostra-se um resumo da análise exploratória dos dados utilizados para predição de DVH para os OAR reto, bexiga, reto\_opt

	D50		V_10-50	
Reto	Modelo linear	Regressão por RF	Modelo linear	Regressão por RF
$R^2$	0.62	0.94	0.62	0.94
Erro médio absoluto	4.63	1.80	4.63	1.80
Raiz do erro quadrático médio	5.84	2.38	5.84	2.38
<b>Bexiga</b>				
$R^2$	0.33	0.89	0.40	0.89
Erro médio absoluto	6.38	3.07	16.09	4.20
Raiz do erro quadrático médio	7.69	2.48	9.78	2.48
<b>Reto_opt</b>				
$R^2$	0.67	0.95	0.74	0.96
Erro médio absoluto	4.04	1.55	25.71	1.56
Raiz do erro quadrático médio	4.97	2.01	6.98	2.87
<b>Bexiga_opt</b>				
$R^2$	0.89	0.98	0.94	0.98
Erro médio absoluto	2.50	0.77	19.84	0.77
Raiz do erro quadrático médio	3.16	1.21	3.56	2.06

Figura 6.1: Resumo da análise dos modelos de predição para D50 e V\_10-50.

e bexiga\_opt. Tem-se, para cada OAR, um gráfico de correlação das variáveis independentes e uma descrição estatística dos dados.

Na Figura 6.2 tem-se, para cada OAR, as métricas para análise dos modelos utilizados para predizer os DVHs. Estas métricas foram calculadas aplicando-se o modelo de aprendizagem de máquina no sub-conjunto de teste das variáveis independentes e comparando os resultados com os valores observados. Através da análise de  $R^2$  e do erro médio absoluto, nota-se que em todas as situações a predição com o algoritmo de Regressão por *Random Forest* com o uso de *grid\_search* teve melhor desempenho que os demais algoritmos. Também, através da análise do erro médio absoluto, que ficou abaixo de 0.5% para *Random Forest* com *grid\_search*, conclui-se que há boa concordância para os valores gerados por este algoritmo.

Ainda para ilustrar o desempenho de cada modelo para predizer o DVH, os gráficos representados da Figura 10 à Figura 13 - contidas no Apêndice - mostram a concordância entre os pontos do DVH predito pelo observado. Idealmente o gráfico seria em formato de reta, com coeficiente angular igual a 1 e coeficiente escalar igual a 0, com coincidência dos pontos. Conclui-se, então, qualitativamente, que o desempenho do algoritmo de Regressão por *Random Forest* com *Grid\_Search* teve desempenho satisfatório e superior aos demais, pois este apresentou uma tendencia visualmente e quantitativamente linear ( $R^2 = 0.99$ ), enquanto os demais algoritmos apresentaram pontos mais dispersos no gráfico. Nota-se também nestas figuras que o algoritmo SVR, para os OAR bexiga e bexiga\_opt, predizeu valores baixos para valores esperados próximos de 100%.

<b>Reto</b>	<b>Modelo linear</b>	<b>Regressão por RF</b>	<b>Regressão por RF com grid_search</b>	<b>SVR</b>
<b>R<sup>2</sup></b>	0.89	0.99	0.99	0.95
<b>Erro médio absoluto</b>	7.34	0.62	0.44	4.07
<b>Raiz do erro quadrático médio</b>	9.22	1.05	0.74	6.05
<b>Bexiga</b>				
<b>R<sup>2</sup></b>	0.74	0.99	0.99	0.68
<b>Erro médio absoluto</b>	8.99	0.42	0.19	7.39
<b>Raiz do erro quadrático médio</b>	12.44	0.92	0.48	13.59
<b>Reto_opt</b>				
<b>R<sup>2</sup></b>	0.85	0.99	0.99	0.95
<b>Erro médio absoluto</b>	9.27	0.57	0.52	3.76
<b>Raiz do erro quadrático médio</b>	11.47	1.03	0.80	6.50
<b>Bexiga_opt</b>				
<b>R<sup>2</sup></b>	0.67	0.99	0.99	0.61
<b>Erro médio absoluto</b>	10.42	0.48	0.23	7.97
<b>Raiz do erro quadrático médio</b>	14.14	1.02	0.55	15.37

Figura 6.2: Resumo da análise dos modelos de predição para dos DVHs.

Para investigar a ocorrência de *overfitting* nos modelos preditores do DVH, utilizou-se *cross-validation* com  $k = 5$  com o modelo de *Random Forest*. Na Figura 6.3 tem-se os resultados. Pode-se concluir que não há *overfitting*, pois os *scores* entre os 5 *folds* estão próximos.

	<b>k=1</b>	<b>k=2</b>	<b>k=3</b>	<b>k=4</b>	<b>k=5</b>	<b>Média</b>
<b>Reto</b>	0.97	0.99	0.98	0.98	0.98	0.98
<b>Bexiga</b>	0.98	0.99	0.99	0.99	0.97	0.98
<b>Reto_opt</b>	0.97	0.99	0.99	0.98	0.98	0.98
<b>Bexiga_opt</b>	0.99	0.98	0.99	0.98	0.96	0.98

Figura 6.3: Resultados da análise de *cross-validation* para a predição dos DVHs com *Random Forest*.

Por fim, aplicou-se o algoritmo de Regressão por *Random Forest* com *grid\_search* em todo o conjunto de dados das variáveis independentes para predição dos DVHs das OAR de interesse. Comparou-se, então, a área do DVH predito pela área do DVH observado. Qualitativamente, pela figura Figura 6.4, pode-se concluir que as áreas dos histogramas não estão discrepantes. Em uma análise quantitativa, aplicou-se a Correlação de Pearson e se fez um teste estatístico, em que a hipótese nula era a de que as áreas dos histogramas predito e real eram iguais. Os resultados estão na Figura 6.5. O *p-value* ficou em torno de 0.980 para todos os OAR, logo não podemos rejeitar a hipótese nula e podemos concluir que as áreas são iguais.



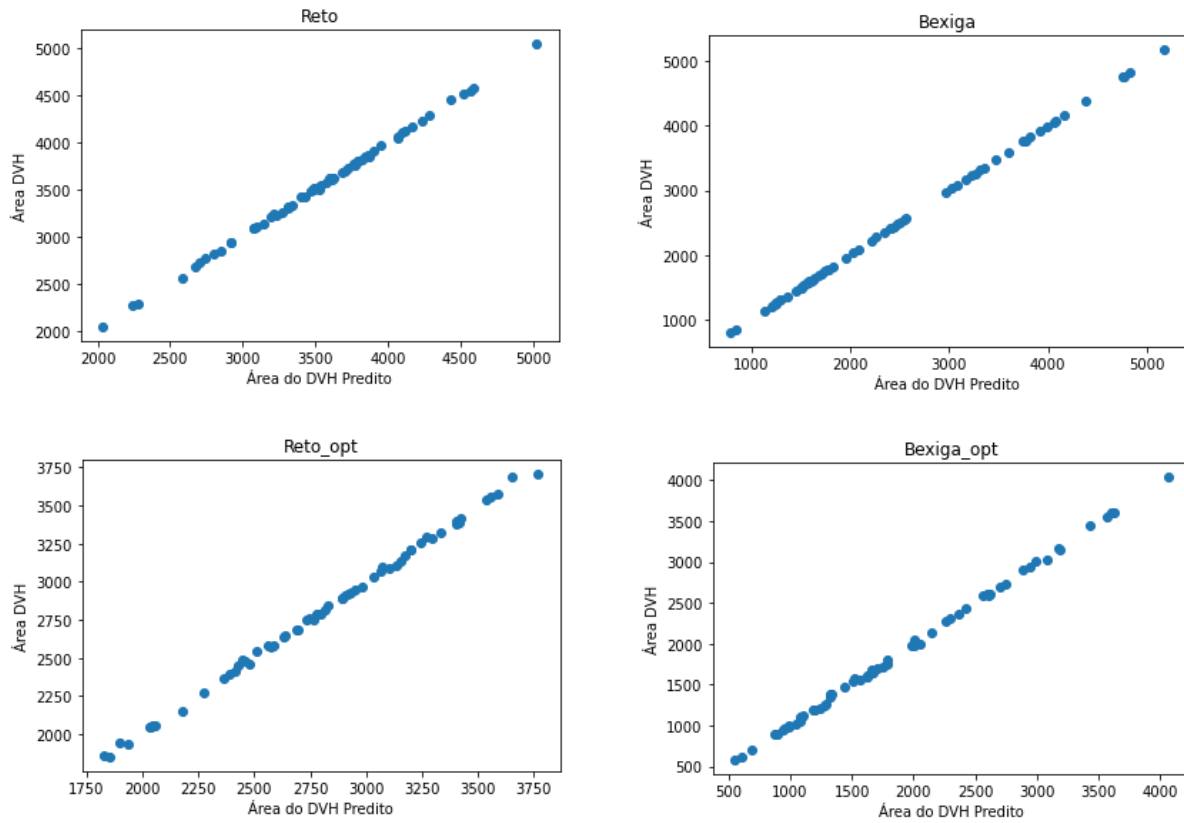


Figura 6.4: Áreas dos DVHs predito vs real para o modelo RF com *grid\_search*.

	Correlação de Pearson	p-value do teste-t
<b>Reto</b>	0.999	0.983
<b>Bexiga</b>	0.999	0.993
<b>Reto_opt</b>	0.998	0.992
<b>Bexiga_opt</b>	0.990	0.999

Figura 6.5: Áreas dos DVHs predito vs real para o modelo RF com *grid\_search*.

Desta forma, o algoritmo de aprendizado de máquina selecionado para a predição de D50 e V<sub>10-50</sub> foi a Regressão por *Random Forest*, enquanto que o algoritmo de aprendizado de máquina selecionado para a predição dos DVHs foi a Regressão por *Random Forest* com *grid\_search*.

## 6.2 Auto-Planning

Os resultados nesta seção servem de guia para analisar se os planejamentos gerados de forma automática possuem qualidade similar aos planejamentos dos especialistas.

O processo de planejamento pelos especialistas foi manual e iterativo com o *software* de planejamento. Em média, o tempo médio gasto pelos especialistas para realizar o planejamento foi de 60 minutos. Este dado foi fornecido pelos próprios especialistas.

Para ilustrar os planejamentos realizados pelos especialistas e a solução gerada pelo auto-planejamento, na Figura 6.6 tem-se os histogramas de reto (na cor marrom referente aos planos gerados pelos especialistas e em destaque na cor verde referente ao gerado pelo auto-planejamento) e bexiga (em amarelo referente aos planos dos especialistas e em destaque na cor vermelha referente ao gerado pelo auto-planejamento) de um dos casos. Para cada OAR, há a sobreposição de cinco curvas de histograma, correspondentes aos cinco planejamentos realizados, em que quatro foram feitos por físicos especialistas e o quinto pelo solução de auto-planejamento desenvolvida neste trabalho.



Figura 6.6: Sobreposição dos DVHs dos cinco Planejamentos para um dos Casos.

Inicialmente, foi analisado se todos os planejamentos haviam conseguido atingir os objetivos do planejamento, em termos de dose entregue no volume alvo e limites de dose nos OAR. Foi constatado que em todos os casos os físicos especialistas atingiram os objetivos. Já a solução de auto-planejamento atingiu em 9 dos 10 casos. Em um dos casos - referido como caso 10 da Figura 6.7 à Figura 6.12, a dose nos OARs estava dentro dos limites, porém o objetivo  $D_{99\%} \geq 57\text{Gy}$  para o PTV\_ovlp não foi alcançado (Figura 2.5), ficando 1% abaixo do esperado.

Analisou-se então as áreas dos histogramas. Isto porque, de modo geral, quanto menor a área de um DVH de um OAR, menor a dose que ele recebe, melhor assim a

qualidade do planejamento - assumindo que se mantém a dose desejada no volume alvo. Normalizando pela área do auto-planejamento, tem-se a razão das áreas dos histogramas de reto e bexiga dos planejamentos gerados pelos físicos especialistas mostrados nas Figura 6.7 e Figura 6.8. Valores maiores que 1 indicam que a área do DVH do respectivo OAR do físico especialista se mostrou maior que a área do DVH do autoplanejamento, ao passo que valores menores que 1 indicam que a área do DVH do respectivo OAR do físico especialista se mostrou menor que a área do DVH do autoplanejamento. Nota-se que as áreas dos histogramas dos OAR dos planejamentos da solução deste trabalho, ficou igual ou ligeiramente inferior às áreas dos planejamentos gerados pelos físicos especialistas.

<b>Caso</b>	<b>Fisico I</b>	<b>Fisico II</b>	<b>Fisico III</b>	<b>Fisico IV</b>
<b>1</b>	1,02	1,21	1,17	1,22
<b>2</b>	0,99	1,08	0,99	1,17
<b>3</b>	1,05	0,99	1,17	1,01
<b>4</b>	1,01	0,89	1,06	0,90
<b>5</b>	1,07	0,96	1,24	1,19
<b>6</b>	1,03	1,03	1,21	0,97
<b>7</b>	1,03	0,98	1,18	1,04
<b>8</b>	1,12	1,12	1,20	1,18
<b>9</b>	0,93	0,95	1,00	0,96
<b>10</b>	0,96	0,97	1,20	0,91
<b>Média</b>	1,02	1,02	1,14	1,06

Figura 6.7: Razão das áreas dos DVHs de reto dos planejamentos dos especialistas (numerador) em relação ao auto-planejamento (denominador).

Já nas figuras Figura 6.9 e Figura 6.10 se vê a comparação ponto a ponto dos histogramas dos planejamentos dos especialistas com a do auto-planejamento. Conta-se, em porcentagem, quantas vezes os pontos dos histogramas do auto-planejamento ficou 5% ou mais acima do histograma do respectivo especialista. Nota-se que, em geral, o auto-planejamento poucos pontos de DVH com valor acima de 5%. Isto aconteceu de forma relevante apenas para o OAR da bexiga no caso 4, em que a comparação mostrou que em relação aos 4 dos planos gerados pelos especialistas, o DVH do auto-planejamento gerou histogramas com porcentagem de pontos acima de 5% entre 16% e 43%. No entanto, essas discrepâncias entre os pontos não passaram de < 10% em comparação com 3 especialistas e os objetivos do planejamento foram satisfeitos, de modo que se tratava ainda de um plano clinicamente viável. Acrescenta-se que, exceto no caso 10, em nenhum dos casos houve um ponto do histograma gerado pelo auto-planejamento com valor maior que 10% em relação ao histograma do especialista para o volume do reto.

<b>Caso</b>	<b>Fisico I</b>	<b>Fisico II</b>	<b>Fisico III</b>	<b>Fisico IV</b>
<b>1</b>	1,03	1,00	1,16	1,13
<b>2</b>	1,02	1,03	0,98	1,03
<b>3</b>	1,00	0,96	0,99	0,94
<b>4</b>	0,90	0,83	0,85	0,82
<b>5</b>	1,06	1,01	1,01	1,07
<b>6</b>	1,05	0,96	1,03	1,04
<b>7</b>	0,96	1,02	1,09	1,05
<b>8</b>	1,02	0,96	1,15	0,91
<b>9</b>	1,04	1,02	0,92	1,07
<b>10</b>	0,99	1,00	0,96	0,98
<b>Média</b>	1,01	0,98	1,01	1,01

Figura 6.8: Razão das áreas dos DVHs de bexiga dos planejamentos dos especialistas (numerador) em relação ao auto-planejamento (denominador).

<b>Caso</b>	<b>Fisico I</b>	<b>Fisico II</b>	<b>Fisico III</b>	<b>Fisico IV</b>
<b>1</b>	0	0	0	3
<b>2</b>	0	0	0	0
<b>3</b>	0	0	0	0
<b>4</b>	0	0	0	0
<b>5</b>	0	0	0	0
<b>6</b>	0	0	0	0
<b>7</b>	0	0	0	0
<b>8</b>	0	0	0	0
<b>9</b>	0	0	0	0
<b>10</b>	14	0	0	21

Figura 6.9: Porcentagem das vezes em que um ponto do DVH de reto do auto-planejamento é superior em 5 % ou mais do mesmo ponto no DVH elaborado pelo especialista.

A seguir, comparou-se a diferença entre as áreas dos histogramas gerados pela solução de auto-planejamento com a área dos histogramas gerados pelos especialistas. Na Figura 6.11 observa-se que, à exceção do DVH de bexiga no caso 4, as áreas dos histogramas dos planejamentos gerados pela solução deste trabalho não esteve sensivelmente acima das áreas dos DVHs dos planejamentos dos especialistas. Nos valores mostrados, valores negativos indicam que a área do DVH do autoplanejamento se mostrou menor que a média das áreas dos DVHs dos especialistas, ao passo que valores positivos indicam que a área do DVH do autoplanejamento se mostrou maior que a média das áreas dos DVHs

<b>Caso</b>	<b>Fisico I</b>	<b>Fisico II</b>	<b>Fisico III</b>	<b>Fisico IV</b>
<b>1</b>	4	0	4	0
<b>2</b>	4	0	4	0
<b>3</b>	4	3	5	0
<b>4</b>	16	44	43	46
<b>5</b>	0	0	0	0
<b>6</b>	0	0	2	0
<b>7</b>	0	0	0	0
<b>8</b>	0	0	0	10
<b>9</b>	0	0	0	0
<b>10</b>	0	3	4	3

Figura 6.10: Porcentagem das vezes em que um ponto do DVH de bexiga do auto-planejamento é superior em 5 % ou mais do mesmo ponto no DVH elaborado pelo especialista.

dos especiliastas.

<b>Caso</b>	<b>Reto</b>	<b>Bexiga</b>
<b>1</b>	-15,5%	-8,1%
<b>2</b>	-5,9%	-1,2%
<b>3</b>	-5,6%	2,6%
<b>4</b>	3,3%	14,9%
<b>5</b>	-11,4%	-4,0%
<b>6</b>	-5,9%	-1,9%
<b>7</b>	-5,8%	-3,0%
<b>8</b>	-15,4%	-0,9%
<b>9</b>	4,0%	-1,1%
<b>10</b>	-1,0%	1,7%

Figura 6.11: Média da diferença percentual entre as áreas dos histogramas dos especialistas em comparação com as áreas dos histogramas gerados pelo autoplanejamento. Valores negativos indicam área do DVH do autoplanejamento menor que a área do DVH do especialista.

Por fim, comparou-se a variabilidade entre as áreas dos histogramas gerados pelos especialistas (Figura 6.12), através da análise das áreas dos DVHs e normalizando pelas

áreas dos histogramas dos planejamentos gerados pelo físico I. Nota-se que os valores variam entre 1,24 e 0,89. Estes resultados mostram que, para um mesmo paciente, pode haver alta variabilidade entre os planejamentos gerados pelos especialistas. Tal fato pode ser minimizado com a adoção de uma solução de autoplanejamento baseado no padrão institucional.

<b>Caso</b>	<b>Reto</b>				<b>Bexiga</b>			
	<b>Físico I</b>	<b>Físico II</b>	<b>Físico III</b>	<b>Físico IV</b>	<b>Físico I</b>	<b>Físico II</b>	<b>Físico III</b>	<b>Físico IV</b>
<b>1</b>	1,00	1,19	1,15	1,20	1,00	0,96	1,12	1,10
<b>2</b>	1,00	1,09	1,00	1,19	1,00	1,01	0,96	1,01
<b>3</b>	1,00	0,95	1,11	0,96	1,00	0,96	0,99	0,94
<b>4</b>	1,00	0,89	1,06	0,90	1,00	0,93	0,94	0,91
<b>5</b>	1,00	0,89	1,16	1,11	1,00	0,95	0,95	1,01
<b>6</b>	1,00	1,00	1,17	0,94	1,00	0,91	0,98	0,99
<b>7</b>	1,00	0,94	1,14	1,00	1,00	1,06	1,14	1,10
<b>8</b>	1,00	1,00	1,07	1,05	1,00	0,95	1,13	0,90
<b>9</b>	1,00	1,02	1,07	1,04	1,00	0,98	0,89	1,03
<b>10</b>	1,00	1,00	1,24	0,95	1,00	1,01	0,97	1,00

Figura 6.12: Variabilidade entre as áreas dos DVHs dos especialistas para os volumes de reto e bexiga, normalizado pelas áreas dos DVHs dos planejamentos gerados pelo físico I.

# Capítulo 7

## Conclusão e Trabalhos Futuros

Este capítulo descreve as conclusões deste trabalho e destaca a contribuição dada. Por fim, comenta sobre os trabalhos futuros.

### 7.1 Conclusão

Neste trabalho foi criado uma aplicação que, utilizando aprendizado de máquina, busca automatizar planejamentos de radioterapia de próstata.

Os modelos foram gerados pela técnica de KBP e os algoritmos escolhidos foram os de Regressão por *Random Forest*. Estes algoritmos, quando aplicados aos dados utilizados no treinamento, predizeram com boa concordância das áreas os novos histogramas (Correlação de Pearson = 0.999), e também com boa concordância entre os pontos preditos e observados ( $R^2 = 0.99$  e erro médio absoluto  $< 0.50\%$ ). Em seguida, foram aplicados a outros casos ainda não vistos na etapa de aprendizado do algoritmo e comparados com os planejamentos de 4 físicos especialistas, tidos aqui como o padrão. Comparando-se os planejamentos, em 9 dos 10 casos a qualidade dos DVHs dos OARs de reto e bexiga resultantes da aplicação de auto-planejamento não foram inferiores em relação aos planejamentos dos especialistas. A comparação das áreas e a comparação ponto-a-ponto dos histogramas mostrou boa concordância, ao mesmo tempo em que os objetivos do planejamento foram alcançados.

No caso em que um dos objetivos de dose entregue ao volume alvo não foi satisfeito (caso 10), uma normalização de 1% no planejamento já faria o objetivo de cobertura do alvo ser satisfeito. Neste mesmo caso 14% e 21% dos pontos do histograma de reto da solução do autoplanejamento estiveram acima de 5% em comparação com os planejamentos gerados pelo físico I e físico IV, respectivamente. No entanto, todos os objetivos de proteção ao OAR foram alcançados. Cita-se ainda que no caso 4 o DVH de bexiga da solução do auto-planejamento teve desempenho inferior em relação aos especialistas,

porém os objetivos de proteção ao OAR também foram satisfeitos. Acredita-se que a causa do algoritmo ter um desempenho inferior nestes casos seja pelo fato do processo de otimização já entrar com valores elevados da função custo desde o início, o que pode atrapalhar a convergência da função custo. Isto porque a fluência dos campos se inicia homogênea e já se tem os pontos objetivos ótimos do DVH. Devido a grande diferença entre os valores e o conseqüente alto valor da função custo, há a possibilidade da função convergir para um mínimo local dentro do espaço das variáveis da função de otimização. Vale lembrar que no processo manual os objetivos são adicionados e editados múltiplas vezes, à medida que a convergência acontece. No cenário do auto-planejamento, todos os objetivos (mesmo que conflitantes inicialmente) são adicionados no início da otimização.

Dados os resultados, pode-se concluir que os modelos estavam adequados, assim como os parâmetros utilizados para treinamento e predição dos histogramas. Na Figura 7.1 tem-se um corte axial da sobreposição dos contornos com a distribuição de dose gerada pela solução de auto-planejamento. Pode-se observar que a solução propicia boa proteção ao OAR em questão (reto), reproduzindo o formato côncavo da distribuição de dose esperado nesse tipo de caso. Qualitativamente tem-se aí uma distribuição de dose similar a observada nos casos clínicos realizados na rotina pelo especialista.

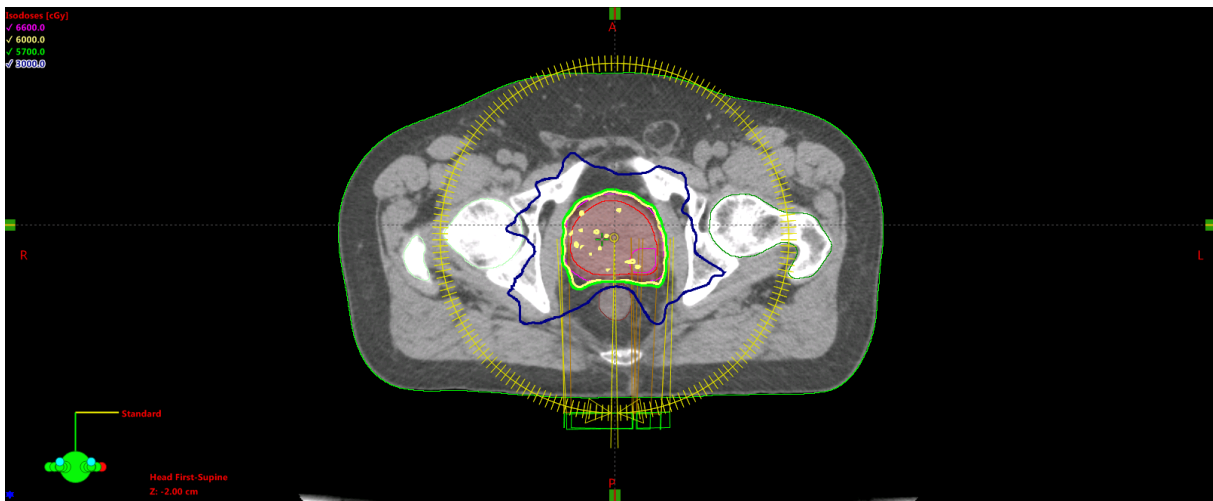


Figura 7.1: Distribuição de dose gerada pela solução de auto-planejamento. Em destaque o formato côncavo das isodose para proteger o OAR reto.

Conclui-se, também, que os objetivos específicos foram alcançados. Os resultados mostraram que o autoplanejamento proveu soluções não inferiores e planejamentos clinicamente aceitáveis em 90% dos casos, sendo necessário pequenos ajustes para que todos os objetivos fossem alcançados. Isto permite aumento da capacidade operacional, já que o tempo de interação do planejador é consideravelmente menor. Também, a adoção de um sistema automático tira a dependencia do planejador quanto a qualidade do planeja-



mento, pois permite padronizar a abordagem dos métodos de planejamento, minimizando a variabilidade entre resultados de DVHs entre planejamentos de diferentes planejadores para um mesmo paciente.

Destaca-se também a facilidade de uso da aplicação. Tendo os delineamentos de OARs e volumes alvo criados, em poucos cliques se tem o planejamento sendo realizado. Uma vez que a aplicação está acoplada ao sistema de planejamento e os modelos de aprendizado de máquina são carregados pelo programa principal, todas as sub-tarefas são invisíveis ao usuário. Isso implica que um profissional não especialista em planejamento pode operar a solução de auto-planejamento. Ao final dos cálculos, tendo uma mensagem gerada a partir da avaliação automática dos objetivos do planejamento, o caso pode ser direcionado para avaliação/aprovação se todos os objetivos forem alcançados, ou para edição por um especialista se todos os critérios não forem obtidos. Assim, tem-se um modelo híbrido de automação com edição do especialista.

Para ser utilizado em uma outra clínica, a premissa principal é que seja utilizado o mesmo sistema de planejamento. Uma vez adequada a nomenclatura dos contornos, não haveria dificuldade para a aplicação ser utilizada em ambiente externo, uma vez que os modelos são aplicados por um executável, não havendo dependência de instalação de bibliotecas. Desta forma, esta solução é aplicável em qualquer departamentos de radioterapia que possua ambiente Varian.

A contribuição deste trabalho é a de se ter uma solução nacional que utiliza Inteligência Artificial para automação de planejamentos. Esta tecnologia aumenta a capacidade de planejamentos de um departamento de radioterapia e pode contribuir tando do ponto de vista de saúde pública, ao auxiliar a ampliar o acesso à radioterapia aos pacientes oncológicos, como também aos serviços privados, ao tornar os processos mais eficientes. No contexto de um país de economia emergente como o Brasil isso tem maior relevância, pois a oferta do serviço público nem sempre atende a demanda da população, e soluções que otimizam e automatizam processos podem ajudar nos ajuste entre oferta e demanda. Tanto para serviços públicos como privados, esta tecnologia é relevante também por se ter uma solução nacional, a qual não está sujeita a variações de câmbio, o que invariavelmente limita investimentos em tecnologia.

## 7.2 Trabalhos futuros

Como trabalho futuro, cita-se outra possível abordagem para predição de DVH: fazê-la em apenas um passo - por exemplo, utilizando regressão funcional-, e não em 2 passos, como é realizado no método utilizado neste trabalho.

Ainda, como trabalho futuro pode-se desenvolver um método que avalie se o algoritmo de otimização e o planejamento gerado estacionou em um mínimo local da função de otimização.

Também, mais casos de próstata serão avaliados, a fim de se entender a qualidade de planejamento que esta solução criada oferece para pacientes com diferentes geometrias entre volume alvo e OARs.

Em seguida, será aplicada esta metodologia e este algoritmo para o desenvolvimento de auto-planejamento em outros sítios que utilizam um nível de dose para o volume alvo.

Por fim, generalizar o algoritmo para lidar com situações em que existam dois ou mais níveis de dose nos volumes alvos no mesmo planejamento.

# Appendices

## .1 Análise Descritiva e Concordância de pontos do DVH

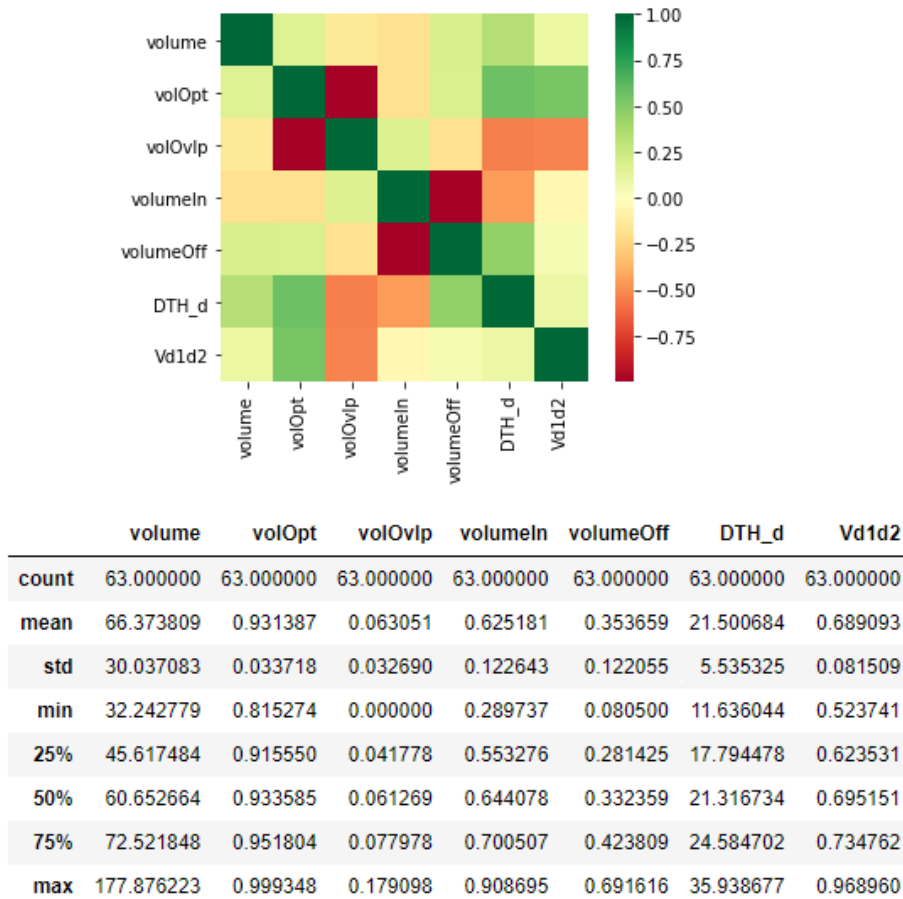


Figura 2: Análise de Correlação (acima) e Estatística Descritiva para a Predição de D50 e V\_10-50 do reto.

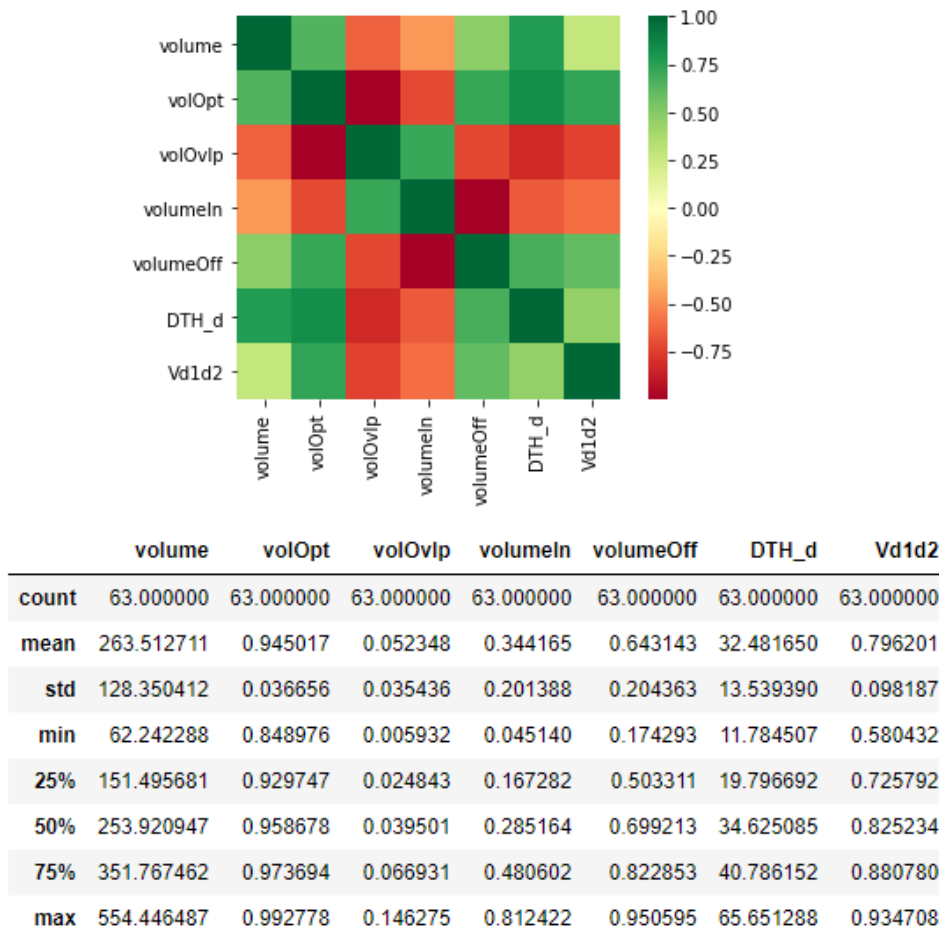
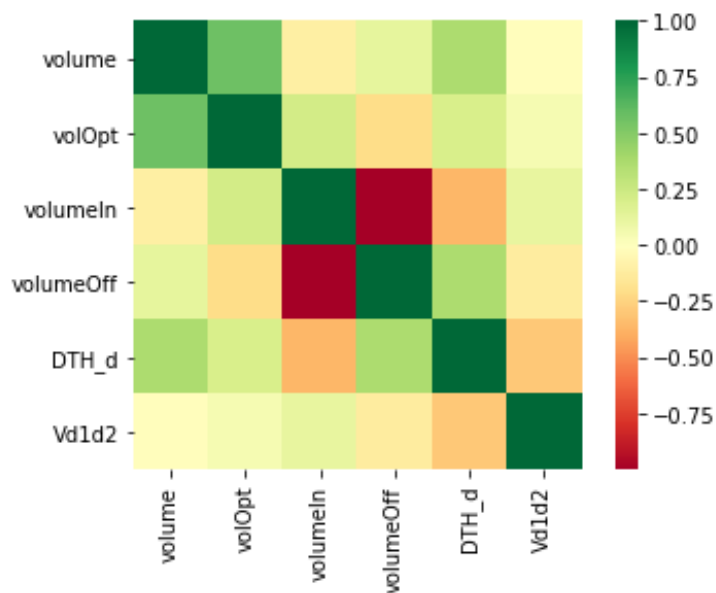
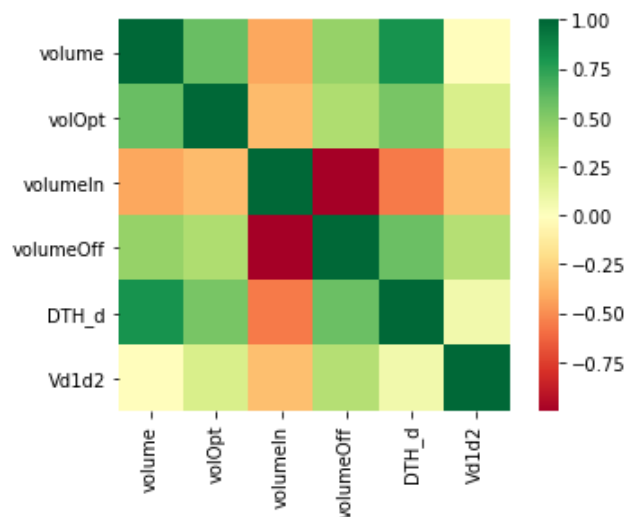


Figura 3: Análise de Correlação (acima) e Estatística Descritiva para a Predição de D50 e V\_10-50 da bexiga.



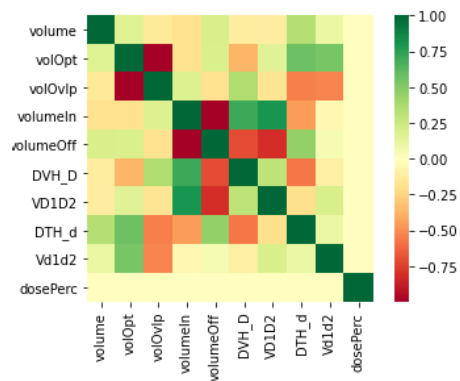
	volume	volOpt	volumeln	volumeOff	DTH_d	Vd1d2
<b>count</b>	62.000000	62.000000	62.000000	62.000000	62.000000	62.000000
<b>mean</b>	59.221794	0.999183	0.582864	0.393690	23.021897	0.778043
<b>std</b>	27.923533	0.000269	0.126650	0.125726	5.254495	0.072705
<b>min</b>	26.954827	0.998164	0.330924	0.095367	13.707789	0.584115
<b>25%</b>	41.335280	0.999068	0.516298	0.324832	20.388093	0.733789
<b>50%</b>	53.326549	0.999222	0.598217	0.375019	22.810247	0.782978
<b>75%</b>	61.900006	0.999375	0.647385	0.459921	25.335314	0.825775
<b>max</b>	165.634717	0.999597	0.891850	0.651091	37.679437	0.968876

Figura 4: Análise de Correlação (acima) e Estatística Descritiva para a Predição de D50 e V\_10-50 do reto\_opt.



	volume	volOpt	volOvlp	volumeln	volumeOff	DTH_d	Vd1d2
<b>count</b>	63.000000	63.000000	63.0	63.000000	63.000000	63.000000	63.000000
<b>mean</b>	245.428320	0.999657	0.0	0.291762	0.692397	33.951430	0.870059
<b>std</b>	127.773434	0.000113	0.0	0.195498	0.199328	13.170001	0.074638
<b>min</b>	56.080992	0.999346	0.0	0.008651	0.196080	13.152008	0.649752
<b>25%</b>	126.865026	0.999591	0.0	0.134758	0.554565	21.720395	0.831624
<b>50%</b>	241.673687	0.999691	0.0	0.240982	0.742635	35.700900	0.897686
<b>75%</b>	328.623407	0.999735	0.0	0.433742	0.853611	41.805623	0.926173
<b>max</b>	544.490187	0.999832	0.0	0.788987	0.981917	66.234551	0.960762

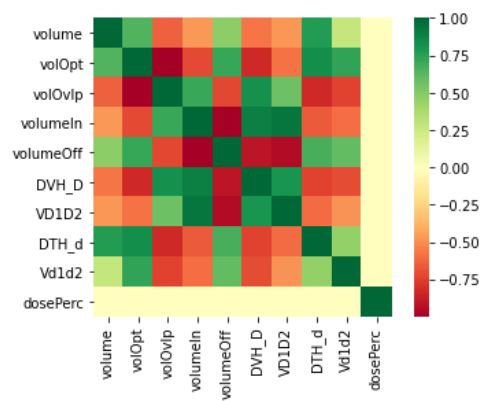
Figura 5: Análise de Correlação (acima) e Estatística Descritiva para a Predição de D50 e V<sub>10-50</sub> da bexiga<sub>opt</sub>.



	volume	volOpt	volOvlp	volumeln	volumeOff	DVH_D	VD1D2	DTH_d	Vd1d2	dosePerc
<b>count</b>	6363.000000	6363.000000	6363.000000	6363.000000	6363.000000	6363.000000	6363.000000	6363.000000	6363.000000	6363.000000
<b>mean</b>	66.373809	0.931387	0.063051	0.625181	0.353659	27.728276	43.487981	21.500684	0.689093	50.000000
<b>std</b>	29.800082	0.033452	0.032432	0.121675	0.121091	9.456463	12.673905	5.491650	0.080866	29.157051
<b>min</b>	32.242779	0.815274	0.000000	0.289737	0.080500	4.777282	18.872381	11.636044	0.523741	0.000000
<b>25%</b>	45.307743	0.914986	0.041519	0.552823	0.278667	23.141864	33.089226	17.569527	0.622573	25.000000
<b>50%</b>	60.652664	0.933585	0.061269	0.644078	0.332359	28.972395	44.120591	21.316734	0.695151	50.000000
<b>75%</b>	75.392425	0.954310	0.078118	0.704540	0.424523	34.071209	51.947443	24.782309	0.735817	75.000000
<b>max</b>	177.876223	0.999348	0.179098	0.908695	0.691616	49.742564	76.372054	35.938677	0.968960	100.000000

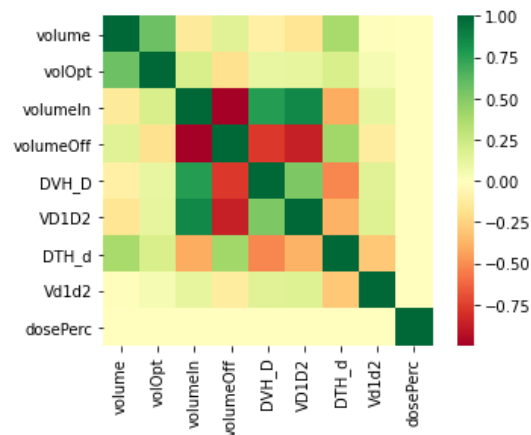
Figura 6: Análise de Correlação (acima) e Estatística Descritiva para a Predição do DVH do reto.





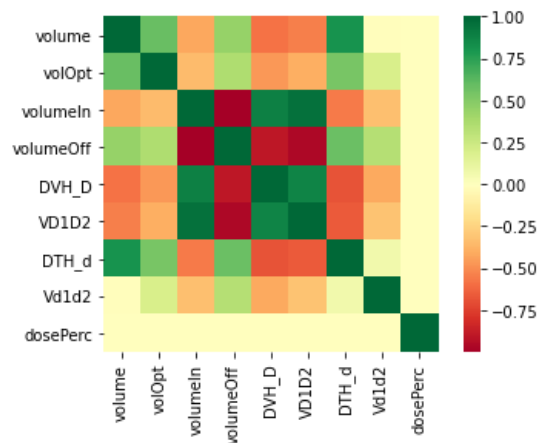
	volume	volOpt	volOvlp	volumeln	volumeOff	DVH_D	VD1D2	DTH_d	Vd1d2	dosePerc
<b>count</b>	6363.000000	6363.000000	6363.000000	6363.000000	6363.000000	6363.000000	6363.000000	6363.000000	6363.000000	6363.000000
<b>mean</b>	263.512711	0.945017	0.052348	0.344165	0.643143	13.502725	27.321502	32.481650	0.796201	50.000000
<b>std</b>	127.337689	0.036367	0.035156	0.199799	0.202750	12.733929	13.018757	13.432561	0.097413	29.157051
<b>min</b>	62.242288	0.848976	0.005932	0.045140	0.174293	2.241431	5.365704	11.784507	0.580432	0.000000
<b>25%</b>	145.676848	0.929591	0.023953	0.166234	0.487083	4.246715	16.273041	19.665670	0.721226	25.000000
<b>50%</b>	253.920947	0.958678	0.039501	0.285164	0.699213	6.813570	24.724040	34.625085	0.825234	50.000000
<b>75%</b>	354.691472	0.974517	0.067148	0.488291	0.824923	23.257822	36.861053	40.862364	0.881730	75.000000
<b>max</b>	554.446487	0.992778	0.146275	0.812422	0.950595	47.937619	66.635941	65.651288	0.934708	100.000000

Figura 7: Análise de Correlação (acima) e Estatística Descritiva para a Predição do DVH da bexiga.



	volume	volOpt	volumeln	volumeOff	DVH_D	VD1D2	DTH_d	Vd1d2	dosePerc
<b>count</b>	6363.000000	6363.000000	6363.000000	6363.000000	6363.000000	6363.000000	6363.000000	6363.000000	6363.000000
<b>mean</b>	59.653318	0.999184	0.577557	0.399059	23.426844	48.983949	23.155090	0.777726	50.000000
<b>std</b>	27.688197	0.000265	0.131452	0.130748	8.918071	13.999069	5.276126	0.071590	29.157051
<b>min</b>	26.954827	0.998164	0.248546	0.095367	4.155774	21.176125	13.707789	0.584115	0.000000
<b>25%</b>	41.294242	0.999066	0.509275	0.324611	19.664923	38.907819	20.386685	0.731375	25.000000
<b>50%</b>	53.361593	0.999225	0.597858	0.375134	24.743977	48.597367	22.853396	0.779917	50.000000
<b>75%</b>	69.029697	0.999380	0.647587	0.466857	29.638150	57.809816	25.992913	0.826366	75.000000
<b>max</b>	165.634717	0.999597	0.891850	0.731933	39.923154	80.334216	37.679437	0.968876	100.000000

Figura 8: Análise de Correlação (acima) e Estatística Descritiva para a predição do DVH do reto\_opt.



	volume	volOpt	volumeln	volumeOff	DVH_D	VD1D2	DTH_d	Vd1d2	dosePerc
<b>count</b>	6363.000000	6363.000000	6363.000000	6363.000000	6363.000000	6363.000000	6363.000000	6363.000000	6363.000000
<b>mean</b>	245.428320	0.999657	0.291762	0.692397	10.734658	30.575831	33.951430	0.870059	50.000000
<b>std</b>	126.765264	0.000112	0.193955	0.197755	9.684552	15.805771	13.066086	0.074050	29.157051
<b>min</b>	56.080992	0.999346	0.008651	0.196080	2.184225	5.507246	13.152008	0.649752	0.000000
<b>25%</b>	123.716911	0.999586	0.130857	0.552529	3.999398	16.965403	21.717958	0.831024	25.000000
<b>50%</b>	241.673687	0.999691	0.240982	0.742635	6.092071	26.765325	35.700900	0.897686	50.000000
<b>75%</b>	328.649651	0.999736	0.435148	0.857559	15.565149	42.559066	41.844593	0.927525	75.000000
<b>max</b>	544.490187	0.999832	0.788987	0.981917	38.833289	78.141225	66.234551	0.960762	100.000000

Figura 9: Análise de Correlação (acima) e Estatística Descritiva para a Predição do DVH da bexiga\_opt.

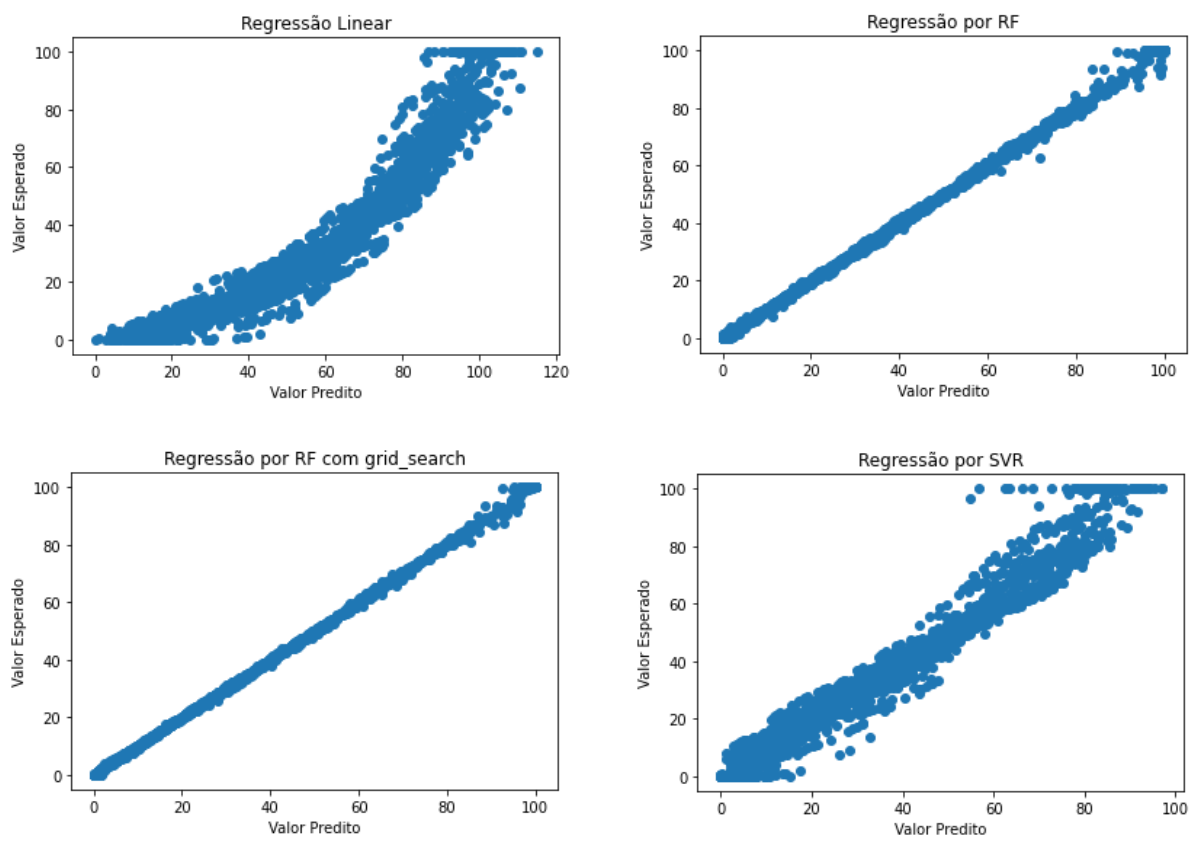


Figura 10: Concordância entre os pontos de DVH de reto predito vs observado.

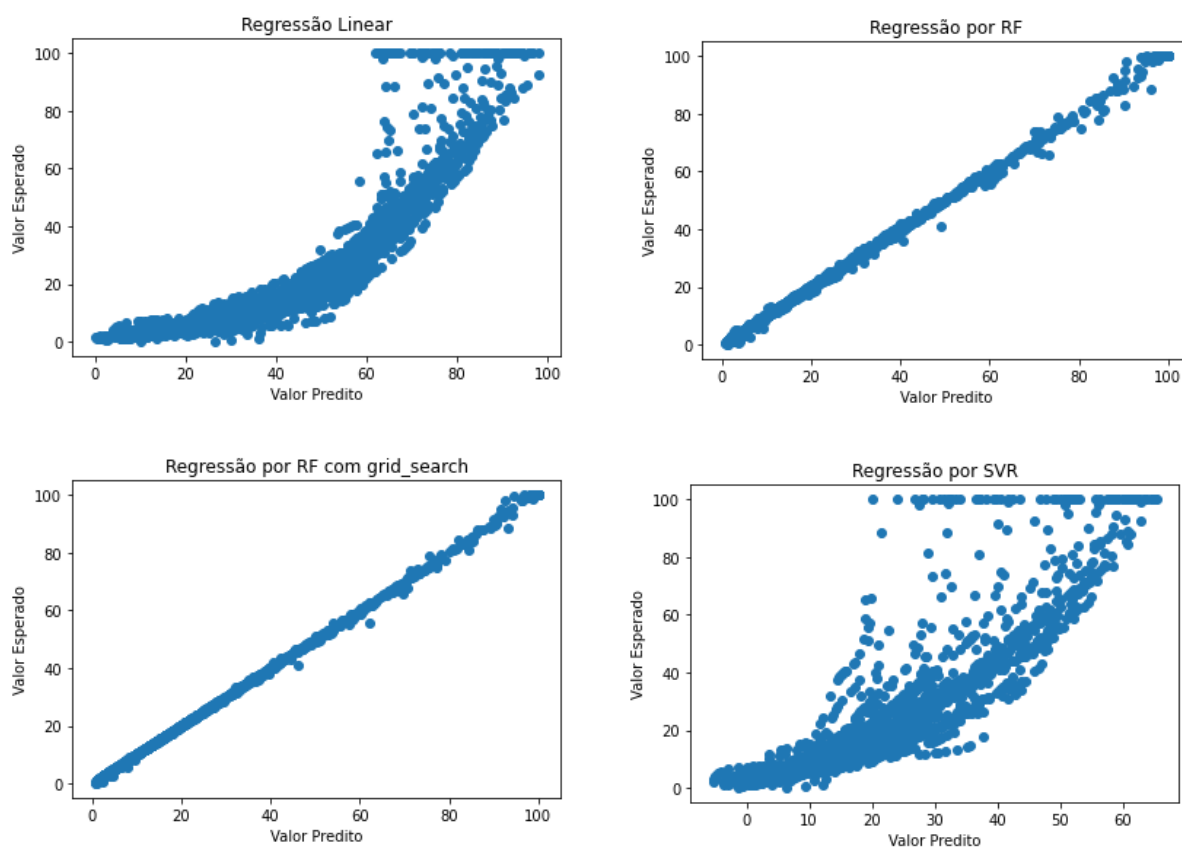


Figura 11: Concordância entre os pontos de DVH de bexiga predito vs observado.

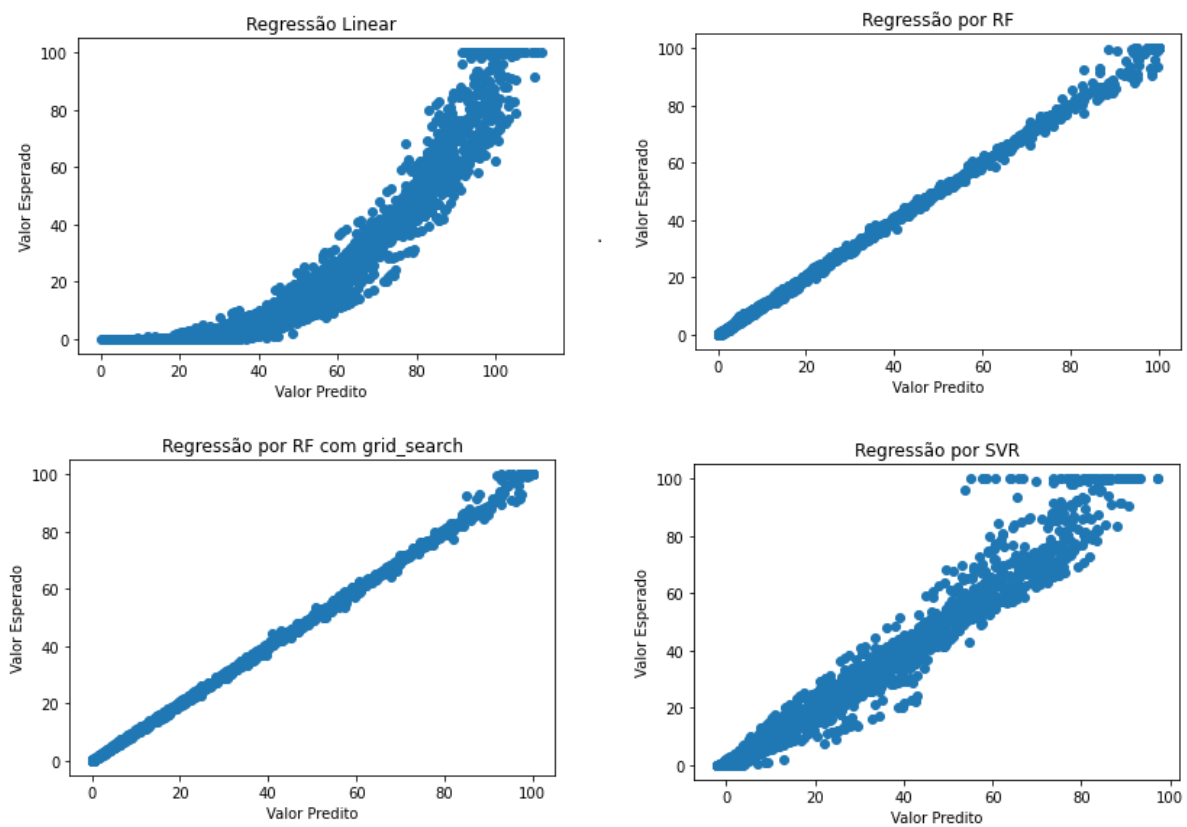


Figura 12: Concordância entre os pontos de DVH de reto\_opt predito vs observado.

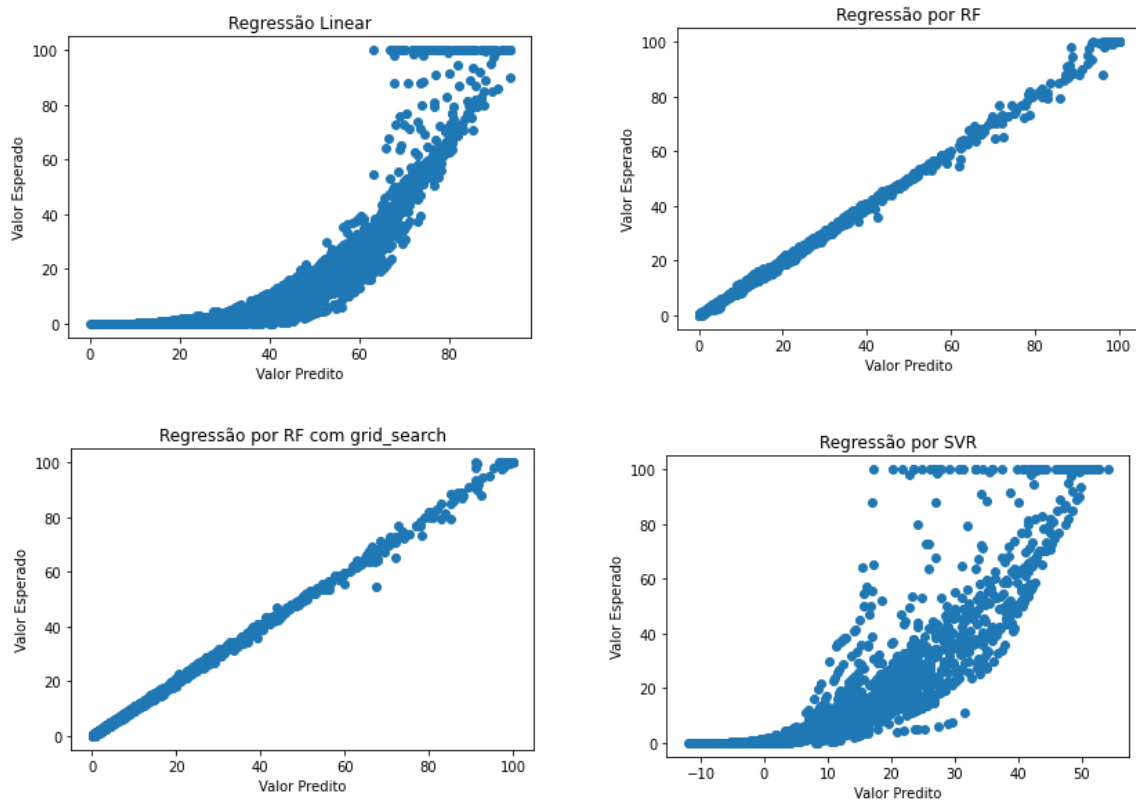


Figura 13: Concordância entre os pontos de DVH de bexiga\_opt predito vs observado.

# Referências

- [1] Yuan, Lulin, Yaorong Ge, W. Robert Lee, Fang Fang Yin, John P. Kirkpatrick e Q. Jackie Wu: *Quantitative analysis of the factors which affect the interpatient organ-at-risk dose sparing variation in imrt plans*. Medical Physics, (11):6868–6878, 2012. x, 18, 19, 25
- [2] Dearnaley, David et al.: *Conventional versus hypofractionated high-dose intensity-modulated radiotherapy for prostate cancer: 5-year outcomes of the randomised, non-inferiority, phase 3 chhip trial*. Lancet Oncology, 17:1047–1060, 2016. 2
- [3] Clark, Grant M., Richard A. Popple, Brendan M. Prendergast, Sharon A. Spencer, Evan M. Thomas, John G. Stewart, Barton L. Guthrie, James M. Markert e John B. Fiveash: *Plan quality and treatment planning technique for single isocenter cranial radiosurgery with volumetric modulated arc therapy*. Practical Radiation Oncology, (2):306–313, 2012. 6
- [4] Khan, Gerbi: *Treatment Planning in Radiation Oncology*, volume 1. Wolters Kluner, 2012. 7
- [5] Babier, Aaron, Justin J. Boutlier, Andrea L. McNiven e Timothy C. Y. Chan: *Knowledge-based automated planning for oropharyngeal cancer*. Medical Physics, 45, 2018. 9, 16
- [6] Shiraishi, Satomi, Jun Tan, Lindsey A. Olsen e Kevin L. Moore: *Knowledge-based prediction of plan quality metrics in intracranial stereotactic radiosurgery*. Medical Physics, 42(2):908–917, 2015. 10, 16
- [7] Boutlier, Justin J., Tim Craig, Michael B. Sharpe e Timothy C. Y. Chan: *Sample size requirements for knowledge-based treatment planning*. Medical Physics, 43(3):1212–1221, 2016. 10, 16
- [8] James, Gareth, Daniela Witten, Trevor Hastie e Robert Tibshirani: *An Introduction to Statistical Learning*, volume 1. Springer, 2013. 12
- [9] Géron, Aurélien: *Hands-on Machine Learning with Scikit-Learn e TensorFlow*, volume 1. O'REILLY, 2017. 12, 13, 14
- [10] al, Vorakarn et., Shiva K. Das, William R. Lee e Joseph Y. Lo: *Knowledge-based imrt treatment planning for prostate cancer*. Medical Physics, 38(5):2515–2522, 2011. 15



- [11] Schimidt, Matthew, Joseph Y. Lo, Shelby Grzetic, Carly Lutzky, David M. Brizel e Shiva K. Das: *Semiautomated head-and-neck imrt planning using dose warping and scaling to robustly adapt plans in a knowledge database containing potentially suboptimal plans*. *Medical Physics*, 42(8):4428–4434, 2015. 15
- [12] Campbell, Warren G. e Moyed Miften: *Neural network dose models for knowledge-based planning in pancreatic sbrt*. *Medical Physics*, 44(12):342–351, 2017. 15
- [13] Ziemer, Benjamin P., Parag Sanghvi, Jona Hattangadi-Gluth e Kevin L. Moore: *Heuristic knowledge-based planning for single-isocenter stereotactic radiosurgery to multiple brain metastases*. *Medical Physics*, 44(10):5001–5009, 2017. 16
- [14] Du, Qian, Chi Zhang, Xiaofeng Zhu, Xiaoying Liang, Vivek Verma, Kenneth Follet, Shuo Wang, Qiyong Fan, Rongtao Ma, Sumin Zhou e Dandan Zheng: *Application of statistical and computational methodology to predict brainstem dosimetry for trigeminal neuralgia stereotactic radiosurgery*. *Medical Physics*, 45(5):1822–1831, 2018. 16
- [15] Younge, Kelly C., Robin B. Marsh, Dawn Owen, Huaizhi Geng, Ying Xiao, Daniel E. Spratt, Joseph Foy, Krithika Suresh, Q. Jackie Wu, Fang Fang Yin, Samuel Ryu e Martha M. Matuszak: *Improving quality and consistency in nrg oncology radiation therapy oncology group 0631 for spine radiosurgery via knowledge-based planning*. *International Journal of Radiation Oncology*, 100(4):1067–1074, 2018. 16
- [16] Nwankwo, Obioma, Hana Mekdash, Dwi Seno Kuncoro Sihono, Frederik Wenz e Gerhard Glatting: *Knowledge-based radiation therapy (kbrt) treatment planning versus planning by experts: validation of a kbrt algorithm for prostate cancer treatment planning*. *Radiation Oncology*, 10(111), 2015. 17
- [17] Chatterjee, Avishek, Monica Serban, Bassam Abdulkarim, Valerie Panet-Raymond, Luis Souhami, George Shenouda, Siham Sabri, Bertrand Jean-Claude e Jan Seuntjens: *Performance of knowledge-based radiation therapy planning for the glioblastoma disease site*. *International Journal of Radiation Oncology*, 99(4):1021–1028, 2017. 17
- [18] Wang, Juanqi, Weigang Hu, Zhaozhi Yang, Chen Xiaohui, Zhiqiang Wu, Xiaoli Yu, Xiaomao Guo, Saiquan Lu, Kaixuan Li e Gongyi Yu: *Is it possible for knowledge-based planning to improve intensity modulated radiation therapy plan quality for planners with different planning experiences in left-sided breast cancer patients?* *Radiation Oncology*, 12(87), 2017. 17
- [19] Tol, Jim P., Alexander R. Delaney, Max Dahele, Ben J. Slotman e Wilko F.A.R. Verbakel: *Evaluation of a knowledge-based planning solution for head and neck cancer*. *International Journal of Radiation Oncology*, 91(3):612–620, 2015. 17
- [20] Delaney, Alexander R., Jim P. Tol, Max Dahele, Johan Johan Cuijpers, Ben J. Slotman e Wilko F.A.R. Verbakel: *Effect of dosimetric outliers on the performance of a commercial knowledge-based planning solution*. *International Journal of Radiation Oncology*, 94(3):469–477, 2016. 17

- [21] Amit, Guy G., Thomas G. Purdie, Alex Levinshtein, Andrew J. Hope, Patricia Lindsay, Andrea Marshal, David A. Jaffray e Vladimir Pekar: *Automatic learning-based beam angle selection for thoracic imrt*. *Medical Physics*, 42(4):1992–2005, 2015. 18
- [22] Varian: *Eclipse Photon and Electron Algorithms 15.5 Reference Guide*, volume 1. Varian Medical Systems, 2017. 27
- [23] Law, Maria Y. Y. e Brent Liu: *Dicom-rt and its utilization in radiation therapy*. *Informatics in Radiology*, 29:655–667, 2009. 29
- [24] Pyyry, Keranen: *Varian APIs, A handbook for programming in the Varian oncology software ecosystem*, volume 1. VARIAN, 2018. 31