

Universidade Estadual de Campinas
Instituto de Matemática, Estatística e Computação Científica
Departamento de Matemática Aplicada

Aplicação da Teoria de Conjuntos *Fuzzy* na predição
do estadiamento patológico do Câncer de Próstata

Autora: Graciele Paraguaia Silveira

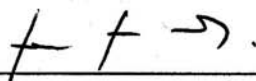
Orientador: Prof. Dr. Laércio Luis Vendite

Campinas - SP
Abril/2007

Aplicação da Teoria de Conjuntos *Fuzzy* na predição do estadiamento patológico do Câncer de Próstata

Este exemplar corresponde a redação final da dissertação devidamente corrigida e defendida por **Graciele Paraguaia Silveira** e avaliada pela comissão julgadora.

Campinas, 09 de abril de 2007.



Prof. Dr. Laércio Luis Vendite
Orientador.

Banca examinadora:

Prof. Dr. Laércio Luis Vendite (IMECC/UNICAMP)

Prof. Dr. Laécio Carvalho de Barros (IMECC/UNICAMP)

Prof. Dr. Ubirajara Ferreira (FCM/UNICAMP)

Dissertação apresentada ao Instituto de Matemática, Estatística e Computação Científica, UNICAMP, como requisito parcial para a obtenção do Título de MESTRE em Matemática Aplicada.

**FICHA CATALOGRÁFICA ELABORADA PELA
BIBLIOTECA DO IMECC DA UNICAMP**

Bibliotecária: Miriam Cristina Alves – CRB8a / 5094

Silveira, Graciele Paraguaia

Si39a Aplicação da teoria de conjuntos fuzzy na predição do
estadiamento patológico do câncer de próstata / Laércio Luis Vendite --
Campinas, [S.P. :s.n.], 2007.

Orientador : Laércio Luis Vendite

Dissertação (mestrado) - Universidade Estadual de Campinas,
Instituto de Matemática, Estatística e Computação Científica.

1. Modelamento matemático. 2. Conjuntos difusos.
3. Biomatemática. 4. Câncer – Modelos matemáticos. I. Vendite,
Laércio Luis. II. Universidade Estadual de Campinas. Instituto de
Matemática, Estatística e Computação Científica. III. Título.

Título em inglês: Application of the fuzzy sets theory in the prediction of the patologic stage of the prostate cancer.

Palavras-chave em inglês (Keywords): 1. Mathematical model. 2. Fuzzy sets. 3. Biomathematics. 4. Cancer – Mathematics models.

Área de concentração: Biomatemática

Titulação: Mestre em Matemática Aplicada

Banca examinadora: Prof. Dr. Laércio Luis Vendite (IMECC-Unicamp)
Prof. Dr. Laécio Carvalho de Barros (IMECC-Unicamp)
Prof. Dr. Ubirajara Ferreira (FCM-Unicamp)

Data da defesa: 09/04/2007

Programa de Pós-Graduação: Mestrado em Matemática Aplicada

Dissertação de Mestrado defendida em 09 de abril de 2007 e aprovada

Pela Banca Examinadora composta pelos Profs. Drs.



Prof. (a). Dr (a). LAÉRCIO LUIS VENDITE



Prof. (a). Dr (a). UBIRAJARA FERREIRA



Prof. (a). Dr (a). LAÉRCIO CARVALHO DE BARROS

“A procura da verdade é difícil e é fácil, já que ninguém poderá desvendá-la por completo ou ignorá-la inteiramente. Contudo, cada um de nós poderá acrescentar um pouco do nosso conhecimento sobre a natureza e, disto, uma certa grandeza emergirá”.

Aristóteles, 350 AC.

Agradecimentos

Ao professor Laércio Luis Vendite, pela orientação e por todo o auxílio prestado.

Ao Dr. Ubirajara Ferreira, pela disposição e pelo empenho em nos ajudar a levar adiante este trabalho, fornecendo-nos os dados de pacientes, que nos foram tão úteis, e auxiliando-nos na análise dos resultados que obtivemos.

Ao professor Laécio Carvalho de Barros, por todos os esclarecimentos que me foram prestados acerca do conteúdo do trabalho.

Aos meus pais Ione Luiza e Jesus Luis, pela educação que me proporcionaram e por todos os anos de estudos em que me sustentaram e me incentivaram.

Aos meus irmãos Fernando Luis e Leidiane, pela força e incentivo.

Ao Raphael por todo o seu amor, atenção e dedicação durante esses três anos que nos conhecemos e juntos estamos.

A minha grande e querida amiga Flávia, pela nossa amizade desde a infância, pelos anos juntas no curso de graduação na UFG/Jataí-GO e por todo o apoio que sempre me foi dado.

E, por fim, a Fernanda Rocha, Luciana Elias, Tacyanne Assis, Denise Bernasoli e Luciana Nakayama, pela amizade e por de alguma forma terem me incentivado a estar aqui hoje.

Resumo

O aumento da incidência de casos de câncer de próstata, nos últimos anos, é um importante problema de saúde pública e um desafio para a ciência médica. Nesta dissertação estudamos a construção de um modelo matemático, que foi desenvolvido para prever o estadiamento patológico do câncer de próstata. A intenção é auxiliar o especialista no processo de tomada de decisão, com relação ao estágio da doença. O modelo consiste num sistema baseado em regras *fuzzy*, que combina os dados pré-cirúrgicos - estado clínico, nível de PSA e grau de Gleason - valendo-se de um conjunto de regras, de natureza lingüística, elaborado a partir das informações presentes nos nomogramas já existentes. Com isso esperava-se obter, na saída do sistema, a chance de o indivíduo, com determinado quadro clínico, estar em cada estágio de extensão do tumor: localizado, localmente avançado e metastático. Foram feitas simulações, com dados de pacientes do Hospital das Clínicas/UNICAMP. Os resultados obtidos foram comparados com as probabilidades de Kattan *et al*, que embora sejam utilizadas nas decisões médicas, são consideradas pessimistas, em relação ao estágio da doença. Com o objetivo de aproximar os resultados, da realidade vivida pelos pacientes, efetuamos algumas modificações na modelagem. Tais mudanças foram suficientes para deixar os resultados mais otimistas e, portanto, mais realísticos.

Abstract

The increase of the incidence of prostate cancer is a important problem of public health and a challenge to medical science. In this dissertation, we studied the construction of a mathematical model wich it was developed to predict the pathologic stage of prostate cancer. The intention is to help specialist on the decision process about stage of the disease. The model consists on a system founded in *fuzzy* laws that it combine the pre-surgicals dates - clinic state, PSA levels and Gleason score - availing of a linguistic laws set made with base on informations of the existents nomograms. Herewith we were hoping to ger person's chance, with clinics characteristics determinates, is in each stage of tumor extension: localized, advanced locally and metastatic. Simulations were made with patient's dates of the Clinics Hospital/UNICAMP. The results were compared with Kattan's probabilities that are used on the medicals decisions. However this probabilities to the disease are considered pessimists. With the aim of approach the results and the reality lived by patients, we did some modifications on the model. This changes were enough to became the results more otimists and therefore more realistics.

Sumário

Introdução	1
1 Câncer - Características do Fenômeno	5
1.1 Introdução	5
1.2 Câncer	6
1.2.1 Marcadores Tumoraes	10
1.3 Câncer de Próstata	11
1.3.1 Diagnóstico	14
1.3.2 Estadiamento	23
1.4 Considerações Finais	26
2 A Teoria de Conjuntos <i>Fuzzy</i>	29
2.1 Introdução	29
2.2 Conjuntos <i>Fuzzy</i>	30
2.3 Lógica <i>fuzzy</i>	33
2.3.1 Variáveis Lingüísticas	34
2.3.2 Relações <i>Fuzzy</i>	35
2.3.3 Regras <i>Fuzzy</i>	37
2.3.4 Inferência	40
2.4 Sistema Baseado em Regras <i>Fuzzy</i>	47

2.5	Possibilidade e Probabilidade	50
2.6	Considerações Finais	55
3	Modelagem <i>Fuzzy</i> para predizer o estágio do Câncer de Próstata	57
3.1	Introdução	57
3.2	Nomogramas para o Câncer de Próstata	59
3.3	O Modelo <i>Fuzzy</i>	64
3.3.1	Simulações	71
3.3.2	Relação entre Probabilidade e Possibilidade	73
3.4	Considerações Finais	79
4	Buscando um melhor desempenho do Modelo <i>Fuzzy</i>	83
4.1	Introdução	83
4.2	As variáveis de Entrada e Saída do Sistema	84
4.3	A Base de Regras e algumas Simulações	85
4.4	Adaptação das Funções de Pertinência	89
4.5	Novas Simulações	90
4.6	Considerações Finais	93
5	Conclusão e Trabalhos Futuros	95
5.1	Conclusão	95
5.2	Trabalhos Futuros	97
	Referências Bibliográficas	99

Introdução

Nas últimas décadas, a humanidade tem se deparado com problemas cada vez mais complexos - em diversas áreas - que necessitam ser tratados em tempo menor e que exigem respostas também complexas.

O desejo de resolver esses problemas impulsionou o avanço da tecnologia dos computadores, tanto em velocidade de processamento, quanto em capacidade de armazenamento de informações, possibilitando sua utilização na modelagem de sistemas multivariáveis.

A modelagem matemática procura descrever fenômenos do mundo real e suas tendências, no decorrer do tempo. Uma das áreas em que é considerada uma ferramenta de grande importância é a medicina.

Os modelos matemáticos em medicina, são classificados, segundo alguns autores, como modelos estocásticos, isto é, modelos que incorporam elementos probabilísticos; ou modelos determinísticos, nos quais a evolução das variáveis é obtida pela estrutura do modelo, dadas as condições iniciais e os valores dos parâmetros [31].

Um modelo matemático é uma representação ou interpretação simplificada da realidade. Contudo, as propriedades inerentes aos procedimentos médicos, biológicos e epidemiológicos, assim como as características das variáveis estabelecidas, devem ser consideradas.

As doenças geralmente são descritas por meio de termos lingüísticos, que

são intrinsecamente vagos, e que muitas vezes constituem variáveis qualitativas em medicina. O tratamento dessas variáveis, utilizando métodos quantitativos, apresenta dificuldades.

Foi a partir de situações como essa, envolvendo incertezas não-estatísticas, que surgiu a Teoria dos Conjuntos *Fuzzy*, introduzida por Lotfi Asker Zadeh, em 1965. A intenção de Zadeh foi flexibilizar a pertinência de elementos aos conjuntos, criando o conceito de grau de pertinência.

A palavra “*fuzzy*”, de origem inglesa, significa incerto, vago, impreciso, subjetivo, nebuloso, obscuro, indistinto e normalmente não é traduzida para o português [5].

O objetivo da teoria de Conjuntos *Fuzzy* é deixar mais próximas, a precisão da matemática clássica e as imprecisões do mundo real. Quando se utiliza essa teoria, a incerteza inerente aos dados e à opinião do especialista podem ser incorporados aos modelos matemáticos [11]. Surgem então, os modelos matemáticos baseados na teoria de Conjuntos *Fuzzy*. Esses modelos vêm sendo utilizados em medicina, desde a década de 70, em várias aplicações, incluindo diagnósticos.

O aumento na incidência de casos de câncer, em países desenvolvidos e em desenvolvimento, nos últimos anos, é um importante problema de saúde pública, que tem preocupado a classe médica.

O processo global de industrialização, ocorrido principalmente no século passado, desencadeou a redefinição de padrões de vida, condições de trabalho, nutrição e consumo [25].

No Brasil, como nos demais países, este processo determinou significativas alterações na demografia, reduzindo taxas de natalidade e mortalidade, e aumentando a expectativa de vida da população. Com isso, a exposição a um grande número de fatores de risco ambientais e a fatores relacionados às

disparidades sociais, também cresceram.

O câncer de próstata é um dos tipos de câncer com maior ocorrência no país e foi escolhido como tema deste trabalho, juntamente com a Teoria de Conjuntos *Fuzzy*, que acreditamos ser apropriada para lidar com as incertezas e subjetividades das informações.

Um dos objetivos do trabalho é fazer uma leitura do modelo matemático *fuzzy*, para prever o estágio patológico do câncer de próstata - construído por Castanho (2005) [11] - descrevendo modificações e atualizações, comentando os resultados obtidos e avaliando o desempenho do modelo.

O segundo objetivo é discutir outras possibilidades de mudanças, nessa modelagem, a fim de obter uma melhor performance do modelo, de acordo com a avaliação de um especialista.

O Capítulo 1 apresenta informações gerais sobre o câncer, e específicas sobre o câncer de próstata, tais como diagnóstico, estadiamento e terapias que podem ser adotadas.

No Capítulo 2 são introduzidos alguns conceitos importantes, relacionados à teoria de Conjuntos *Fuzzy*. As definições e algumas propriedades das distribuições de probabilidade e possibilidade, que serão úteis na modelagem, também são apresentadas.

O Capítulo 3 foi destinado à construção do sistema baseado em regras *fuzzy*, para prever o estadiamento patológico do câncer de próstata. Como já mencionamos, fizemos atualizações e algumas modificações, na modelagem feita por Castanho (2005). Além disso, os resultados de simulações e a comparação com as probabilidades existentes nos nomogramas serão apresentados.

No Capítulo 4, são discutidas novas mudanças que implementamos, na construção do sistema, visando aproximar um pouco mais os resultados das

simulações, da realidade clínica dos pacientes com câncer de próstata.

Finalmente, no Capítulo 5, apresentamos a conclusão da dissertação e algumas sugestões de trabalhos futuros.

Capítulo 1

Câncer - Características do Fenômeno

1.1 Introdução

Os casos de câncer têm aumentado de maneira considerável, principalmente a partir do século passado, configurando-se, na atualidade, como um dos mais importantes problemas de saúde pública mundial.

Segundo a *World Health Organization*, são diagnosticados, anualmente, mais de 11 milhões de novos casos de câncer no mundo. A estimativa é de que 16 milhões de novos casos surjam em 2020 [51].

No Brasil, o Instituto Nacional do Câncer - INCA - é o órgão do Ministério da Saúde responsável pelas ações nacionais integradas, voltadas para a prevenção e controle da neoplasia.

Em maio de 2005, o INCA inaugurou o Banco Nacional de Tumores e DNA, cuja função será reunir as informações necessárias para a elaboração do perfil genético da população brasileira, possibilitando estudos voltados ao aprimoramento do diagnóstico e do tratamento do câncer.

Entre os brasileiros, o câncer é a segunda causa de mortes, superado apenas pelas doenças do sistema circulatório, como diabetes, infarto, hipertensão, entre outros [27].

As estimativas do INCA, para 2006, mostraram que cerca de 472 mil novos casos de câncer seriam diagnosticados, sendo 234 mil entre os homens e 238 mil entre as mulheres. O câncer de próstata apresentou a segunda taxa mais elevada, perdendo apenas para o câncer de pele não melanoma. A ocorrência estimada foi de 47.280 novos casos, o que corresponde a um risco de 51 casos a cada 100 mil homens.

1.2 Câncer

Câncer é o nome dado a uma classe de doenças caracterizadas pelo crescimento descontrolado de células anormais do organismo. As células normais do corpo vivem, se dividem e morrem de forma controlada. As células cancerosas são diferentes, não obedecem a esses controles e se dividem sem parar. Além disso, não morrem como as células normais e continuam a produzir mais células anormais (Sasse, 2007) [43].

A atividade de cada célula ou tecido é dirigida pelo seu DNA. No decorrer da vida, o DNA sofre alterações denominadas mutações, causadas por erros que ocorrem durante sua duplicação, necessária à divisão celular. Na maioria das vezes, o próprio DNA detecta e conserta seus erros. Entretanto, o mecanismo de reparo não funciona para algumas mutações do DNA, que ocorrem em genes específicos e, portanto, podem provocar o crescimento desordenado das células, constituindo a origem de tumores (acúmulo de células cancerosas) ou *neoplasias* malignas [26].

As seguintes características são comuns a toda célula tumoral:

- perda do controle do crescimento celular e de suas funções no tecido saudável;
- infiltração nos tecidos adjacentes;
- formação de metástases (propagação irregular pelo organismo), por via hematogênica, linfática ou superfície serosa.

De acordo com o INCA [27], a extensa maioria dos casos de câncer (80% a 90%) está, de alguma forma, associada a fatores ambientais. Entende-se por ambiente o meio em geral, o ambiente de consumo (alimentos e medicamentos), o ambiente ocupacional (industrialização - agentes químicos, físicos e biológicos) e o ambiente sócio-cultural (estilo e hábitos de vida). Os casos de câncer que se devem, exclusivamente, a fatores hereditários, familiares e étnicos são raros.

Em suas etapas iniciais, o crescimento do câncer é relativamente silencioso e os poucos sintomas que apresentam assemelham-se aos produzidos por enfermidades benignas ou até mesmo banais. Somente quando o tumor se expande é que surgem manifestações que permitem o diagnóstico.

Segundo Fair, Heston e Cordon-Cardo (1997) [21], estima-se que 1g de tumor contenha aproximadamente 1 bilhão (10^9) de células e que seu volume seja de 1cm^3 . Supondo que o tumor se desenvolve a partir de uma única célula, 30 duplicações implicam em um volume de 1cm^3 . Uma massa de 1Kg, que é o tamanho máximo que um tumor pode alcançar, é o resultado de 40 duplicações (Ver tabela 1.1).

É importante observar que a progressão no tamanho do tumor, não depende somente das divisões celulares, também inclui sua habilidade em acessar a vasculatura (garantir suplemento sanguíneo) e se disseminar para locais distantes.

Duplicações	Células	Gramas de tumor	Deteção
0	1	Nanograma	—
10	10^3	Micrograma	—
20	10^6	Miligrama	Microscópica
30	10^9	Gramma	Palpável
40	10^{12}	Quilograma	Debilitação
			Morte

Tabela 1.1: Relação entre n° de duplicações, n° de células e massa do tumor.

Diferentes tipos de câncer têm comportamentos diferentes e exigem tratamentos distintos, até mesmo quando se trata de câncer do mesmo órgão [26]. O primeiro objetivo é remover todas as células cancerosas, de modo que as funções do tecido normal não sejam danificadas. Caso permaneça alguma, um novo tumor pode se desenvolver fazendo com que a doença reapareça.

A seguir descrevemos as modalidades terapêuticas destinadas à pacientes com câncer:

- Cirurgia: é um tratamento local e o mais definitivo, quando o tumor é localizado. A finalidade é remover todo o tecido doente e o mínimo possível de tecido saudável. No entanto, para alguns tipos de câncer apenas a cirurgia não é suficiente, devido à disseminação de células cancerosas local ou difusamente. É considerada radical quando a lesão inteira é retirada, com margens de segurança adequadas e com intenção de cura. Em algumas situações, para manter a função de um determinado órgão ou para prevenir complicações, deve-se optar por uma cirurgia menos radical. Nesses casos, outras terapias como a radio ou a quimioterapia podem ser associadas.

- Radioterapia: consiste na aplicação de diferentes tipos de raios-X, com a finalidade de destruir as células doentes. É utilizada em cânceres localizados, tanto para reduzir um tumor antes da cirurgia, como para eliminar eventuais células cancerosas após a cirurgia. Atualmente, a radioterapia é uma técnica de alta precisão, que atinge de forma eficaz as células doentes e afeta minimamente órgãos e tecidos normais [26].
- Quimioterapia: consiste na utilização de medicamentos que têm ação citotóxica (causam danos às células), administrados por via oral ou injetados. Diferente da radioterapia, os medicamentos usados na quimioterapia podem atingir qualquer parte do corpo, através da corrente sanguínea. Pode ser adotada como tratamento principal (leucemias, câncer de testículo, linfomas), mas normalmente é adjuvante, após cirurgia ou radioterapia. Em doenças avançadas pode ser paliativo.
- Hormonoterapia: trata-se da utilização de hormônios que possuem influência restritiva no crescimento de alguns tumores. Essa terapia pode ser adotada como tratamento exclusivo, mas geralmente é associada às demais.
- Imunoterapia: esse tratamento baseia-se na ativação do sistema imune, contra as células malignas, já que há fortes indícios de que o nosso sistema de defesa tem papel importante na resposta ao câncer. Pode ser utilizada de forma combinada com a radio e a quimioterapia.
- Braquiterapia (câncer de próstata): consiste no implante de sementes de iodo 125, diretamente na próstata. A aplicação é feita através de agulhas (varia entre 14 a 16 delas) colocadas no períneo. Este material libera uma radiação, dentro do órgão, capaz de matar as células tumorais. Como a aplicação ocorre diretamente no local afetado, os

resultados são mais precisos (não atinge órgãos e tecidos sadios). As semente são revestidas de platina e, após certo tempo, perdem a função no organismo, não causando problemas futuros. No caso de tumores avançados, a substância é o irídio 192, aplicada através de agulhas, no períneo, mas sem que o material fique alojado no local [13].

As terapias indicadas para o tratamento de câncer, frequentemente têm efeitos adversos e, algumas vezes, falham. Segundo Sasse (2007) [43], se a cura não é possível, o objetivo passa a ser a remissão da doença, deixando o paciente longe de seus efeitos e de hospitalizações. Caso sejam remotas as chances de remissão, a opção pelo tratamento deve priorizar a melhoria da qualidade de vida do paciente, e não apenas o prolongamento de uma vida sofrida.

Castanho (2005) [11], ressalta que definir a cura do câncer é muito difícil, pois não existem métodos para comprovar que a última célula cancerosa foi eliminada do organismo. Estatisticamente, cura é o estado no qual o risco da doença é o mesmo, tanto para um grupo de pessoas que tiveram câncer, quanto para os que não tiveram. O sucesso da terapia e as chances de sobrevivência dependem do tipo de câncer, de quão precoce ele é descoberto e da escolha das modalidades de tratamento que mais se adequam ao paciente.

1.2.1 Marcadores Tumorais

Os marcadores tumorais são substâncias produzidas pelos tumores, ou pelos organismos em resposta à presença tumoral, com expressão ou quantificação diferencial no sangue, urina ou tecidos de pacientes com neoplasia.

De acordo com Souto e colaboradores (2004) [47], o marcador tumoral ideal deve possuir os seguintes atributos:

- Alta sensibilidade - ser detectado precocemente, quando poucas células cancerosas estiverem presentes;
- Alta especificidade - não ser detectado em indivíduos saudáveis ou em doenças benignas;
- Ser órgão-específico, ou seja, específico para um determinado tipo de câncer;
- Dosagem fácil e baixo custo;
- Utilidade no diagnóstico, estadiamento e monitorização do tratamento.

Nas últimas décadas, vários marcadores tumorais foram descritos e nenhum deles possui todos os atributos mencionados acima.

Segundo Pradal e colaboradores (1999) [41], embora a aplicabilidade e interpretação corretas, dos marcadores tumorais, ainda não tenham atingido domínio amplo, eles muito contribuíram para a compreensão e melhor controle do processo neoplásico.

1.3 Câncer de Próstata

A próstata é uma glândula do sistema urogenital masculino, que se localiza logo abaixo da bexiga e é atravessada pela uretra (ver figura 1.1). É responsável pela produção de parte do líquido seminal, que protege e nutre os espermatozóides. Também produz o PSA (Antígeno Prostático Específico), cuja função biológica é a lise do coágulo seminal, auxiliando na reprodução humana.

O câncer de próstata é a neoplasia mais freqüente diagnosticada em homens e uma das principais causas de morte, por doença maligna, no Brasil

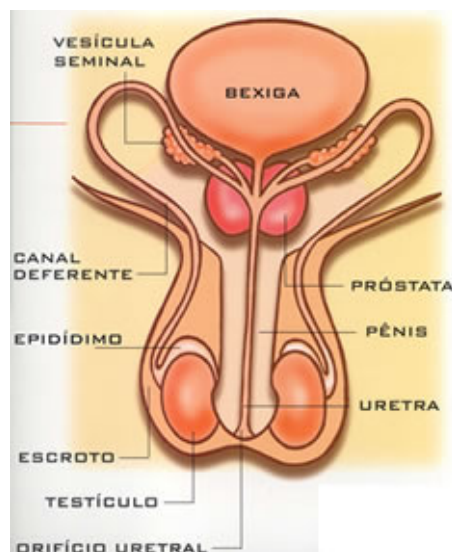


Figura 1.1: Ilustração do sistema reprodutor masculino.

(Ferreira e Nardi, 1999) [22]. A origem do câncer de próstata ainda é desconhecida; entretanto, presume-se que alguns fatores possam influenciar o seu desenvolvimento, a saber:

- Idade: é um tumor raro antes dos 50 anos e o risco aumenta progressivamente com a idade;
- Fator genético: a possibilidade de aparecimento da doença é 2 vezes maior, quando o pai ou irmão já foi acometido pelo problema. No caso de 2 ou 3 parentes de 1º grau, os riscos aumentam 5 e 11 vezes, respectivamente [22];
- Raça: há uma variação em diferentes grupos étnicos. O câncer de próstata é extremamente raro em asiáticos e tem sua prevalência aumentada entre negros americanos [33];
- Dieta:
 - levantamentos epidemiológicos revelam diferenças no consumo de gor-

dura animal, nas áreas de alto-risco (países escandinavos) e baixo-risco (extremo oriente);

- consumo de *licopeno*, antioxidante presente no tomate e derivados, diminui as chances de surgir um câncer de próstata [24];
- outros agentes importantes, no combate ao tumor, são a vitamina A, vitamina E, selênio e alguns cereais e grãos - como a soja, ervilha, lentilha - que têm um componente chamado *isoflavona*. Tais substâncias teriam efeito protetor contra a progressão do câncer [24];
- consumo de vegetais ricos em *isotiocianatos* (principalmente brócolis e couve-flor), diminui o risco de neoplasia na próstata. Essas substâncias ativam enzimas capazes de detoxificar carcinógenos [17].

Outros potenciais fatores, que podem oferecer risco são citados por Ferreira e Nardi (1999) [22], como a poluição do ar e a exposição a altos níveis de cádmio e zinco (antagonistas do selênio); Pienta (1997) [40] cita o metabolismo hormonal alterado e a vasectomia; a exposição a pesticidas, principalmente os compostos por *metilbrometo*, é apontada por Alavanja e colaboradores (2003) [4].

Embora a próstata seja constituída por vários tipos de células, a maioria dos cânceres de próstata tem origem nas células das glândulas que produzem o líquido seminal (porções mais periféricas e posterior). Eles são chamados de adenocarcinomas. O restante compreende casos de sarcomas, carcinomas epidermóides e de células transicionais.

Segundo Ferreira e Nardi (1999), a história natural do câncer de próstata é variável e imprevisível. Alguns tumores apresentam potencial maligno tão pronunciado, que as metástases surgem antes que ocorram sinais locais ou sintomas sistêmicos. Em outros casos, permanecem localizados na glândula - em estado latente - durante toda a vida do paciente, sem nenhuma

repercussão clínica.

O câncer de próstata pode ser classificado como se segue:

- confinado ou localizado na glândula: se todo o câncer estiver confinado na glândula prostática;
- localmente avançado: se o câncer se espalhou para fora da próstata;
- metastático: se o tumor se espalhou pelo organismo. A doença atinge por extensão direta os órgãos contíguos como a bexiga, as vesículas seminais e a parede pélvica ou mesmo os linfonodos regionais pélvicos. Outras vezes adquire a capacidade de se disseminar para órgãos distantes, acometendo principalmente os ossos, pulmão e fígado.

Quando o estágio da doença é confinado, as opções de tratamento incluem a cirurgia (prostatectomia radical), radioterapia externa, braquiterapia, terapia hormonal ou a combinação deles. Para o câncer que já produziu metástase, o tratamento é paliativo, ou seja, visa diminuir a progressão do tumor e aliviar os sintomas do paciente. Nesse caso, o tratamento inclui a hormonoterapia e quimioterapia.

1.3.1 Diagnóstico

Na fase inicial é comum o câncer de próstata não apresentar sintomas. Na doença avançada, os sintomas podem manifestar-se pela obstrução do fluxo urinário, com conseqüente piora da função renal. De acordo com Ferreira e Nardi (1999) [22], ainda na fase avançada, podem ocorrer dores ósseas decorrentes de metástases.

Para o diagnóstico precoce, do carcinoma da próstata, são indicados o exame clínico (toque retal) e a dosagem do Antígeno Prostático Específico

(PSA). Os resultados podem sugerir a existência da doença e indicarem a realização da ultra-sonografia transretal. Esta ultra-sonografia, por sua vez, poderá mostrar a necessidade de se realizar a biópsia prostática transretal, que fornecerá o diagnóstico definitivo [27]. Outros exames como a tomografia computadorizada e a cintilografia óssea podem ser utilizados para evidenciar o avanço da doença.

Toque Retal

O toque retal é realizado, rotineiramente, para avaliar a extensão local do carcinoma da próstata. Na fase inicial, pode revelar a presença de um nódulo endurecido na glândula. Entretanto, é importante observar que algumas lesões benignas, como nódulos da hiperplasia benigna, cálculos prostáticos e áreas de infarto, podem produzir áreas endurecidas e simular o câncer.

Por outro lado, o toque normal não exclui a presença de câncer. Ferreira e Nardi (1999), ressaltam que boa parte dos cânceres (cerca de 20%) se origina nas porções mais centrais da glândula, inacessíveis ao exame digital do reto.

Nos casos avançados, a glândula apresenta consistência dura, algumas vezes pétrea, podendo-se notar endurecimento das vesículas seminais.

Em síntese, sua sensibilidade diagnóstica varia de 40% a 80%, dependendo do tipo de câncer que está sendo avaliado. Falha em 23% a 45% das neoplasias detectadas, em biópsias feitas, porque estavam elevados os níveis de PSA (Srougi, 1999) [48]. Além disso, em mais de 50% dos pacientes, os cânceres detectados pelo toque retal estão patologicamente avançados e, dessa forma, o tratamento tem poucas chances de sucesso (Carter e Partin, 1997) [10].

Marcadores Tumoriais em Câncer de Próstata

I. Fosfatase Ácida Prostática - FAP

A fostatase ácida foi o primeiro marcador tumoral utilizado para o estadiamento clínico, dos pacientes com câncer de próstata. FAP é uma fosfomonoesterase não-específica, pois não é encontrada exclusivamente na próstata. É sintetizada pelas células epiteliais prostáticas, sob controle androgênico. Segundo Souto e colaboradores (2004) [47], sua utilidade é limitada, pois mesmo após a remoção cirúrgica da próstata, níveis séricos do marcador podem ser detectados, e uma série de doenças renais, ósseas e esqueléticas pode alterá-los.

Valores anormais da FAP e valores na metade superior da faixa normal sugerem acometimento ósseo em mais de 80% dos casos. No entanto, níveis normais não implicam em ausência da doença extra-prostática.

II. Antígeno Prostático Específico - PSA

O Antígeno Prostático Específico (PSA) é uma glicoproteína presente no epitélio prostático e é responsável pela liquefação do coágulo do esperma, após a ejaculação.

Para Souto e colaboradores (2004), o PSA é o que mais se aproxima daquele que seria um marcador ideal, comparativamente a todos os marcadores hoje disponíveis em oncologia. Está em uso clínico para diagnóstico e seguimento do câncer prostático desde 1986.

O PSA é “próstata - específico”, ou seja é produzido principalmente pelo tecido prostático, mas não é um teste “câncer - específico”. Isso significa que sua presença em níveis séricos elevados nem sempre é indicativo de câncer de próstata. Sabe-se que as concentrações de PSA podem alterar-se em diversas patologias prostáticas e em situações distintas, como mostra a tabela 1.2.

Segundo Souto e colaboradores (2004), acredita-se que o PSA seja liberado para a corrente sangüínea, principalmente quando há alguma ruptura

Causas de Alterações do PSA	
Comprovadas	Discutidas
Câncer de Próstata	Toque Retal
Hiperplasia Benigna	Ejaculação
Prostatites	Exercícios (ciclismo)
Isquemias, infartos prostáticos	Neoplasia prostática intraepitelial (PIN)
Manipulação - biópsias prostáticas, ressecções, cistoscopias	—
Medicações - antiandrogênicos	—

Tabela 1.2: Possíveis causas de alterações dos níveis séricos de PSA [47].

nos mecanismos que o mantêm no tecido prostático, como isquemias, infartos, processos inflamatórios e câncer. A hiperplasia benigna da próstata (HBP) eleva os níveis de PSA sérico devido ao aumento do número de células produtoras de PSA. Cada grama de tecido com hiperplasia benigna eleva o PSA sérico em 0,3 ng/ml, ao passo que cada grama de tecido com câncer elevaria o mesmo em 3,0 ng/ml.

Na classificação “clássica”, presente na literatura, os valores normais de PSA variam de 0 a 4 ng/ml, níveis entre 4 e 10 ng/ml são considerados levemente elevados, entre 10 e 20 ng/ml são moderadamente elevados e acima de 20 ng/ml são altamente elevados.

Ferreira e Nardi (1999) [22] afirmam que, como regra geral, deve-se considerar uma margem de erro em torno de 20%, tanto para casos de pacientes portadores de câncer, cujo PSA encontra-se normal, quanto nos casos de HBP cujo PSA apresenta-se elevado. Entretanto, resultados acima de 10 ng/ml são altamente sugestivos de neoplasia e devem ser sempre investigados.

O nível de PSA correlaciona-se com a extensão do tumor. Quanto maior

o índice de PSA encontrado, maiores são as possibilidades de doença extra-prostática. Porém, somente o nível de PSA não é suficiente para determinar o estágio do câncer, pois existem pacientes com mesmo PSA em estágios distintos.

Estudos apontam que cerca de 80% dos pacientes com câncer de próstata apresentam, ao diagnóstico, nível de PSA acima de 4 ng/ml. No entanto, Catalona, Smith e Ornstein (1997) [12] mostraram, em suas pesquisas, que 22% dos homens com câncer de próstata têm PSA sérico entre 2,6 e 4 ng/ml. Zhu e colaboradores (2005) [52] também mostraram que cânceres de próstata clinicamente significantes, no estágio localizado, foram detectados com frequência, em homens que apresentavam PSA entre 2,6 e 4 ng/ml e fizeram biópsia.

De acordo com Dini e Koff (2006) [18], o número de biópsias da próstata tem aumentado devido ao rastreamento da doença com PSA, que é um marcador que apresenta baixa especificidade. Vários pesquisadores preconizam um ponto de corte menor (2,5 ng/ml), para a indicação de biópsia pelo PSA, com o objetivo de detectar um número adicional de casos potencialmente curáveis, mas certamente elevando o número de biópsias desnecessárias.

Quando o PSA situa-se entre 4 e 10 ng/ml, obter uma sensibilidade aceitável significa perder especificidade, isto é, para manter taxas aceitáveis de diagnóstico, nessa faixa, muitos pacientes são submetidos a exames desnecessários. Segundo Dini e Koff (2006), assume-se que cerca de 11% a 39% dos pacientes, com PSA entre 4 e 10 ng/ml, apresentam câncer à biópsia transretal.

Para otimizar a eficácia da medição do PSA, como teste diagnóstico, outros conceitos foram propostos, sempre visando incrementar a especificidade do teste e evitar biópsias desnecessárias.

1. Velocidade do PSA: é a taxa de aumento do PSA com o tempo. Estudos revelaram que a velocidade de elevação do PSA difere, consideravelmente, entre grupos de pacientes com HBP, câncer e grupos de controle, sugerindo que taxas maiores do que 0,75 ng/ml ao ano representam alto risco de câncer de próstata.

2. Densidades do PSA: no cálculo da densidade, o PSA sérico (*ng/ml*) é dividido pelo volume prostático (*cm³*), determinado por ultra-sonografia transretal. O ponto de corte sugerido tem sido 0,15, com valores superiores indicando presença de câncer. Críticas sobre a utilidade desse teste vêm sendo feitas, devido à variabilidade na avaliação do volume prostático e às diferenças de PSAs com mesmo volume glandular, em casos de HBP [47].

3. PSA esperado: é o quociente do volume prostático (avaliado pela ultra-sonografia) por 12 - ou didaticamente por 10. É considerado normal uma glândula com 90g produzir 3 ng/ml de PSA, porém uma glândula com 30g não pode produzir 9 ng/ml. Nesse último caso, seria indicada a realização da biópsia.

4. PSA ajustado à idade: a idade é um fator importante na elevação dos níveis de PSA. O uso do PSA ajustado à idade visa aumentar a sensibilidade da detecção do câncer nas camadas mais jovens da população e aumentar a especificidade nas mais idosas. No entanto, segundo Souto e colaboradores (2004) [47], a maioria dos autores concorda que a redução da sensibilidade nas faixas etárias acima de 60 anos é problemática, e recomenda a manutenção do uso do limite de 4 ng/ml entre os mais idosos.

Valores máximos diferentes do PSA (ng/ml), a saber 0-2,5; 0-3,5; 0-4,5 e 0-6,5, foram propostos para as categorias etárias 40-49; 50-59; 60-69 e 70-79 anos, respectivamente [35]. Entretanto, dados recentes indicam que esses valores de referência devem ser revisados, pois possuem sensibilidade limitada

e levam a uma excessiva proporção de resultados falso-positivos.

Chun e colaboradores (2006) [15], em um estudo com pacientes canadenses, concluíram que o nível de PSA médio, na maior parte das faixas etárias, não atingiu 1 ng/ml. Os valores de PSA encontrados foram 0,70, 0,90, 0,97 e 1,47 ng/ml, referentes às categorias etárias 40-49, 50-59, 60-69 e 70-79 anos, respectivamente. Esses autores apontam a falta de variabilidade racial como limitação de sua pesquisa e ressaltam que esse fato dificulta sua aplicabilidade em outras populações.

Contudo, assim como Dini e Koff (2006) [18], esses pesquisadores reafirmam a necessidade de cada vez mais se procurar caracterizar o perfil da doença em cada região, tendo em vista a variabilidade das populações no Brasil e no mundo.

5. PSA livre: trata-se de uma dentre as três formas moleculares principais do PSA, no plasma. Vários estudos demonstraram que a proporção de PSA livre, em relação ao PSA total, é inferior nos pacientes com câncer. As pesquisas também estabeleceram que a porcentagem de PSA livre é mais predictiva, do diagnóstico de câncer, do que os níveis de PSA total em pacientes com PSA total entre 4 e 10 ng/ml [47]. Não se sabe o ponto de corte ideal - foram sugeridos valores entre 15% e 25%.

De modo geral, o teste de PSA tem grande utilidade no acompanhamento dos tratamentos instituídos para o câncer de próstata. Se o nível de PSA no sangue retornar ao normal, após o tratamento, então há um indicativo de completa remoção do tumor. Uma queda inicial seguida de nova elevação indica tumor residual ou metástase e um aumento nos níveis de PSA, após o período de normalização, indica recorrência.

Segundo Souto e colaboradores (2004), após um mês da prostatectomia radical, os níveis séricos de PSA devem tornar-se indetectáveis. Também

após os tratamentos radioterápicos ou hormonais, o PSA tende a sofrer uma drástica redução e elevações em seu nível denotam progressão. Já os pacientes com doença metastática, que respondem ao tratamento hormonal, apresentam uma redução importante do PSA, que costuma a equivaler a 80% dos valores prévios, após 30 dias do início do bloqueio hormonal.

Ultra - Som Transretal - USTR

A indicação para ultra-som da próstata deve ocorrer após a realização do teste de PSA e do toque retal. Trata-se de um exame que auxilia na identificação de áreas suspeitas, não detectadas ao toque, e é útil para guiar a agulha da biópsia.

Segundo Ferreira e Nardi (1999) [22], o câncer localizado mostra-se como uma área suspeita em até 60% dos casos. No entanto, existem outras causas, não malignas, que explicam as áreas suspeitas ao USTR, como infarto prostático, cistos e áreas de fibrose.

O USTR pode auxiliar o estadiamento local da doença, fornecendo informações como presença de extensão extra-prostática e envolvimento das vesículas seminais.

Biópsia da Próstata

A biópsia prostática é um procedimento cirúrgico, mediante o qual é coletada uma amostra de tecido que será examinada. A biópsia sextante equivale à retirada de seis fragmentos e a biópsia sextante estendida corresponde à retirada de doze fragmentos, no mínimo. Esse número pode ser ampliado, dependendo da situação clínica e da ultra-sonografia do paciente.

A maioria das biópsias são realizadas com o auxílio do ultra-som. A sonda de ultra-som é introduzida no reto, possibilitando a visualização da glândula

prostática. No caso de identificação de alguma anormalidade (lesão), a agulha será posicionada nesta área, a fim de obter uma amostra de tecido. Biópsias múltiplas podem auxiliar na determinação da causa e extensão da doença.

Segundo a Sociedade Brasileira de Urologia [46], as indicações para a biópsia prostática são: presença de nódulos (detectados no toque retal), níveis de PSA elevados (geralmente acima de 4ng/ml) e, mais recentemente, indica-se também em pacientes mais jovens (com idade inferior a 55 anos) e que tenham PSA acima de 2,5 ng/ml, e naqueles em que a densidade de PSA $> 0,15$ e a velocidade anual $> 0,75$ ng/ml.

O grau de anormalidade das células cancerosas, quando comparadas com as células normais, é denominado grau do tumor. O patologista, através da biópsia, distingue tumor benigno de maligno e identifica o grau de diferenciação celular - ou seja, descreve quão similares, em aparência e função, são as células saudáveis e as células neoplásicas.

A graduação histológica do carcinoma da próstata tem correlação com o estadiamento clínico tendo, portanto, valor prognóstico (Billis, 2003) [8]. Os tumores da próstata são bastante heterogêneos, com áreas de maior e menor diferenciação, e por isso o grau histológico é obtido considerando as áreas de maior anaplasia.

Atualmente o sistema de Gleason, que enfatiza a arquitetura glandular, é o mais utilizado e tem recebido grande aceitação devido à alta correlação com o prognóstico e relativa simplicidade [8]. No sistema de Gleason os tumores são classificados em 5 graus, sendo que o grau 1 representa as lesões mais diferenciadas - que têm comportamento menos agressivo - e o grau 5 simboliza as mais indiferenciadas - que apresentam um comportamento mais agressivo. A figura 1.2 mostra a representação desse sistema.

O diagnóstico final na escala de Gleason é dado pela soma dos graus do

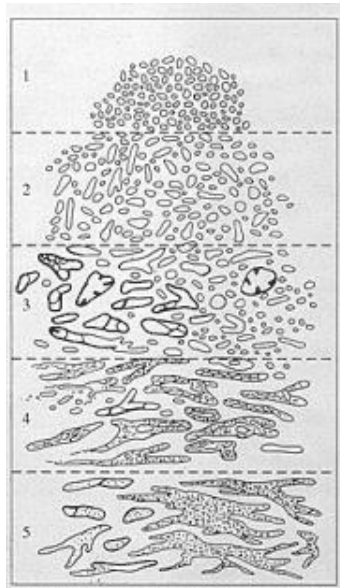


Figura 1.2: Sistema de Gleason para graduação histológica do carcinoma da próstata [34].

primeiro e do segundo padrões histológicos predominantes. Dessa forma, as neoplasias mais diferenciadas são caracterizadas como grau $2 = 1 + 1$ e as mais agressivas são classificadas como grau $10 = 5 + 5$.

De acordo com Billis (2003), em casos de biópsia prostática de agulha, recomenda-se que se existirem mais de dois graus presentes, seja usado como primeiro padrão o mais extenso e como segundo, o pior.

1.3.2 Estadiamento

Após a confirmação histológica do diagnóstico de câncer de próstata, a avaliação da extensão da doença deve ser feita para que o paciente seja encaminhado às terapias apropriadas. A avaliação correta é fundamental, pois os riscos de se submeter aos tratamentos intensivos se justificam, se existirem chances reais de sucesso.

Estadiamento Clínico

O estadiamento clínico consiste na avaliação da extensão da doença, feita a partir dos resultados dos exames do toque retal, marcadores tumorais, grau do tumor e exames de imagem, como ultra-som, cintilografia óssea, radiografia, entre outros.

O sistema TNM (tumor, nodo e metástase), para a classificação de tumores malignos, foi desenvolvido por Pierre Denoix (França), entre os anos de 1943 e 1952. Em 1954, a União Internacional Contra o Câncer (UICC) criou um comitê para prosseguir os estudos nesse campo e estender a técnica geral, de classificação do câncer, para todas as localizações anatômicas. A primeira edição foi publicada em 1958 e, desde então, edições reformuladas e aperfeiçoadas vêm sendo apresentadas [45].

A tabela 1.3 descreve a classificação TNM, presente na sexta edição publicada pelo INCA, em 2004. As regras para classificação e estadiamento correspondem exatamente àquelas que aparecem na sexta edição do Manual para Estadiamento do Câncer, da *American Joint Committee for Cancer Staging (AJCC)*.

Nessa abordagem, o tumor primário (T) foi classificado em quatro estágios, com seus respectivos sub-estágios. A denominação T1 indica que o tumor possui dimensões microscópicas e não pode ser palpado no toque retal. A designação T2 representa os tumores que podem ser tocados, mas parecem estar inteiramente localizados na glândula prostática. O tumor T3 se disseminou além da próstata, atingindo os tecidos circundantes ou as vesículas seminais. O câncer T4 está fixo ou possui extensão ainda maior, atingindo outras estruturas adjacentes. O câncer N já apresentou metástase nos linfonodos pélvicos e M significa metástases em regiões distantes.

Estágio	Descrição
T1a	Não palpável; 5% (ou menos) da amostra é maligna
T1b	Não palpável; mais de 5% da amostra é maligna
T1c	Tumor identificado por biópsia (ex: devido a PSA elevado)
T2a	Palpável; envolve uma metade (ou menos) de um dos lobos
T2b	Palpável; envolve mais da metade de um dos lobos, mas não os dois
T2c	Tumor que envolve ambos os lobos
T3a	Extensão extra-prostática (uni- ou bilateral)
T3b	Tumor que invadiu vesícula(s) seminal (ais)
T4	Tumor fixo ou que invadiu outras estruturas adjacentes
N	Tumor que se espalhou pelos linfonodos regionais
M	Metástase à distância

Tabela 1.3: Estado Clínico do câncer de próstata - Tabela TNM.

A figura 1.3 mostra uma ilustração dos estágios T1 - tumor localizado, T3a - extensão extra-prostática unilateral e N - metástase nos linfonodos regionais.

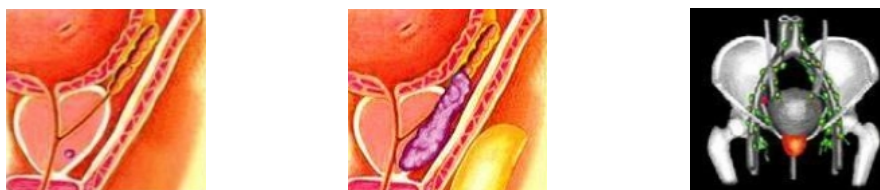


Figura 1.3: Tumor localizado (T1), extensão extra-prostática unilateral (T3a) e metástase nos linfonodos (N).

Estadiamento Patológico

O estadiamento patológico avalia a extensão da doença através do exame dos linfonodos pélvicos e da próstata, após a retirada desses tecidos - por

meio da prostatectomia radical. É mais preciso, comparado ao estadiamento clínico, pois o volume do tumor, a expansão extra-prostática e o envolvimento das vesículas seminais podem ser determinados. Carter e Partin (1997) [10], afirmam que essa é a razão pela qual comparações, referentes à sobrevivência, entre terapias cirúrgicas e não-cirúrgicas continuarão limitadas.

O estágio patológico do câncer de próstata é classificado em:

- órgão-confinado ou localizado: quando a doença está confinada na glândula prostática;
- extra-prostático: quando a doença se espalhou além da próstata, nas não envolveu as vesículas seminais;
- envolvimento das vesículas seminais: quando a doença já atingiu as paredes musculares das vesículas seminais;
- envolvimento dos linfonodos pélvicos.

1.4 Considerações Finais

O câncer de próstata configura um problema de saúde pública. A Sociedade Brasileira de Urologia promove campanhas de esclarecimento e instituiu, desde o ano de 2004, o dia 17 de novembro como o “Dia Nacional de Combate ao Câncer de Próstata”.

Além da difusão de informações, o objetivo da 3ª Campanha Nacional, realizada em 2006, foi cobrar ações governamentais em relação a medicamentos e tempo de espera por cirurgia - para quem recorre ao Sistema Único de Saúde (SUS).

Os índices envolvendo a doença são alarmantes. Mesmo assim, muitos homens ainda relutam em aceitar a realização dos exames, principalmente o

toque retal, por vergonha, medo ou preconceito.

Os médicos alertam que o exame preventivo deve ser feito a partir dos 45 anos de idade, caso não haja diagnóstico anterior na família. Se houver, a recomendação é que o faça a partir dos 40 anos. Este exame é útil na identificação de nódulos ou de áreas suspeitas, como regiões endurecidas da próstata.

As dosagens séricas de PSA podem auxiliar na detecção da doença. No entanto, esse teste ainda não permite a descoberta precoce de todos os casos de câncer. Da mesma forma, as biópsias podem não revelar a presença de tumores na glândula prostática, retardando o diagnóstico.

Avaliar a extensão do tumor é outra dificuldade enfrentada pelos clínicos. A avaliação correta é essencial, pois existe uma grande chance de cura para o câncer de próstata, se este for descoberto precocemente e ainda estiver confinado na glândula.

Por todos esses fatores que citamos, torna-se importante desenvolver um modelo matemático, que combine os dados fornecidos pelos exames, para prever o estágio patológico dos pacientes. A ferramenta escolhida, para a construção desse modelo, foi a teoria dos conjuntos *fuzzy*, devido a sua capacidade de lidar com as incertezas e subjetividades envolvidas no problema.

Capítulo 2

A Teoria de Conjuntos *Fuzzy*

2.1 Introdução

A subjetividade intrínseca de variáveis, embora seja utilizada em nosso cotidiano, transmitida e compreendida linguisticamente entre interlocutores, tem invariavelmente permanecido fora do tratamento matemático tradicional (Barros e Bassanezi, 2001) [6].

No entanto, para modelar fenômenos do mundo real é preciso considerar imprecisões, incertezas, subjetividades e conceitos vagos, estruturas que não são conjuntos no sentido clássico, mas conjuntos em que as propriedades que os definem são incertas, com fronteiras indefinidas, nos quais a transição de pertinência para não - pertinência é gradual.

Assim, com o intuito de estabelecer uma relação entre a precisão da matemática clássica e a imprecisão do mundo real, foi introduzida a teoria dos conjuntos *fuzzy*, em 1965, pelo matemático de origem iraniana Lotfi Asker Zadeh, professor da Universidade de Berkley - Estados Unidos. Seu principal objetivo era proferir um tratamento matemático a certos termos lingüísticos subjetivos, como “aproximadamente”, “em torno de”, entre outros.

Segundo Barros e Bassanezi (2006) [5], esse seria um primeiro passo no sentido de se programar e armazenar conceitos vagos em computadores, tornando possível a produção de cálculos com informações imprecisas, a exemplo do que faz o ser humano.

Esses autores acrescentam ainda que, para obter a formalização matemática de um conjunto *fuzzy*, Zadeh baseou-se no fato de que qualquer conjunto clássico pode ser caracterizado por uma função - sua função característica - cuja definição será dada a seguir, juntamente com alguns conceitos fundamentais da teoria.

2.2 Conjuntos *Fuzzy*

Definição 2.2.1 *Seja $\mathbf{U}$ um conjunto e A um subconjunto de $\mathbf{U}$. A função característica de A é dada por*

$$\chi_A(x) = \begin{cases} 1 & \text{se } x \in A \\ 0 & \text{se } x \notin A. \end{cases}$$

Desse modo, χ_A é uma função cujo domínio é $\mathbf{U}$ e a imagem está contida no conjunto $\{0, 1\}$, onde $\chi_A = 1$ indica que x é um elemento de A e $\chi_A = 0$ indica que x não está em A .

Essa noção é generalizada se, ao invés de estar restrito ao conjunto com dois elementos $\{0, 1\}$, o contradomínio da função característica for o intervalo real $[0, 1]$.

Definição 2.2.2 *Um subconjunto fuzzy A , do universo $\mathbf{U}$ (clássico), é caracterizado por uma função de pertinência $\mu_A : \mathbf{U} \longrightarrow [0, 1]$, onde $\mu_A(x)$ indica o grau de pertinência do elemento x em A ,*

$$A = \{(x, \mu_A(x)) \mid x \in \mathbf{U} \text{ e } \mu_A : \mathbf{U} \longrightarrow [0, 1]\}.$$

O valor $\mu_A(x) \in [0, 1]$ indica o grau com que o elemento x , de $\mathbf{U}$, está no conjunto *fuzzy* A . Quanto mais um elemento pertence a A , mais próximo de 1 será seu grau de pertinência. Os valores $\mu_A(x) = 0$ e $\mu_A(x) = 1$ indicam, respectivamente, a não pertinência e a pertinência completa de x , ao conjunto *fuzzy* A .

Segundo Klir e Yuan (1997) [30], o grau de pertinência $\mu_A(x)$, de um elemento x em A , pode ser interpretado como o grau de compatibilidade do atributo associado com A e o elemento x . Ortega (2001) [36], acrescenta que o grau de pertinência pode ser considerado como uma medida que expressa a possibilidade de que um dado elemento seja membro de um conjunto *fuzzy*.

Seguem abaixo algumas definições importantes.

Definição 2.2.3 *O suporte de um conjunto fuzzy A é o conjunto clássico de todos os elementos $x \in \mathbf{U}$, cuja função de pertinência tem valor diferente de zero:*

$$\text{supp}(A) = \{x \in \mathbf{U} \mid \mu_A(x) > 0\}.$$

Definição 2.2.4 *O conjunto de elementos que pertencem a um conjunto fuzzy A , pelo menos com um grau α , é chamado conjunto α - nível ou conjunto α - cut:*

$$A_\alpha = \{x \in A \mid \mu_A(x) > \alpha\}$$

Definição 2.2.5 *O núcleo (core) de um conjunto fuzzy A é o conjunto clássico de $\mathbf{U}$:*

$$\text{core}(A) = \{x \in \mathbf{U} \mid \mu_A(x) = 1\}.$$

Definição 2.2.6 *A altura de um conjunto fuzzy A é o maior grau de pertinência obtido, por qualquer elemento no conjunto, ou seja:*

$$hgt(A) = \sup_{x \in \mathbf{U}} \mu_A(x).$$

Definição 2.2.7 *A é dito normal se, e somente se, $\exists x \in \mathbf{U}$, tal que:*

$$\mu_A(x) = 1.$$

Essa definição implica que, num conjunto *fuzzy* normal A, $hgt(A) = 1$.

Normalização é a operação que converte um conjunto subnormal, não vazio, em sua versão normalizada, dividindo a função de pertinência original pela altura de A:

$$Norm[\mu_A(x)] = \frac{\mu_A(x)}{hgt(A)} \quad (2.1)$$

Concentração é a operação por meio da qual a função de pertinência se torna mais concentrada ao redor dos pontos com altos graus de pertinência:

$$Con[\mu_A(x)] = (\mu_A)^p(x) \quad (2.2)$$

onde $p > 1$.

Definição 2.2.8 *Para o conjunto vazio ϕ , tem-se $\mu_\phi(x) = 0, \forall x \in \mathbf{U}$ e , para o conjunto universo, tem-se, $\forall x, \mu_{\mathbf{U}}(x) = 1$.*

Definição 2.2.9 *Para qualquer conjunto fuzzy A, definido num universo finito $\mathbf{U}$, a cardinalidade é definida por*

$$|A| = \sum_{x \in U} \mu_A(x).$$

Quando $\mathbf{U}$ não é finito, $|A|$ nem sempre existe. Entretanto, se A tem suporte finito [19],

$$|A| = \sum_{x \in supp(U)} \mu_A(x).$$

As três operações básicas em conjuntos clássicos - complemento, intersecção e união - podem ser generalizadas, para conjuntos *fuzzy*, e estão definidas a seguir:

Definição 2.2.10 *O complementar $\bar{A}$, de um conjunto fuzzy A , com respeito ao universo $\mathbf{U}$, é definido, para todo $x \in \mathbf{U}$, por*

$$\mu_{\bar{A}}(x) = 1 - \mu_A(x).$$

Definição 2.2.11 *Dados dois conjuntos fuzzy A e B , a intersecção $A \cap B$ e a união $A \cup B$ são definidas, para todo $x \in \mathbf{U}$, por*

$$\mu_{A \cap B}(x) = \mu_A(x) \wedge \mu_B(x) \quad e \quad \mu_{A \cup B}(x) = \mu_A(x) \vee \mu_B(x),$$

onde $\wedge$ e $\vee$ denotam o operador mínimo e o operador máximo, respectivamente.

2.3 Lógica *fuzzy*

Lógica é a designação para o estudo de sistemas prescritivos de raciocínio, ou seja, sistemas que definem como se “deveria” pensar, usando a razão dedutivamente e indutivamente.

A lógica clássica trabalha com proposições que são verdadeiras ou falsas. A lógica *fuzzy*¹, é definida, segundo Santos (2003) [42], como a lógica que suporta os modos de raciocínio que são aproximados, ao invés de exatos, como estamos naturalmente acostumados a trabalhar. Nesta lógica, o raciocínio exato corresponde a um caso limite do raciocínio aproximado, sendo interpretado como um processo de composição nebulosa.

¹Também chamada de lógica difusa ou lógica nebulosa

Para Bilobrovec (2005) [9], a característica especial da lógica *fuzzy* é a de representar uma forma inovadora no manuseio de informações imprecisas, de forma muito distinta da teoria de probabilidades. A lógica *fuzzy* provê um método de traduzir expressões verbais, vagas, imprecisas e qualitativas, comuns na comunicação humana, em valores numéricos.

Assim, sua importância está na habilidade de lidar com proposições que não apresentam limites claramente definidos, ou seja, a lógica *fuzzy* não impõe limites bruscos, mas proporciona graus de pertinência de elementos a uma determinada categoria.

2.3.1 Variáveis Lingüísticas

Um conjunto *fuzzy* permite representar conceitos vagos, expressos em linguagem natural. Castanho (2005) [11], ressalta que essa representação não depende apenas do conceito, também depende do contexto no qual está inserido. Vários conjuntos *fuzzy* representando conceitos lingüísticos como *alto*, *médio* ou *baixo* são frequentemente empregados para definir o estado de uma variável. Tal variável é denominada variável lingüística ou variável *fuzzy*.

As variáveis e suas classificações lingüísticas são modeladas por conjuntos *fuzzy*, isto é, funções de pertinência que podem assumir variadas formas - linear por partes (triangular ou trapezoidal), quadrática, gaussiana, entre outras. A forma apropriada é determinada no contexto de uma aplicação em particular.

Os métodos para se obter essas funções de pertinência são: apelos intuitivos, ajustes de curvas, interpolações e até mesmo redes neurais [5].

Os rótulos, de conjuntos *fuzzy*, que descrevem uma variável lingüística são chamados termos primários, como por exemplo, *baixo*, *médio*, *alto*, entre outros. Podem ainda ser aplicados modificadores lingüísticos, a esses termos

primários, como *muito, pouco, mais ou menos*, dentre outros. Os termos primários também podem ser relacionados por meio da negação *não*, ou através dos conectivos *e* e *ou*.

A utilização de variáveis *fuzzy*, ou variáveis lingüísticas, facilita a transição gradual entre estados, pois estas possuem uma capacidade natural de expressar observações e medidas incertas. Além disso, com estas variáveis é possível converter informações qualitativas em formas que podem ser implementadas computacionalmente.

2.3.2 Relações *Fuzzy*

A compreensão de muitos fenômenos exige o estudo de relações ou interações entre os elementos de vários conjuntos. Segundo Barros e Bassanezi (2006) [5], uma relação clássica indica se há ou não associação entre dois objetos, enquanto que uma relação *fuzzy*, além de apontar se há ou não tal associação, indica o grau desta relação.

Uma proposição da forma, “ Se x é A e y é B , então z é C ou z é D ”, pode ser modelada como conjuntos *fuzzy*. Para isso, é necessário modelar os conectivos *ou* e *e*, assim como a condição *se ... então ...*. Para modelar o conectivo *e*, há uma classe de conectivos denominada *t-norma* e para modelar o conectivo *ou*, existem as *s-normas*, também conhecidas por *t-conormas*.

Definição 2.3.1 *Uma operação binária $t : [0, 1] \times [0, 1] \rightarrow [0, 1]$ é uma t-norma se satisfaz as seguintes condições:*

1. *Comutatividade: $xy = yx$;*
2. *Associatividade: $xt(ytz) = (xty)tz$;*
3. *Monotonicidade: Se $x \leq y$ e $w \leq z$ então $xw \leq yz$;*

4. *Condições de Fronteira:* $0tx = 0$ e $1tx = x$.

Definição 2.3.2 *Uma operação binária $s : [0, 1] \times [0, 1] \rightarrow [0, 1]$ é uma s -norma se satisfaz as seguintes condições:*

1. *Comutatividade:* $xsy = ysx$;
2. *Associatividade:* $xs(ysz) = (xsy)sz$;
3. *Monotonicidade:* Se $x \leq y$ e $w \leq z$ então $xsw \leq ysz$;
4. *Condições de Fronteira:* $0sx = x$ e $1sx = 1$.

A proposição “Se x é A , então y é B ” descreve uma relação entre as variáveis x e y . Uma relação clássica entre conjuntos clássicos A e B pode ser expressa por

$$\mathcal{R} = \{(x, y) \mid (x, y) \in A \times B\},$$

onde $A \times B$ representa o produto cartesiano entre A e B .

Definição 2.3.3 *Uma relação fuzzy $\mathcal{R}$, em $A \times B$, é um conjunto fuzzy caracterizado por uma função de pertinência $\mu_{\mathcal{R}}$, que associa a cada par ordenado (x, y) , um grau de pertinência em $\mathcal{R}$,*

$$\mathcal{R} = \{((x, y), \mu_{\mathcal{R}}(x, y)) \mid (x, y) \in A \times B, \mu_{\mathcal{R}}(x, y) \in [0, 1]\}.$$

Generalizando, se $\mathbf{U}$ é o produto cartesiano de n universos de discurso, $U_1 \times U_2 \times \dots \times U_n$, então, uma relação fuzzy $\mathcal{R}$ em $\mathbf{U}$ é um subconjunto fuzzy de $\mathbf{U}$, caracterizado por uma função de pertinência $\mu_{\mathcal{R}}(u_1, \dots, u_n)$, com $u_i \in U_i$, $i = 1, \dots, n$.

Definição 2.3.4 *O produto cartesiano $\mathcal{R}$, dos subconjuntos fuzzy, $A_1, A_2, \dots, A_n$ de $\mathbf{U}_1, \mathbf{U}_2, \dots, \mathbf{U}_n$ é a relação fuzzy cuja função de pertinência é*

$$\mu_{\mathcal{R}}(u_1, u_2, \dots, u_n) = \min [\mu_{A_1}(u_1), \mu_{A_2}(u_2), \dots, \mu_{A_n}(u_n)].$$

2.3.3 Regras *Fuzzy*

Uma proposição é um tipo particular de sentença, a saber, aquela que afirma ou nega um predicado de um sujeito. Como já foi dito anteriormente, a principal diferença entre proposições clássicas e proposições *fuzzy* é que, numa proposição *fuzzy* a verdade ou falsidade é uma questão de grau.

As proposições *fuzzy* são estruturas vastamente utilizadas em várias abordagens da teoria *fuzzy*. Conceitualmente, estas regras descrevem situações específicas que podem ser submetidas a análise de um painel de especialistas, e cuja inferência nos conduz a algum resultado desejado (Ortega, 2001) [36].

Uma proposição *fuzzy* básica pode ser expressa na forma canônica

$$p: x \text{ é } A,$$

onde, x é o sujeito que pode ser interpretado como uma variável lingüística e A é um conjunto *fuzzy* que representa um predicado *fuzzy*, ou seja, o estado da variável.

Uma proposição composta é construída a partir de conjunções,

$$p: x_1 \text{ é } A_1 \text{ e } x_2 \text{ é } A_2 \text{ e } \dots \text{ e } x_n \text{ é } A_n,$$

ou disjunções,

$$q: x_1 \text{ é } A_1 \text{ ou } x_2 \text{ é } A_2 \text{ ou } \dots \text{ ou } x_n \text{ é } A_n,$$

onde $A_1, A_2, \dots, A_n$ são conjuntos *fuzzy*, nos universos $\mathbf{U}_1, \mathbf{U}_2, \dots, \mathbf{U}_n$ e $x_1, x_2, \dots, x_n$ são variáveis.

Uma proposição condicional é uma regra do tipo, “Se (antecedente), então (conseqüente)”, onde o antecedente e o conseqüente são proposições *fuzzy*. Por exemplo:

$$p: \text{Se } x_1 \text{ é } A_1 \text{ e } \dots \text{ e } x_n \text{ é } A_n, \text{ então } y_1 \text{ é } B_1 \text{ e } \dots \text{ e } y_m \text{ é } B_m,$$

q: Se x_1 é A_1 ou ... ou x_n é A_n , então y_1 é B_1 ou ... ou y_m é B_m ,

onde $A_1, A_2, \dots, A_n$ são conjuntos *fuzzy* em $\mathbf{U}_1, \mathbf{U}_2, \dots, \mathbf{U}_n$ e $B_1, B_2, \dots, B_m$ são conjuntos *fuzzy* em $\mathbf{V}_1, \mathbf{V}_2, \dots, \mathbf{V}_m$ e $x_1, x_2, \dots, x_n$ e $y_1, y_2, \dots, y_m$ são variáveis.

As regras p e q induzem relações P e Q em $\mathbf{U}_1 \times \mathbf{U}_2 \times \dots \times \mathbf{U}_n \times \mathbf{V}_1 \times \mathbf{V}_2 \times \dots \times \mathbf{V}_m$, ou seja:

$$P(x_1, \dots, x_n, y_1, \dots, y_m) = f(P_a(x_1, \dots, x_n), P_c(y_1, \dots, y_m)),$$

$$Q(x_1, \dots, x_n, y_1, \dots, y_m) = f(Q_a(x_1, \dots, x_n), Q_c(y_1, \dots, y_m)),$$

onde f pode ser uma implicação, conjunção (t-normas) ou disjunção (s-normas) e P_a, Q_a e P_c, Q_c são as relações induzidas pelos antecedentes e conseqüentes das regras p e q . Logo, segundo Pedrycz e Gomide (1998) [39], as regras podem ser expressas como proposições:

$$p: (x_1, x_2, \dots, x_n, y_1, y_2, \dots, y_m) \text{ é } P,$$

$$q: (x_1, x_2, \dots, x_n, y_1, y_2, \dots, y_m) \text{ é } Q.$$

Uma regra pode ainda, ser ponderada, isto é ter a forma: “*Se (antecedente), então (conseqüente) [peso]*”, onde *peso* é um número real, no intervalo $[0, 1]$, que descreve a confiança da regra (Castanho, 2005) [11].

O peso é aplicado ao antecedente da regra. Um peso igual a 0 indica que a regra não é importante e pode até ser desprezada, enquanto um peso igual a 1 não modifica sua importância e não interfere na implicação. Graus de importância intermediários podem ser atribuídos ajustando os valores dos pesos no intervalo unitário.

A relação *fuzzy* é calculada incluindo-se o produto entre a relação induzida pelo antecedente e o peso, ou seja:

$$P(x_1, \dots, x_n, y_1, \dots, y_m) = f(P_a(x_1, \dots, x_n) * peso, P_c(y_1, \dots, y_m)).$$

A figura 2.2 mostra uma representação gráfica da regra: “Se x é A então y é B ”.

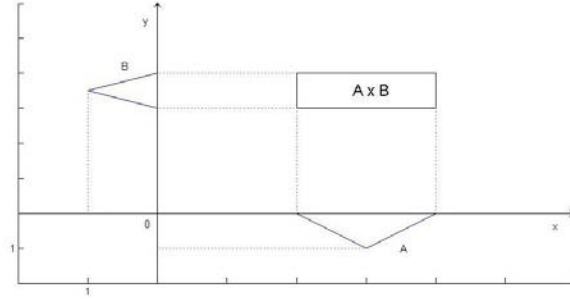


Figura 2.1: Representação gráfica da proposição p : Se x é A então y é B .

As regras do tipo “se ... então ...” podem ser modeladas através de conjunções, disjunções ou implicações *fuzzy*:

1. Conjunção *fuzzy* - é uma função $f_t : [0, 1]^2 \rightarrow [0, 1]$, definida por:

$$f_t(\mu_A(x), \mu_B(y)) = \mu_A(x) \wedge \mu_B(y), \quad \forall (x, y) \in \mathbf{U} \times \mathbf{V}.$$

O operador de Mamdani, que veremos adiante, é um exemplo de conjunção *fuzzy*.

2. Disjunção *fuzzy* - é uma função $f_s : [0, 1]^2 \rightarrow [0, 1]$, definida por:

$$f_s(\mu_A(x), \mu_B(x)) = \mu_A(x) \vee \mu_B(y), \quad \forall (x, y) \in \mathbf{U} \times \mathbf{V}.$$

3. Implicação *fuzzy* - é uma função $f_i : [0, 1]^2 \rightarrow [0, 1]$, que obedece as seguintes condições:

- Monotônica no segundo argumento:

$$\mu_B(y_1) \leq \mu_B(y_2) \Rightarrow f_i(\mu_A(x), \mu_B(y_1)) \leq f_i(\mu_A(x), \mu_B(y_2));$$

- Dominância da falsidade: $f_i(0, \mu_B(y)) = 1$;
- Neutralidade da verdade: $f_i(1, \mu_B(y)) = \mu_B(y)$.

Embora possa ser representada como conjunção ou implicação, uma regra *fuzzy* é matematicamente, descrita por uma relação *fuzzy* entre os conjuntos que descrevem o antecedente e o conseqüente.

2.3.4 Inferência

O método de inferência é onde se definem quais são os *conectivos lógicos* usados para estabelecer a relação *fuzzy* que modela a base de regras. O sucesso do sistema *fuzzy* depende deste método, já que ele fornecerá a saída (controle) *fuzzy* à ser adotada pelo controlador, a partir de cada entrada *fuzzy* (Amendola, Souza e Barros, 2005) [3].

Pressupondo que R é uma relação *fuzzy* em $\mathbf{X} \times \mathbf{Y}$ e A' e B' são conjuntos *fuzzy* em $\mathbf{X}$ e $\mathbf{Y}$, respectivamente, então, se R e A' são dados, obtém-se B' da equação

$$\mu_{B'}(y) = \sup_{x \in \mathbf{X}} \min [\mu_{A'}(x), \mu_R(x, y)], \quad \forall y \in \mathbf{Y}. \quad (2.3)$$

Uma proposição *fuzzy* da forma, “p: Se x é A então y é B ”, pode ser interpretada como uma relação *fuzzy*,

$$\mu_R(x, y) = f [\mu_A(x), \mu_B(y)], \quad (2.4)$$

onde f denota uma relação *fuzzy* que descreve o antecedente e o conseqüente.

Seja a proposição q da forma “ q : x é A' ”. Considerando a proposição p como uma regra e a proposição q como um fato, temos o esquema:

Regra: Se x é A então y é B

Fato: x é A'

Conclusão: y é B'

Esse procedimento é chamado *modus ponens* generalizado (Zimmermann, 1993) [53]. No esquema B' é calculado por (2.3) e R é determinado por (2.4). Quando $A' = A$ e $B' = B$ e os conjuntos são clássicos, o esquema torna-se o *modus ponens* clássico.

A inferência, em proposições com múltiplos antecedentes, inclui operações de conjunção (com o conectivo *e*) ou disjunção (com o conectivo *ou*), antes da conclusão de B' . Em geral, o método de inferência é aplicado à várias regras, gerando muitas conclusões B'_i . A conclusão final do sistema é a disjunção de todos os B'_i

O processo de inferência baseado no *modus ponens* não é único. Os mais divulgados são o Método de Mamdani e o Método de Kang-Takagi-Sugeno (KTS).

Método de Mamdani

No método de inferência de Mamdani, a relação associada com uma regra particular é obtida por meio de uma conjunção do antecedente e o conseqüente da regra e, numa coleção de regras, a agregação é realizada através da união das relações individuais.

Consideremos o raciocínio *fuzzy* - como no esquema a seguir - com n regras combinadas com os conectivos *e* e *ou*:

No esquema, A_{jk} e C_j são subconjuntos *fuzzy* nos universos de discurso, $\mathbf{X}_k$ e $\mathbf{Z}_k$, e k é o número de conjuntos *fuzzy* no antecedente.

Regra 1: Se x_1 é A_{11} , ..., x_k é A_{1k} então z_1 é C_1

ou

Regra 2: Se x_1 é A_{21} , ..., x_k é A_{2k} então z_2 é C_2

ou

$\vdots$

$\vdots$

ou

Regra n : Se x_1 é A_{n1} , ..., x_k é A_{nk} então z_n é C_n

Fato: $A = x_1$ é A_1 e x_2 é A_2 e ... e x_k é A_k

Conclusão: z é $C = R(A)$

O método de Mamdani propõe uma relação *fuzzy* binária $\mathcal{M}$, entre x e z , e baseia-se na regra de composição de inferência max-min. A seguir estão os passos para determinar C , segundo Barros e Bassanezi (2006) [5]:

1. Em cada regra R_j , da base de regras *fuzzy*, a condicional “se x é A_j **então** z é C_j ” é modelada pela aplicação $\wedge$ (mínimo);
Se a regra for ponderada, o peso w_j deve ser incluído no grau de ativação da regra,
2. Adota-se a t-norma $\wedge$ (mínimo) para o conectivo lógico “e”;
3. Para o conectivo lógico “ou” adota-se a t-conorma $\vee$ (máximo), que conecta as regras *fuzzy*.

A relação *fuzzy* $\mathcal{M}$ é o subconjunto *fuzzy* de $\mathbf{X} \times \mathbf{Z}$ cuja função de pertinência é

$$\mu_{\mathcal{M}}(x, z) = \max_{1 \leq i \leq r} (\mu_{R_i}(x, z)) = \max_{1 \leq i \leq r} [\mu_{A_j}(x) \wedge \mu_{C_j}(z)],$$

onde r é o número de regras e, A_j e C_j são os subconjuntos *fuzzy* da regra j . Os valores $\mu_{A_j}(x)$ e $\mu_{C_j}(z)$ são interpretados como os graus com que x e z estão nos subconjuntos *fuzzy* A_j e C_j , respectivamente.

A figura 2.3 ilustra as saídas parciais, do método de inferência de Mamdani, para duas entradas reais, x_0 e y_0 , uma saída e duas regras.

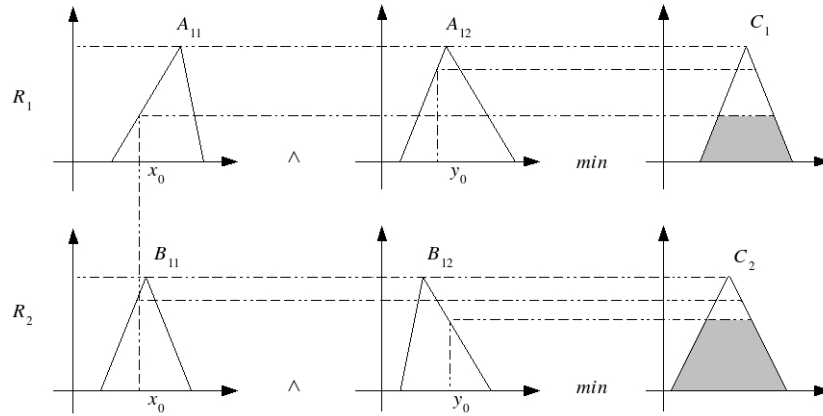


Figura 2.2: Saídas parciais no método de Mamdani.

A saída final C é dada pela união das saídas parciais, como mostra a figura 2.4.

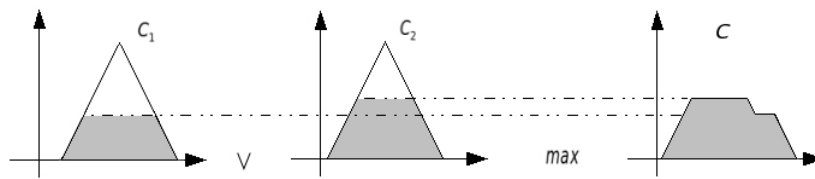


Figura 2.3: Saída final do controlador *fuzzy* de Mamdani.

O resultado do método de Mamdani, último gráfico da figura 2.4, representa a função de pertinência μ_C do controle C que foi obtido por meio dos procedimentos descritos. Para se obter um número real, que represente o conjunto, é necessário utilizar algum método de defuzzificação.

Método de Kang-Takagi-Sugeno

Em muitos aspectos este método é similar ao método de Mamdani. A principal diferença entre a inferência *fuzzy* do tipo Mamdani e a do tipo Sugeno é que, nesse último, o conseqüente de cada regra é dado explicitamente por uma função dos valores de entrada desta regra.

O método de raciocínio, proposto por Takagi e Sugeno, está associado a um conjunto de regras do tipo, “*se ... então*”, cujos conseqüentes são representados por funções lineares (Ortega, 2001) [36]. Um conjunto de regras nesse modelo possui a seguinte forma:

Regra (i): Se x_1 é A_{i1} , ..., x_k é A_{ik} então $z_i = c_{i0} + c_{i1}x_1 + \dots + c_{ik}x_k$,

onde $i = 1, \dots, n$; n é o número de regras *fuzzy*; c_{ij} , com $j = 0, \dots, k$, são os parâmetros do conseqüente e A_{ij} são subconjuntos *fuzzy*.

Dada uma entrada $(x_1, x_2, \dots, x_n)$, o resultado final deste método é a média ponderada das saídas de cada regra, isto é

$$z = \frac{\sum_{i=1}^n m_i z_i}{\sum_{i=1}^n m_i} = \frac{\sum_{i=1}^n m_i (c_{i0} + c_{i1}x_1 + \dots + c_{ik}x_k)}{\sum_{i=1}^n m_i}, \quad (2.5)$$

onde m_i é o peso ou grau de ativação da i -ésima regra

$$m_i = \min (\mu_{A_{ij}}(x_j)).$$

As funções lineares nos conseqüentes podem ser substituídas por funções não-lineares [11]. Nesse caso, as regras são do tipo:

Regra (i): Se x_1 é A_{i1} , ..., x_k é A_{ik} então $z_i = f_i(x_1, \dots, x_k)$,

nas quais as saídas são combinadas da mesma forma que no caso linear, via expressão (2.5).

Segundo Ortega (2001), a grande vantagem desse modelo consiste no seu poder de representação, especialmente para descrever processos tecnológicos. Ele permite a decomposição de um sistema complexo em um conjunto de sub-sistemas mais simples. Em uma ativação realística, seria praticamente impossível a decomposição de um sistema em sub-sistemas clássicos, uma vez que o conhecimento disponível é fragmentado e não há fronteiras bem delimitadas para os sistemas.

Métodos de Decodificação

Os métodos de decodificação ou *defuzzificação*, como também são chamados, selecionam um número real que seja representativo do conjunto *fuzzy*. Podemos citar, dentre outros métodos que têm sido propostos, o Centro de Gravidade (ou Centróide, ou Centro de Área), Centro dos Máximos, Média dos Máximos e Método das Alturas.

O *Método do Centro de Gravidade* é semelhante à média aritmética para uma distribuição de dados. A diferença é que os pesos são os valores $\mu_C(z_i)$, que apontam o grau de compatibilidade do valor z_i , com o conceito modelado pelo conjunto *fuzzy* C [5].

O centro de gravidade fornece a média das áreas de todas as figuras que representam os graus de pertinência de um conjunto *fuzzy*. Esse método é considerado natural do ponto de vista do senso comum, porém, a computação requerida é um tanto complexa (Mizumoto, 1995) [32].

As equações (2.6) e (2.7) referem-se aos domínios discreto e contínuo, respectivamente.

$$G(C) = \frac{\sum_{i=0}^n z_i \mu_C(z_i)}{\sum_{i=0}^n \mu_C(z_i)}. \quad (2.6)$$

$$G(C) = \frac{\int_{\mathbb{R}} z \mu_C(z) dz}{\int_{\mathbb{R}} \mu_C(z) dz} \quad (2.7)$$

No *Centro dos Máximos*, o ponto representativo do conjunto é obtido como uma média entre as regiões de maior possibilidade, entre os possíveis valores da variável. Neste caso temos:

$$C(C) = \frac{i + j}{2} \quad (2.8)$$

onde

$$\begin{aligned} i &= \inf \left\{ z \in \mathbb{R} : \mu_C(z) = \max_z \mu_C(z) \right\} \\ &\quad \text{e} \\ j &= \sup \left\{ z \in \mathbb{R} : \mu_C(z) = \max_z \mu_C(z) \right\}. \end{aligned}$$

O *Método da Média dos Máximos* é utilizado como defuzzificador para domínio discreto. Sua definição é dada por

$$M(C) = \frac{\sum z_i}{\sum i}, \quad (2.9)$$

onde z_i são os elementos de maiores pertinências ao conjunto *fuzzy* C , isto é, para cada i tomamos

$$\mu_C(z_i) = \max_z \mu_C(z).$$

No *Método das Alturas* a saída é uma média ponderada dos pontos z_i , pelas alturas de C_i , ou seja:

$$A(C) = \frac{\sum_{i=1}^n m_i \text{hgt}(C_i)}{\sum_{i=1}^n \text{hgt}(C_i)}, \quad (2.10)$$

onde n é o número de regras ativadas e m_i é o grau de ativação da regra.

Esse método pode ser considerado como um caso especial do método de Kang-Takagi-Sugeno, no qual a saída de um controlador *fuzzy* é uma função ou um número real.

2.4 Sistema Baseado em Regras *Fuzzy*

Os sistemas baseados em regras *fuzzy* - SBRF - devido a natureza multidisciplinar, recebem várias denominações, tais como: sistema de inferência *fuzzy*, sistema especialista *fuzzy*, modelo *fuzzy*, controlador lógico *fuzzy* ou simplesmente sistema *fuzzy* [11].

A lógica *fuzzy* é uma técnica que incorpora a forma humana de pensar, em um sistema de controle. Segundo Shaw e Simões (1999) [44], um controlador *fuzzy* típico pode ser projetado para comportar-se conforme o raciocínio dedutivo, isto é, o processo que as pessoas utilizam para inferir conclusões baseadas em informações que elas já conhecem.

A habilidade de explicitar, linguisticamente, relações que, ou são muito complexas ou não são suficientemente bem entendidas, para serem descritas por modelos matemáticos precisos, é uma vantagem dos sistemas *fuzzy*.

Shaw e Simões (1999) citam o *Princípio da Incompatibilidade* de Zadeh, o qual afirma que

“Conforme a complexidade de um sistema aumenta, nossa habilidade de fazer afirmações precisas e significativas sobre seu comportamento diminui, até um limiar em que a precisão e relevância tornam-se praticamente características mutuamente exclusivas”.

A estrutura básica, de um sistema baseado em regras *fuzzy*, consiste em quatro módulos principais: um processador de entrada, uma base de conhe-

cimento, uma máquina de inferência e um processador de saída, como está representado na figura 2.5.

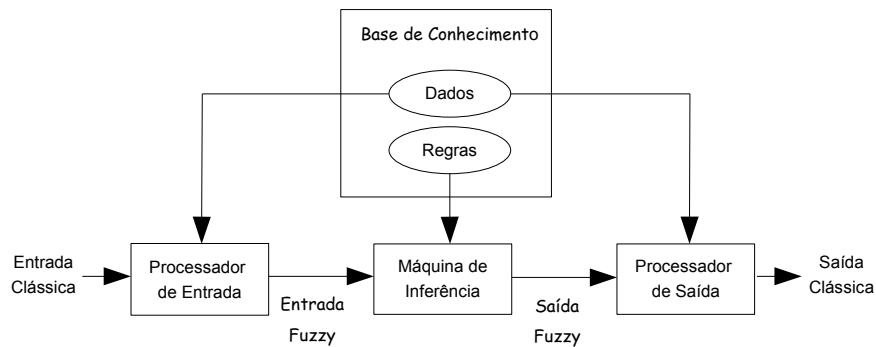


Figura 2.4: Estrutura básica de um sistema baseado em regras *fuzzy*.

Processador de Entrada

Identificadas as variáveis relevantes, de entrada e saída do sistema, e o intervalo de valores de cada variável, são atribuídos termos lingüísticos, que descrevem os estados destas variáveis. Estes termos são traduzidos - em um domínio apropriado - por uma função de pertinência, em um subconjunto *fuzzy*. Essa fase é chamada *fuzzificação* ou codificação.

As funções de pertinência representam os aspectos fundamentais de todas as ações teóricas e práticas dos sistemas *fuzzy*. Estas funções podem ser geradas de várias maneiras, incluindo apelos intuitivos e auxílio de especialistas.

Klir e Yuan (1997) [30], apresentam dois métodos baseados na teoria matemática de ajuste de curvas: o método de interpolação de Lagrange e o método de mínimos quadrados. Civanlar e Trussell (1986) [16], determinaram funções de pertinência, para conjuntos *fuzzy*, a partir da função densidade de probabilidade, de uma característica definida de seus elementos, no universo

de discurso. Shaw e Simões (2001) [44], ressaltam que técnicas especiais usando redes neurais ou algoritmos genéticos são utilizadas para a geração de funções de pertinência.

Base de Conhecimento

A base de conhecimento contém um conjunto de regras *fuzzy*, denominado base de regras e um conjunto de funções de pertinência, que formam a base de dados.

A base de regras *fuzzy* é formada por uma coleção de proposições do tipo “se ... então ...”, que podem ser construídas de acordo com as informações de um especialista ou de um banco de dados.

Máquina de Inferência

O método de inferência determina a forma operacional do modelo lingüístico. Trata-se de um mapeamento, que define uma transformação do valor *fuzzy* de entrada em um valor de saída.

Existem vários métodos de inferência, de acordo com a concepção de regra *fuzzy* (conjunção ou implicação) e com a definição dos operadores de união e intersecção. Devemos portanto, escolher aquele que, em algum sentido, melhor se adapta ao sistema que se está modelando. O método mais utilizado é o método de Mamdani.

Processador de Saída

Na maioria das aplicações práticas de sistemas *fuzzy*, a saída final deve ser um número real, ao invés de um conjunto *fuzzy*.

A decodificação ou defuzzificação é um procedimento que nos permite interpretar a distribuição de possibilidades, da saída de um modelo lingüístico *fuzzy*, de forma quantitativa, ou seja, ele nos fornece um valor numérico

representativo, que captura o significado essencial dessa distribuição de possibilidades (Ortega, 2001) [36].

2.5 Possibilidade e Probabilidade

A teoria de conjuntos *fuzzy* é baseada no fato de que os conjuntos existentes no mundo real não possuem limites precisos. A “característica *fuzzy*” implica em existência de imprecisão, incerteza e definições qualitativas (Shaw e Simões, 2001) [44].

Modelar problemas do mundo real envolve essas incertezas, que podem ser de dois tipos distintos: incertezas que provêm da falta de conhecimento relacionado a conceitos bem definidos, ou ainda, podem resultar da vagueza inerente aos próprios conceitos.

Nesse contexto, um conceito importante é o de medida *fuzzy*. Uma medida *fuzzy* representa a incerteza com que se afirma que um dado elemento pertence a um conjunto clássico.

Definição 2.5.1 *Seja Ω um conjunto não vazio e $\mathbf{P}(\Omega)$ o conjunto das partes de Ω . Uma medida fuzzy é uma função*

$$\varphi : \mathbf{P}(\Omega) \rightarrow [0, 1]$$

que satisfaz as seguintes condições:

1. $\varphi(\emptyset) = 0$ e $\varphi(\Omega) = 1$ (*condições de contorno*);
2. $\forall A, B \in \mathbf{P}(\Omega)$, se $A \subseteq B$ então $\varphi(A) \leq \varphi(B)$ (*monotonicidade*).

Possibilidade, intuitivamente, refere-se à percepção do grau de facilidade de obtenção ou realização de algo, enquanto probabilidade está associada à

freqüência com que algo ocorre. Assim, o que é possível pode não ser provável e o que é improvável não é necessariamente, impossível.

Ortega (2001) [36], reafirma que a distribuição de possibilidades é diferente da distribuição de probabilidades, pois estas lidam com diferentes tipos de incertezas. É importante ressaltar que o conceito de possibilidade não envolve, de modo algum, a noção de experimentação repetida e por isso é um conceito não-estatístico.

Em termos de possibilidade, dizer que uma variável aleatória X , em $\mathbf{U}$, poderia estar em um subconjunto clássico A de $\mathbf{U}$ é equivalente a

$$\Pi(X = x) = \pi_X(x) = \begin{cases} 1, & \text{se } x \in A \\ 0, & \text{se } x \notin A \end{cases} \quad (2.11)$$

De acordo com Ortega (2001), assim como ocorre com os conjuntos clássicos, quando uma variável lingüística é definida, ela é restrita a um conjunto de valores. O que diferencia as duas abordagens é justamente a noção de valores possíveis e impossíveis, que na lógica *fuzzy* é expressa por diferentes graus.

A extensão da representação (2.11), para um conjunto *fuzzy*, quando A é normal, é

$$\Pi(X = x) = \pi_X(x) = \mu_A(x).$$

Uma proposição *fuzzy* da forma, “ X é A ”, onde X é uma variável que toma valores num universo de discurso $\mathbf{U}$ e A é um subconjunto *fuzzy* de $\mathbf{U}$, induz a distribuição de possibilidade Π_X , que é igual a A , isto é

$$\Pi_X = A.$$

Assim, a distribuição de possibilidade de X é um conjunto *fuzzy* e serve para definir a possibilidade com que X pode assumir algum valor específico

em $\mathbf{U}$. Ou seja, se $x \in \mathbf{U}$ e $\mu_A : \mathbf{U} \rightarrow [0, 1]$ é a função de pertinência de A , então a possibilidade de que $X = x$, dada a proposição, “ X é A ”, é

$$\pi \{X = x \mid X \text{ é } A\} = \mu_A(x), x \in \mathbf{U}.$$

A é o conjunto de mais ou menos possíveis valores para X .

Dada uma distribuição de possibilidade Π_X , a possibilidade de x pertencer a um conjunto clássico B é definida como:

$$\pi(X \in B) = \sup_{x \in B} \pi_X(x).$$

Essa definição é consistente, diante da noção intuitiva de que a possibilidade de ocorrência de qualquer evento, dentre vários, corresponde à possibilidade de ocorrência daquele que é mais possível.

As medidas de possibilidade e probabilidade são medidas *fuzzy* e serão definidas a seguir.

Definição 2.5.2 *Uma medida de possibilidade Π , sobre Ω , é uma função de $\mathbf{P}(\Omega)$ em $[0, 1]$, tal que:*

1. $\Pi(\phi) = 0$;
2. $\Pi(\Omega) = 1$;
3. $\Pi(\bigcup_i A_i) = \sup_i \Pi(A_i)$, para qualquer família $\{A_i\}$ de subconjuntos de Ω ,

onde $\mathbf{P}(\Omega)$ é o conjunto das partes de Ω .

Definição 2.5.3 *Uma medida de probabilidade P é uma função real que determina, para todo A em Ω , uma probabilidade $P(A)$, tal que:*

1. $P(A) \geq 0, \quad \forall A \in \Omega$;

2. $P(\Omega) = 1$;

3. $P(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$, onde A_i é uma coleção de eventos disjuntos.

Quando informações sobre um fenômeno são dadas em termos probabilísticos e possibilísticos, as duas descrições devem ser, de alguma forma, consistentes (Klir e Yuan, 1995) [30]. Isto é, dada uma medida de probabilidade P e uma medida de possibilidade Π , definidas em $\mathbf{P}(\Omega)$, as duas medidas devem satisfazer alguma condição de consistência.

Embora possibilidade e probabilidade sejam medidas distintas, existe uma relação entre elas. Segundo Dubois e Prade (1980) [19], o grau de possibilidade de um evento é maior ou igual ao seu grau de probabilidade, ou seja,

$$P(A) \leq \Pi(A) \quad \forall A \subset \Omega,$$

onde P é a medida de probabilidade associada a p ,

$$P(A) = \sum_{x \in A} p(x)$$

e Π é a medida de possibilidade associada a π ,

$$\Pi(A) = \sup_{x \in A} [\pi(x)].$$

É essencial, para qualquer transformação probabilidade/possibilidade, obedecer o princípio de consistência acima. Klir e Yuan (1995), afirmam que essas transformações são úteis em problemas práticos, como por exemplo, na construção de funções de pertinência de um conjunto *fuzzy*, a partir de dados estatísticos e na combinação de informações possibilísticas e probabilísticas em sistemas especialistas.

Um outro aspecto que diferencia as duas teorias é o fato da teoria de probabilidades não considerar subjetividades. Para citar o diagnóstico médico, podemos dizer que o raciocínio médico parece estar muito mais baseado

em graus de possibilidade do que de probabilidade, uma vez que seria humanamente impossível, para o médico, guardar todas as informações exatas relacionadas à frequência dos sintomas e à prevalência das doenças em uma dada população.

Ortega (2001), acrescenta ainda que os médicos raramente expressam suas impressões em valores numéricos; são usados termos lingüísticos tanto para se expressar quanto para associar cognitivamente os sintomas/doenças.

Transformações probabilidade/possibilidade

Sejam $X = \{x_1, x_2, \dots, x_n\}$, $p_i = p(x_i)$ e $\pi_i = \pi(x_i)$, com $i = 1, \dots, n$ e, ainda, $\pi_1 = 1$ e $\sum_{i=1}^n p_i = 1$ sejam normalizações de possibilidade e probabilidade, respectivamente. Suponha que os elementos de X estejam ordenados de tal forma que, as distribuições de possibilidade

$$\Pi = \langle \pi_1, \pi_2, \dots, \pi_n \rangle$$

e as distribuições de probabilidade

$$P = \langle p_1, p_2, \dots, p_n \rangle$$

sejam sempre seqüências não crescentes. Então, as transformações mais simples, $P \leftrightarrow \Pi$, são baseadas na razão escalar $\pi_i = p_i \alpha_i$, $\forall i$, onde α é uma constante positiva e são expressas pelas equações:

$$\pi_i = \frac{p_i}{p_1}, \quad (2.12)$$

$$p_i = \frac{\pi_i}{\sum_{i=1}^n \pi_i}. \quad (2.13)$$

2.6 Considerações Finais

A teoria de conjuntos *fuzzy* possui propriedades que a tornam apropriada para lidar com os vários níveis de imprecisão e incerteza, tão evidentes na área médica, de forma mais crítica e realista.

Os termos lingüísticos, que descrevem as características das doenças - em particular o câncer, que é o objeto de estudo nesse trabalho - são intrinsecamente vagos, e muitas vezes são variáveis qualitativas em medicina. Nesse sentido, a lógica *fuzzy* apresenta-se como uma alternativa poderosa, capaz de se aproximar da forma de raciocinar do médico.

Capítulo 3

Modelagem *Fuzzy* para prever o estágio do Câncer de Próstata

3.1 Introdução

O interesse, de profissionais e pesquisadores das mais diferentes áreas, em resolver problemas cada vez mais complexos tem produzido o desenvolvimento dos computadores - cada vez mais eficientes - e impulsionado os avanços da Inteligência Artificial.

A informática médica acompanha esta evolução e o computador auxilia o médico nas mais diversas atividades. Segundo Ortega (2001) [36], as contribuições da engenharia médica, no desenvolvimento de sistemas especialistas, são inegáveis. A lógica *fuzzy* tem participado efetivamente desse processo e isso se deve aos tipos de incertezas envolvidos nos procedimentos médicos, biológicos e epidemiológicos.

De acordo com Castanho (2005) [11], a natureza imprecisa e, muitas vezes contraditória, da informação ou a sua falta devem ser consideradas em toda decisão médica, pois podem ter consequências importantes, ou mesmo vitais,

para o paciente. Além disso, o processo pelo qual o médico usa seu conhecimento para inferir um diagnóstico, a partir das informações apresentadas pelo paciente e dos resultados de exames, é um processo complexo, baseado em conhecimento expreso linguisticamente e na presença de diversos tipos de incerteza e imprecisão. Diante disso, a teoria dos conjuntos fuzzy é apropriada para desenvolver sistemas especialistas que auxiliarão os médicos em suas decisões.

Kalmanson e Stegall (1975) [29], indicaram a teoria *fuzzy* como uma alternativa no processo de tomada de decisão, por considerarem-na adequada às reais necessidades clínicas e experimentais. Adlassnig (1980 e 1986) [1] [2], descreveu o CADIAG - 2, um dos primeiros sistemas de diagnóstico computadorizado, que utiliza a teoria dos conjuntos *fuzzy* para formalizar as relações médicas e modelar o processo de diagnóstico. Aplicações dessa teoria também foram feitas em modelagem médica, como epidemiologia (Barros, Leite e Bassanezi, 2003; Ortega, 2001) [7] [36] e evolução da AIDS (Jafelice, 2003) [28].

Castanho (2005) [11] construiu, pela primeira vez, um modelo matemático, que consiste num Sistema Baseado em Regras *Fuzzy* - SBRF - com a finalidade de prever o estágio patológico do câncer de próstata. Esse modelo é parte integrante da tese de doutorado, dessa autora, defendida na FEEC/UNICAMP, em 2005.

Nossa intenção, nesse capítulo, é fazer uma leitura desta modelagem *fuzzy*, mostrando e discutindo modificações feitas na construção do modelo, e apresentando algumas simulações de pacientes e os resultados obtidos. Ressaltamos que toda a implementação do modelo foi feita no MATLAB 7.

3.2 Nomogramas para o Câncer de Próstata

Tradicionalmente, o julgamento clínico tem formado a base para as estimativas de riscos, aconselhamento de pacientes e tomadas de decisões. Entretanto, segundo Stephenson e Kattan (2006) [49], as pessoas têm dificuldades para prever resultados, devido às linhas e tendências que existem, para todos os estágios do processo de predição.

De acordo com esses autores, os clínicos não tratam todos os casos igualmente e assim, determinados casos podem permanecer latentes ou exercer, impropriamente, uma larga influência em futuras predicções de resultados.

Na tentativa de obter maior precisão, os pesquisadores têm desenvolvido ferramentas predictivas, baseadas em modelos estatísticos. Stephenson e Kattan (2006) ressaltam que esses modelos têm mostrado performance semelhante ou, em muitas situações, até melhor do que o julgamento clínico.

No caso do câncer de próstata, uma das dificuldades que o especialista enfrenta, após diagnosticá-lo, é estabelecer o estágio em que este se encontra para que possa administrar a terapia apropriada.

Uma taxa de cura de mais de 85% é estimada, pela Sociedade Americana de Câncer, se o tumor estiver confinado na próstata e a glândula for removida através da prostatectomia radical (Wang *et al*, 2006) [50]. Quando a doença apresenta metástase, o objetivo passa a ser o alívio dos sintomas e, portanto, outras terapias serão administradas.

Segundo Gann e Han (2005) [23], a grande maioria dos cânceres de próstata são detectados incidentalmente, por meio de ultrasonografias transretais ou em cirurgias realizadas, em razão da hiperplasia benigna da próstata. Partin e colaboradores (1997) [37], afirmam que aproximadamente 60% dos pacientes, diagnosticados clinicamente como portadores de tumor confinado na próstata, na realidade apresentam a doença em estágio avançado, não

sendo mais possível eliminá-la por cirurgia ou radiação.

Assim, a avaliação correta do estágio do câncer é essencial, pois os clínicos recomendarão o tratamento aos seus pacientes e os riscos, das terapias curativas intensivas, justificam-se apenas se existirem razoáveis possibilidades de êxito.

O estadiamento do câncer de próstata torna-se, significativamente, mais preciso, se combinados os parâmetros de extensão local da doença (estado clínico - conferir tabela 1.3), o nível de PSA e o escore de Gleason (Carter e Partin, 1997) [10].

Uma alternativa que vem sendo utilizada, nos últimos anos, é o desenvolvimento de modelos multivariáveis contínuos, chamados *nomogramas*. Stephenson e Kattan (2006) [49], definem um nomograma como sendo uma representação gráfica, de uma fórmula matemática ou algoritmo, que incorpora diferentes predictores, modelados como variáveis contínuas, para prever um resultado particular.

Partin e colaboradores (1997 e 2001) [37] [38], construíram tabelas de probabilidades, baseadas no estado clínico do tumor, PSA e grau de Gleason, pré-cirúrgicos. Esses autores utilizaram informações precisas do estágio patológico, de um grande número de indivíduos submetidos à prostatectomia radical e, por meio de uma regressão multinomial log-linear, foi encontrada a probabilidade de a doença estar em um dos seguintes estágios: câncer localizado na próstata, localmente avançado ou metastático - com envolvimento das vesículas seminais e/ou dos linfonodos pélvicos.

É importante observar que, Partin *et al* consideraram os estágios do câncer de próstata mutuamente exclusivos, ou seja, a probabilidade de a doença estar localmente avançada foi calculada, desconsiderando o fato de que é possível, nessas condições, que o tumor já tenha atingido as vesículas seminais.

O mesmo ocorreu no cálculo da probabilidade de ter envolvimento dos linfonodos pélvicos, onde foi excluída a possibilidade de ter, simultaneamente, o envolvimento das vesículas seminais.

Kattan e colaboradores (2006), também desenvolveram nomogramas - baseados no estado clínico do tumor, nível de PSA e grau de Gleason - que fornecem as probabilidades, para os diferentes estágios do câncer de próstata. O modelo estatístico de Kattan *et al* foi construído a partir da análise de regressão logística modificada, restrita a splines cúbicos.

Quando se utiliza o modelo de regressão não modificada, como foi o caso de Partin *et al*, assume-se que mudanças incrementais representam o mesmo significado, sobre um espectro de valores. Por exemplo, assume-se que uma variação no nível de PSA, de 2 ng/ml para 4 ng/ml, pode representar o mesmo impacto que uma mudança de 302 ng/ml para 304 ng/ml.

De acordo com Stephenson e Kattan (2006), a aplicação de splines cúbicos possibilita a flexibilidade do nomograma, permitindo que variáveis contínuas mantenham as conexões não-lineares. Com isso, foi possível considerar os estágios da doença não mutuamente exclusivos, ou seja, os autores levaram em conta as sobreposições dos estágios.

Esses e outros nomogramas, desses autores, estão disponíveis ao público em geral, no site do *Memorial Sloan-Kettering Cancer Center*, para download (<http://www.mskcc.org>) ou mesmo para efetuar online o cálculo das probabilidades, referentes aos estágios do câncer de próstata.

A figura 3.1 ilustra a utilização dos nomogramas, considerando um determinado quadro clínico.

Stephenson e Kattan (2006), destacam ainda que outros métodos de modelagem, tais como redes neurais artificiais, podem proporcionar maior flexibilidade do que os métodos estatísticos tradicionais. No entanto, as dificul-

Memorial Sloan-Kettering Cancer Center

Prostate Nomogram - Pre-Treatment

[Change Treatment Stage](#)

Pre-treatment PSA:	15
Biopsy Primary Gleason:	3
Biopsy Secondary Gleason:	3
Biopsy Gleason Sum:	6
1992 Clinical Tumor Stage:	T2a
1997 Clinical Tumor Stage:	T2a
Prescribed External Radiation Dose (64.8 - 86.4 Gy):	
Neo-Adjuvant Hormones:	☐
Neo-Adjuvant Radiation:	☐

[Clear](#) [Calculate](#)

Results

Organ Confined Disease	38%
Extra Capsular Penetration	52%
Seminal Vesicle Involvement	5%
Lymph Node Involvement	4%
5yr Progression Free Probability Radical Prostatectomy	73%
5yr Progression Free Probability External Beam Radiation Therapy	NA
5yr Progression Free Probability Brachytherapy	75%

[Print](#)

ADDITIONAL TOOLS

[Volume](#) [Life Expectancy](#) [PSA Doubling Time](#)

HELP

[Glossary](#) [FAQ](#)

[About](#) | [Website](#) | [Disclaimer](#) | [Contact Us](#) | [Restart](#)

©2007 Memorial Sloan-Kettering Cancer Center

Figura 3.1: Cálculo das probabilidades, para os diferentes estágios do câncer de próstata, considerando um determinado quadro clínico.

dades encontram-se na reprodução e interpretação desses modelos.

Após essa breve descrição dos nomogramas, vamos comentar suas vantagens e limitações.

Segundo Chun e colaboradores (2006) [14], controvérsias envolvem a questão do valor clínico dos nomogramas. Sabe-se que todo nomograma depende da amostra de pacientes, utilizada em sua construção. Isso quer dizer que a ferramenta predictiva, desenvolvida em uma determinada época, pode não fornecer predicções igualmente satisfatórias, em pacientes contemporâneos. Além disso, se a amostra de pacientes possuir um número maior de indivíduos de uma determinada raça, o nomograma pode não ser aplicável a indivíduos de outras origens.

Um outro fato importante é que as variáveis utilizadas, por Kattan *et al* e por Partin *et al*, são variáveis incertas. Como vimos anteriormente, o estado clínico é obtido pelo exame do toque retal, que depende da experiência

e percepção do especialista. Algumas vezes, a ultra-sonografia transretal é indicada, como exame complementar, pois pode auxiliar na identificação de áreas suspeitas. Contudo, esse procedimento também detecta nódulos intra-prostáticos de outra natureza, como: infarto prostático, cistos e áreas de fibrose (Ferreira e Nardi, 1999) [22].

Existe uma relação entre aumento nos níveis de PSA e estágio patológico. Entretanto, segundo Epstein (1997) [20], níveis elevados de PSA nem sempre estão associados a características patológicas avançadas (câncer localmente avançado, com envolvimento de vesículas seminais ou de linfonodos) e níveis baixos não necessariamente sugerem doença localizada na glândula prostática.

No primeiro capítulo vimos que, na classificação clássica, o nível de PSA no sangue é considerado normal até 4 ng/ml, embora a linha de corte entre normal e alterado não seja totalmente clara. Um estudo de Catalona, Smith e Ornstein (1997) [12], mostrou que 22% dos homens com câncer de próstata, têm níveis de PSA entre 2,6 e 4 ng/ml. Zhu e colaboradores (2005) [52], em uma outra pesquisa, mostraram que biópsias feitas em homens, com níveis de PSA entre 2,6 e 4 ng/ml, frequentemente detectaram câncer de próstata, clinicamente significativo, no estágio localizado.

Além disso, nos dois nomogramas que descrevemos, os níveis de PSA foram considerados nos intervalos ≤ 4 , $(4, 10]$, $(10, 20]$, > 20 e, dessa forma, dois pacientes com níveis de PSA iguais a 10 e 10,1 ng/ml, com mesmo estado clínico e mesmo escore de Gleason, têm diferentes probabilidades de estar em um determinado estágio patológico. Já um paciente com nível de PSA igual a 20 ng/ml, tem a mesma probabilidade daquele com 10,1 ng/ml.

O grau de Gleason é atribuído por um patologista, após a análise de uma amostra do tumor, que pode ser bastante heterogêneo. Tem-se, portanto, um

valor numérico preciso representando uma situação subjetiva.

Com o objetivo de lidar com essas incertezas e atenuar a transição entre as classificações dessas variáveis, foi desenvolvido um sistema baseado em regras *fuzzy*, para prever o estadiamento patológico do câncer de próstata.

3.3 O Modelo *Fuzzy*

Como visto na seção 2.5, um sistema baseado em regras *fuzzy* - SBRF - compreende quatro módulos principais: um módulo de fuzzificação ou codificador, que representa as variáveis de entrada e saída do sistema, por conjuntos *fuzzy*; um módulo de inferência; uma base de regras e um módulo de defuzzificação ou decodificador, que transforma a saída, que é um conjunto *fuzzy*, em um valor numérico (crisp).

A figura 3.2 ilustra o esquema do SBRF que utilizamos.

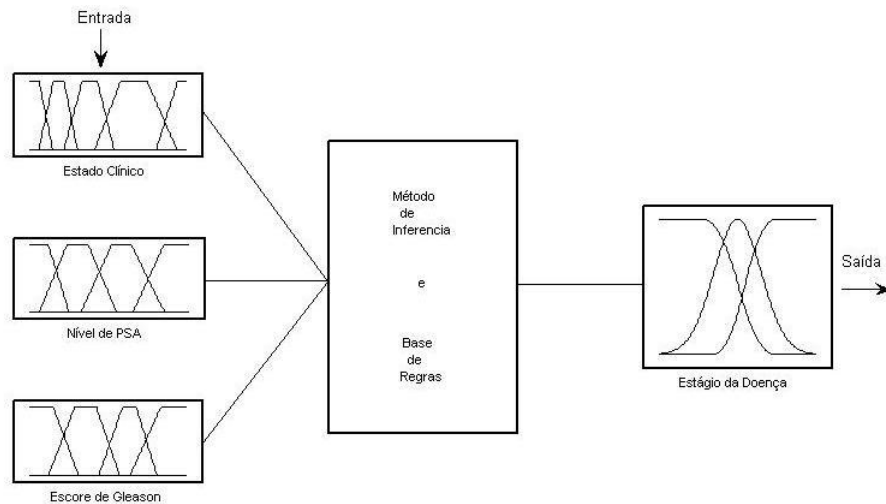


Figura 3.2: Estrutura básica de um sistema baseado em regras *fuzzy*.

Na modelagem *fuzzy*, as variáveis de entrada - *Estado Clínico*, *Nível de PSA* e *Escore de Gleason* - e a variável de saída - *Estágio da Doença* - foram

consideradas como variáveis lingüísticas e seus valores como conjuntos *fuzzy*, em seus respectivos domínios. Essas variáveis que utilizamos são as mesmas usadas por Castanho (2005), em seu modelo *fuzzy*. Partin *et al* (1997) [37] e Stephenson e Kattan (2006) [49], também as utilizaram na construção de seus respectivos nomogramas, como visto na seção anterior.

As variáveis quantitativas são fuzzificadas diretamente e as variáveis qualitativas são arranjadas numa escala linear (por exemplo, de 0 a 1), onde são quantificadas as observações do especialista. Segundo Castanho (2005) [11], tal procedimento nos permite classificar as observações nas fronteiras de uma maneira natural.

Para a variável lingüística *Estado Clínico* foram atribuídos cinco estágios, traduzidos por conjuntos *fuzzy*, com base na tabela TNM (conferir tabela 1.3). Assim como feito por Castanho (2005), os estágios T1a, T1b e T1c foram considerados apenas como T1 (não palpável), pois o que os difere na tabela são o nível de PSA e o grau de diferenciação do tumor, que também são variáveis deste modelo. A figura 3.3 mostra a representação gráfica dessa variável.

O estado clínico relaciona-se com o tamanho do tumor e, portanto, o domínio [0 1] representa o percentual de comprometimento da próstata, ou seja, 0 representa uma próstata sem tumor, enquanto 1 representa a próstata totalmente tomada pelo tumor. A partir do estágio T3, o tumor já se espalhou além da próstata e, por essa razão, os estágios seguintes não serão considerados na modelagem.

Na classificação clássica, o nível de PSA é considerado normal - até 4 ng/ml - níveis entre 4 e 10 ng/ml são considerados levemente elevados, entre 10 e 20 ng/ml são moderadamente elevados e acima de 20 ng/ml são altamente elevados. Assim, para a variável *Nível de PSA*, utilizamos os ter-

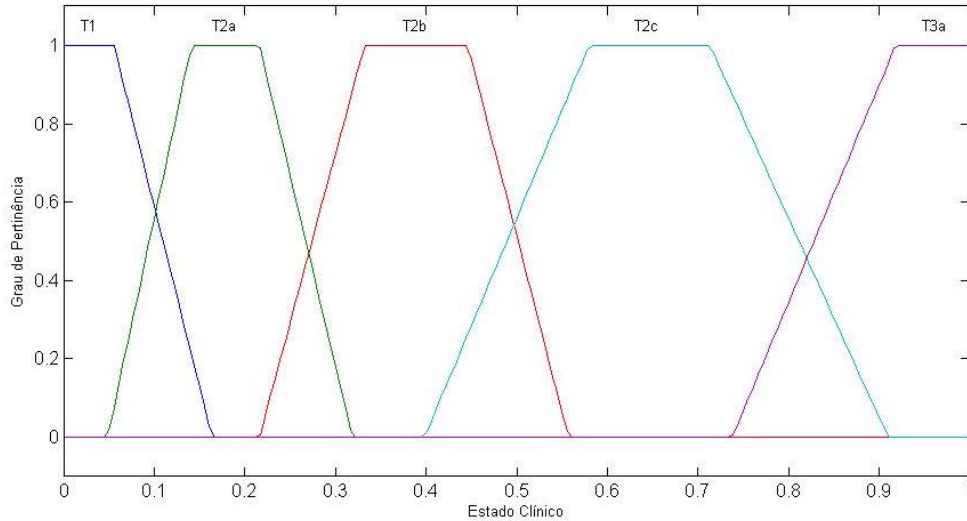


Figura 3.3: Funções de pertinência dos conjuntos *fuzzy*, assumidos pela variável lingüística *Estado Clínico*.

mos lingüísticos *Normal*, *Levemente Elevado*, *Moderadamente Elevado* e *Altamente Elevado*, como ilustrado na figura 3.4.

A variável *Escore de Gleason*, na modelagem de Castanho (2005), foi classificada com os termos lingüísticos: *Bem Diferenciado*, para representar as lesões menos agressivas - graus 2, 3 e 4 - *Moderadamente Diferenciado* - graus 5 e 6 - *Pouco Diferenciado* - grau 7 - e *Indiferenciado*, para representar as lesões mais agressivas - graus 8, 9 e 10.

Optamos por fazer uma modificação nessa classificação: o termo lingüístico *Moderadamente Diferenciado*, que representa os graus de Gleason 5 e 6, foi subdividido em *Médio Diferenciado* (grau 5) e *Médio-Baixo Diferenciado* (grau 6).

Essa mudança foi feita pois, como veremos adiante, para construir a base de regras do sistema, precisamos das probabilidades - para os diferentes estágios da doença - presentes nos nomogramas (de Stephenson e Kattan (2006),

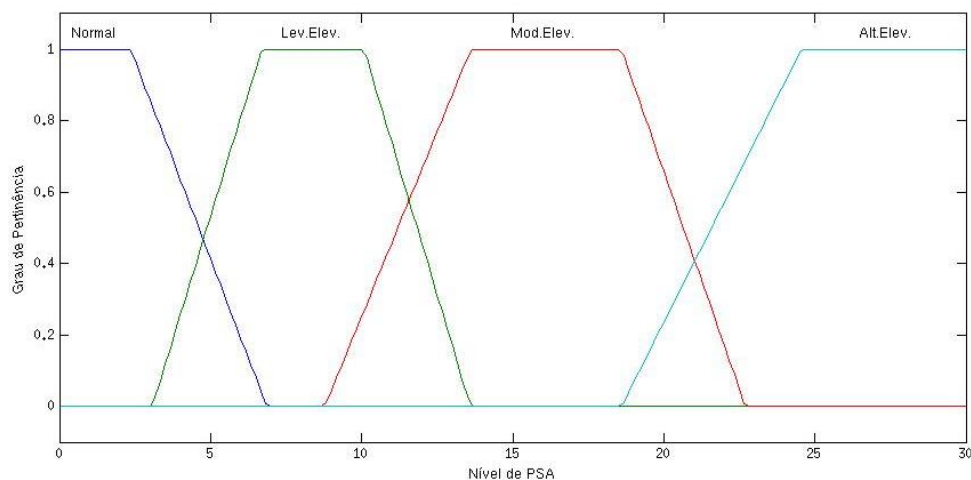


Figura 3.4: Funções de pertinência dos conjuntos *fuzzy*, assumidos pela variável lingüística *Nível de PSA*.

em nosso caso), que resultam de todas as diferentes combinações entre estado clínico, nível de PSA e escore de Gleason; as combinações envolvendo os graus 5 e 6, apresentaram resultados bastante distintos. Então, na tentativa de discriminar um maior número de possibilidades, a classificação para esses graus foi separada.

A figura 3.5 mostra a representação gráfica da variável Escore de Gleason.

Para a variável de saída - *Estágio da Doença* - foram atribuídos, por Castanho (2005), os seguintes termos: *Confinado*, indicando que todo o câncer está confinado na glândula prostática; *Capsular*, se, mesmo que o câncer já esteja fora da glândula prostática, as vesículas seminais e os linfonodos pélvicos estiverem livres do tumor e *Vesícula + Linfonodos*, se o câncer já invadiu a parede muscular de uma ou ambas as vesículas seminais e/ou os linfonodos pélvicos.

Após contatos com o especialista, prof. Dr. Ubirajara Ferreira, da Faculdade de Ciências Médicas (FCM) e do Hospital das Clínicas (HC), da

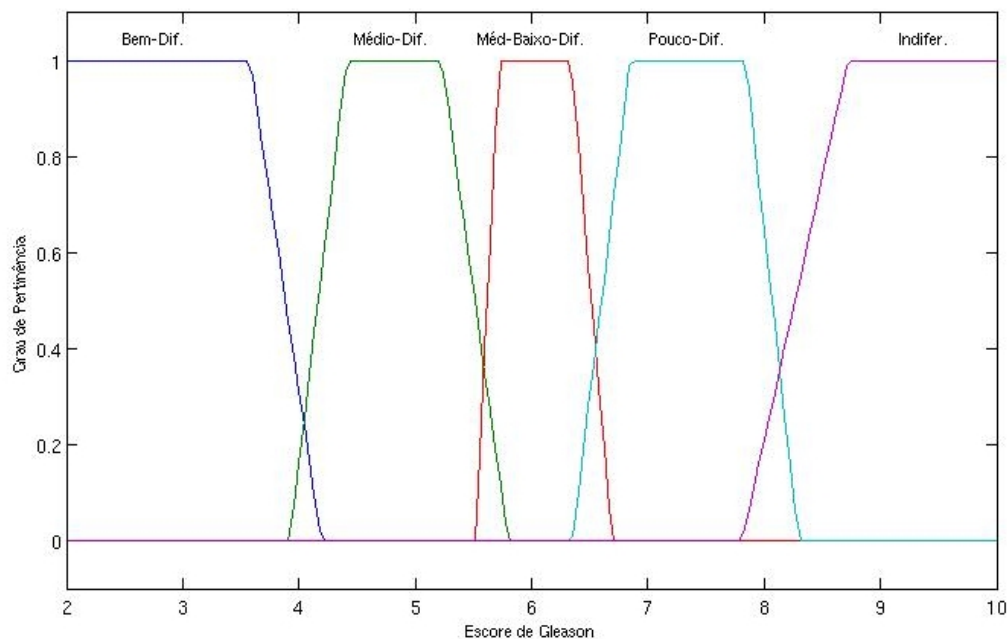


Figura 3.5: Funções de pertinência dos conjuntos *fuzzy*, assumidos pela variável lingüística *Escore de Gleason*.

UNICAMP, obtivemos a informação de que, atualmente, os especialistas utilizam os termos lingüísticos, *doença confinada* ou *localizada*, *doença localmente avançada* e *doença extra-prostática* ou *metastática*, para designarem os três estágios do câncer de próstata, descritos acima. A razão para essas denominações é que, recentemente, os especialistas descobriram que a glândula prostática não possui uma cápsula (envoltório fibroso).

Diante disso, classificamos a variável de saída - *Estágio da Doença* - com os termos: *Localizado*, *Localmente Avançado* e *Metastático*. Essa variável é qualitativa e, por esse motivo, foi escolhida uma escala de 0 a 1 para indicar a extensão do câncer: quanto mais próximo de 0, menor o tamanho do tumor e, quanto mais próximo de 1, maior o comprometimento das estruturas adjacentes à próstata.

As funções de pertinência dos conjuntos *fuzzy* foram construídas de forma

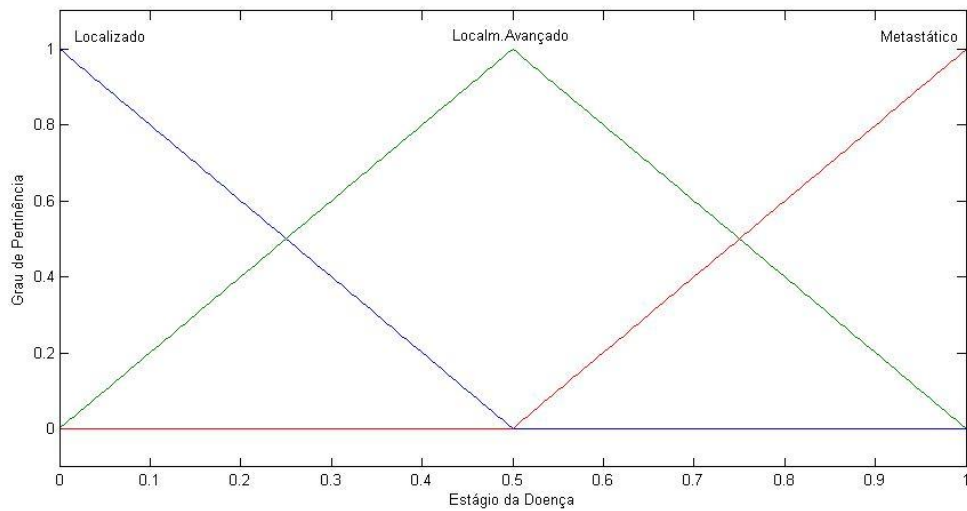


Figura 3.6: Funções de pertinência dos conjuntos *fuzzy*, assumidos pela variável lingüística *Estágio da Doença*.

triangular, como mostra a figura 3.6.

A escolha dessa forma, para as funções de pertinência da saída, foi uma espécie de “chute” inicial. Nas seções seguintes veremos que é possível, a partir dos resultados que serão encontrados e dos nomogramas que utilizamos, adaptar essas funções e melhorar a saída do sistema.

A base de regras é construída, normalmente, de acordo com as informações de um especialista. Este conjunto de regras *fuzzy* desempenha o papel da função matemática, para obter a saída do sistema.

Castanho (2005) [11] construiu sua base de regras tendo em vista as informações contidas nos nomogramas de Partin *et al* (1997) [37]. Segundo essa autora, os médicos utilizavam tais nomogramas, para predizer o estágio patológico da doença, quando disponíveis os dados pré-cirúrgicos dos pacientes.

Na modelagem que fizemos, a base de regras foi elaborada a partir das informações contidas nos nomogramas de Stephenson e Kattan (2006) [49]. Essa mudança foi feita pois, de acordo com o especialista, anteriormente ci-

tado, atualmente os médicos utilizam esses nomogramas, quando necessário.

A seguir apresentamos, como exemplo, o procedimento adotado na construção de uma das regras.

Nos nomogramas, a probabilidade de um paciente com *Estado Clínico* igual a T2a, *Nível de PSA* menor que 4 ng/ml (normal) e *Escore de Gleason* igual a 5 (médio diferenciado), estar com a doença *localizada* é de 68%, 30% de estar *localmente avançada* e 2% de estar com a doença *metastática*. Com essas informações foi construída a regra número 6 da tabela 3.1:

“Se *Estado Clínico* é T2a, *Nível de PSA* é Normal e *Escore de Gleason* é Médio Diferenciado, então *Estágio da Doença* é Localizado (1)”.

De modo inteiramente análogo, isto é, fazendo todas as diferentes combinações entre estado clínico, nível de PSA e escore de Gleason, levando em conta os termos lingüísticos atribuídos à cada uma dessas variáveis e verificando as probabilidades - constantes nos nomogramas - para os diferentes estágios da doença, foi construída a base de regras, formada por 156 regras.

Observamos que, assim como na modelagem de Castanho (2005), as regras foram ponderadas quando a diferença de probabilidade, de estar em dois estágios distintos, foi menor do que 20%. Assim, por exemplo, para um paciente cujo *Estado Clínico* é T2b, *Nível de PSA* é normal (< 4) e *Escore de Gleason* é 5, temos nos nomogramas: 57% para *Localizado*, 40% para *Localmente Avançado* e 3% para *Metastático*. Com isso, foram construídas as regras 07 e 08 (ver tabela 3.1).

A regra 07 tem peso igual a 0,6 e a regra 08 tem peso 0,4, indicando que, com esses dados clínicos, a chance de o câncer estar localizado é um pouco maior do que de estar localmente avançado.

Definida a base de regras, precisamos *traduzi-la* matematicamente, como

N	EST.CLÍN.	PSA	GLEASON	ESTÁGIO	PESO
03	T2b	Normal	Bem Dif.	Localizado	1
06	T2a	Normal	Médio Dif.	Localizado	1
07	T2b	Normal	Médio Dif.	Localizado	0,6
08	T2b	Normal	Médio Dif.	Loc. Avan.	0,4
40	T1	Lev. Elev.	Bem Dif.	Localizado	1
61	T3a	Lev. Elev.	Méd-B. Dif.	Loc. Avan.	1
112	T3a	Mod. Elev.	Pouco Dif.	Loc. Avan.	0,45
113	T3a	Mod. Elev.	Pouco Dif.	Metastático	0,55
134	T2a	Alt. Elev.	Méd-B. Dif.	Loc. Avan.	1
154	T2c	Alt. Elev.	Indif.	Metastático	1

Tabela 3.1: Algumas das 156 regras, da Base de Regras construída. Na 1^o coluna consta o n^o da regra e, na última, o peso utilizado para ponderá-la.

vimos na seção 2.4, por meio das técnicas da lógica *fuzzy*, a fim de obtermos a saída do sistema. Para isso, utilizamos o Método de Inferência de Mamdani.

A defuzzificação do conjunto *fuzzy*, que representa a saída do sistema, foi feita pelo Método do Centro de Gravidade, para que, enfim, a saída obtida fosse um número real. Esse valor pertence ao intervalo $[0, 1]$ e a ele corresponde uma pertinência aos conjuntos *fuzzy*: Localizado, Localmente Avançado e Metastático, que representam o estágio da doença (ver figura 3.6).

3.3.1 Simulações

É possível determinar o grau de compatibilidade com determinado estágio da doença, isto é, o grau de pertinência ao conjunto *fuzzy*, que representa aquele estágio, a partir dos dados pré-cirúrgicos dos pacientes.

Foram feitas simulações do modelo e alguns resultados encontram-se na tabela 3.2. Os dados considerados são de pacientes do Hospital das Clínicas da UNICAMP, obtidos em setembro de 2006, junto ao especialista que contactamos, mencionado na seção anterior.

Est.Clin.	PSA	Gleason	Saída	Localiz.	Loc. Avan.	Metast.
T1	12	5	0,393	0,21	0,79	0,00
T2a	3	7	0,446	0,11	0,89	0,00
T2a	3,5	6	0,230	0,54	0,46	0,00
T2b	15,4	9	0,590	0,00	0,82	0,18
T2c	18,2	7	0,550	0,00	0,90	0,10
T3a	17,2	7	0,570	0,00	0,86	0,14
T3a	23,7	7	0,805	0,00	0,40	0,60

Tabela 3.2: Pertinência aos conjuntos *fuzzy* que representam os estágios da doença: Localizado na Próstata, Localmente Avançado e Metastático.

Assim, um paciente com tumor palpável (T2a), nível de PSA igual a 3,5 e escore de Gleason igual a 6, tem pertinência maior ao conjunto *Localizado* (0,54). Já outro paciente, com mesmo estado clínico (T2a), nível de PSA menor (3), porém com escore de Gleason maior (7), tem pertinência maior ao conjunto *Localmente Avançado* (0,89).

Os resultados obtidos, em sua maioria, mostraram-se coerentes com as probabilidades fornecidas pelos nomogramas, no sentido de que, o estágio que tem maior probabilidade de ocorrência, também tem maior grau de pertinência. Por exemplo, para o paciente cujo estado clínico é T2c, nível de PSA é 18,2 e escore de Gleason é 7, a probabilidade correspondente à pertinência 0,90, para o estágio localmente avançado, é 0,45, que é a maior dentre as probabilidades para esse quadro clínico.

Entretanto, no caso dos dois pacientes, acima citados, a probabilidade correspondente à pertinência 0,54 é 0,64 - que é a maior dentre as probabilidades - e à pertinência 0,89 é 0,45, que não é a maior das probabilidades.

Com o intuito de estabelecer uma relação entre pertinência ao conjunto e probabilidade, Castanho (2005) encontrou uma forma de adaptar as funções de pertinência, que descrevem a saída do sistema, a partir da probabilidade de ocorrência de cada estágio.

Vamos proceder de forma análoga e adaptar as funções, que descrevem os conjuntos *fuzzy* da saída (ver figura 3.6), utilizando as probabilidades de Stephenson e Kattan (2006).

3.3.2 Relação entre Probabilidade e Possibilidade

Os nomogramas mostram a *probabilidade* de um paciente, com determinado quadro clínico - nível de PSA, estado clínico e escore de Gleason - estar em um dos estágios: tumor localizado, localmente avançado ou metastático (envolvimento das vesículas seminais e/ou linfonodos pélvicos).

Assim, temos uma função densidade de probabilidade de X , onde a variável, aleatória discreta X , representa os pacientes com determinado quadro clínico, cujo espaço amostral S é definido por:

$$S = \{Localizado, Loc.Avan., Metast.\} = \{x_1, x_2, x_3\}.$$

Cada resultado possível x_i , é associado a um número,

$$p(x_i) = P(X = x_i),$$

denominado probabilidade de x_i , que satisfaz as seguintes condições:

1. $p(x_i) \geq 0, \forall i$;

$$2. \sum_{i=1}^3 p(x_i) = 1, \text{ ou seja, } p(\text{Localizado}) + p(\text{Loc.Avan.}) + p(\text{Metast.}) = 1.$$

A figura 3.7 mostra a representação da distribuição de probabilidade, dos pacientes com um quadro clínico específico.

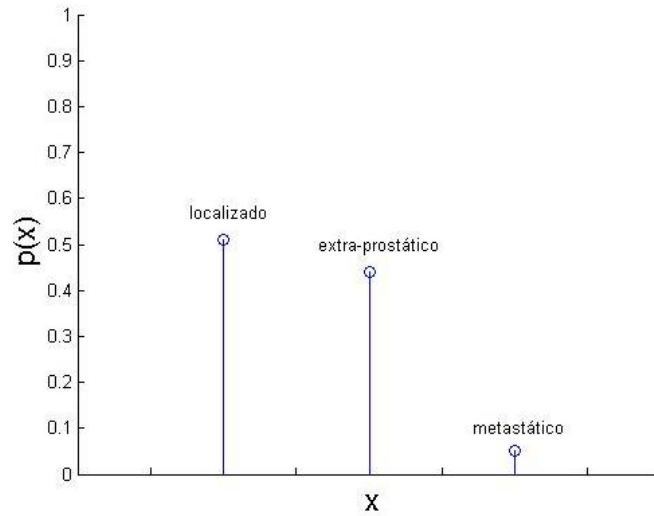


Figura 3.7: Distribuição de probabilidade, da variável aleatória X , que representa os pacientes com Estado Clínico T2a, níveis de PSA entre 4,1 e 10 ng/ml e escore de Gleason 6.

Combinando os diferentes estados clínicos (T1, T2a, T2b, T2c, T3a), níveis de PSA (0 a 30 ng/ml) e escores de Gleason (2-4, 5, 6, 7, 8-10), de uma amostra de indivíduos, obtivemos, através do sistema baseado em regras *fuzzy*, um número real α , pertencente ao intervalo $[0, 1]$, para cada combinação. Cada número α encontrado representa uma subpopulação de pacientes, com mesmo quadro clínico e, portanto, tem uma distribuição de probabilidade como descrita anteriormente.

Para cada estágio da doença, uma curva foi ajustada aos pontos que representam a probabilidade, fornecida pelos nomogramas, de o paciente estar

naquele estágio. O ajuste foi feito pelo método dos mínimos quadrados. É possível construir, a partir desses ajustes, uma nova função de pertinência aos conjuntos *fuzzy*, que representam a saída do sistema.

O processo foi repetido, isto é, foram simuladas - no SBRF - várias amostras de indivíduos, com diferentes quadros clínicos e novas curvas foram ajustadas às probabilidades de cada estágio. Isso foi feito até que não houvesse diferença significativa entre a saída anterior e a nova.

Na figura 3.8 estão representadas as funções de pertinência aos conjuntos *fuzzy*, da saída do sistema, obtidas com os ajustes de curvas.

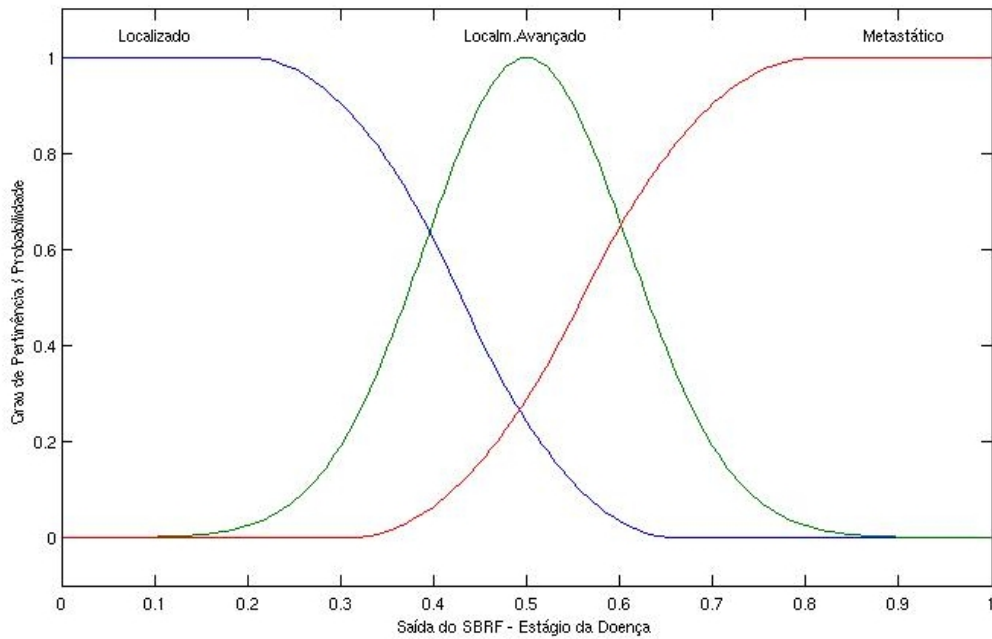


Figura 3.8: Funções de pertinência que descrevem os conjuntos *fuzzy*, que representam os estágios do câncer.

As funções de pertinência aos conjuntos *fuzzy*, *Localizado*, *Localmente Avançado* e *Metastático*, representam o grau com o qual o paciente se enquadra a determinado estágio da doença.

Por exemplo, de acordo com o que vimos na seção 2.6, a proposição

Estágio é Localizado permite que esse grau seja visto como a *possibilidade* de que a doença esteja localizada na glândula prostática. Dessa forma, essa proposição define uma distribuição de possibilidade de X , onde X representa os pacientes em um estágio da doença. Esta distribuição está na figura 3.8.

Relação Probabilidade/Possibilidade

Vimos, no capítulo anterior, que uma relação entre probabilidade e possibilidade deve obedecer ao princípio de consistência,

$$P(A) \leq \Pi(A), \quad \forall A \subset X$$

ou seja, a probabilidade de um evento ocorrer é sempre menor ou igual à possibilidade.

Utilizando a saída do sistema baseado em regras *fuzzy*, como mostrado na figura 3.7, foram feitas outras simulações, considerando os dados de pacientes do Hospital das Clínicas, da UNICAMP. As tabelas 3.3 e 3.4 contêm alguns resultados, obtidos em termos de possibilidade e transformados em probabilidades, através da equação (2.13), vista na seção 2.6.

Como exemplo, na primeira tabela consta um paciente com estado clínico T2a, isto é, com tumor palpável, envolvendo uma metade (ou menos) de um dos lobos. O nível de PSA é 11 e a biópsia revela um escore de Gleason igual a 6.

De acordo com o SBRF, a possibilidade de que esse paciente tenha câncer de próstata localizado na glândula é 0,13, de que tenha a doença localmente avançada é 0,84 e de que apresente metástase é 0,01. Convertendo essas possibilidades em probabilidades tem-se 33%, 58% e 9%, respectivamente. Esse resultado é coerente com os nomogramas, nos quais as probabilidades são 38%, 52% e 9%, respectivamente.

		SBRF		Kattan <i>et al</i>
Paciente	Estágio	Possib.	Probab.	Nomograma
Est. Clínico T2a	Localizado	0,47	33%	38%
PSA 11	Loc. Avançado	0,84	58%	52%
Gleason 6	Metastático	0,13	9%	9%
Est. Clínico T2b	Localizado	0,24	16%	22%
PSA 7,8	Loc. Avançado	1	65%	54%
Gleason 7	Metastático	0,29	19%	24%
Est. Clínico T2a	Localizado	0,11	8%	10%
PSA 27,9	Loc. Avançado	0,90	61%	51%
Gleason 7	Metastático	0,46	31%	39%
Est. Clínico T3a	Localizado	0,07	5%	6%
PSA 15,1	Loc. Avançado	0,81	57%	40%
Gleason 7	Metastático	0,54	38%	54%
Est. Clínico T1	Localizado	0,65	50%	61%
PSA 12	Loc. Avançado	0,62	46%	33%
Gleason 5	Metastático	0,05	4%	6%
Est. Clínico T2a	Localizado	0,24	16%	22%
PSA 16,69	Loc. Avançado	1	65%	55%
Gleason 7	Metastático	0,29	19%	23%
Est. Clínico T2a	Localizado	0,32	70%	64%
PSA 2,9	Loc. Avançado	0,34	29%	34%
Gleason 6	Metastático	0,01	1%	2%

Tabela 3.3: Comparação dos resultados, obtidos pelo SBRF, com as probabilidades de Kattan et al [49]. A coluna *Probab.* é a transformação da possibilidade em probabilidade, via equação (2.13).

		SBRF		Kattan <i>et al</i>
Paciente	Estágio	Possib.	Probab.	Nomograma
Est. Clínico T2c	Localizado	0,20	13%	15%
PSA 11,3	Loc. Avançado	0,99	65%	45%
Gleason 7	Metastático	0,34	22%	40%
Est. Clínico T3a	Localizado	0,07	5%	6%
PSA 15	Loc. Avançado	0,81	57%	40%
Gleason 7	Metastático	0,54	38%	54%
Est. Clínico T1	Localizado	0,86	75%	47%
PSA 5,24	Loc. Avançado	0,28	24%	44%
Gleason 6	Metastático	0,01	1%	9%
Est. Clínico T2a	Localizado	0,82	70%	64%
PSA 2,8	Loc. Avançado	0,34	29%	34%
Gleason 6	Metastático	0,01	1%	2%
Est. Clínico T1	Localizado	1	98%	62%
PSA 0,26	Loc. Avançado	0,02	2%	32%
Gleason 5	Metastático	0	0%	2%
Est. Clínico T2c	Localizado	0,16	11%	15%
PSA 12	Loc. Avançado	0,97	64%	45%
Gleason 7	Metastático	0,38	25%	40%
Est. Clínico T3a	Localizado	0,16	10%	6%
PSA 5,71	Loc. Avançado	0,96	63%	34%
Gleason 8	Metastático	0,40	27%	59%

Tabela 3.4: Comparação dos resultados, obtidos pelo SBRF, com as probabilidades constantes nos nomogramas de Kattan et al [49].

Na segunda tabela encontram-se alguns resultados que se mostraram mais otimistas, em relação ao estágio da doença, do que as probabilidades presen-

tes nos nomogramas.

Por exemplo, para um paciente cujo estado clínico é T1 - tumor não palpável - nível de PSA é 5,24 e escore de Gleason é 6, o SBRF fornece como possibilidades: 0,86 para doença localizada; 0,28 para doença localmente avançada e 0,01 para o câncer metastático. Transformando esses valores em probabilidades tem-se, respectivamente, 75%, 24% e 1%.

Observemos que nos nomogramas de Kattan *et al* [49] constam as seguintes probabilidades: 47% para doença localizada; 44% para o câncer localmente avançado e 9% para metástases.

Nesse caso, portanto, os resultados obtidos pelo SBRF apresentaram uma probabilidade maior de a doença estar localizada na glândula e uma menor probabilidade de que haja metástase. Além disso, houve uma diminuição na probabilidade de que o câncer esteja localmente avançado.

Embora tenham ocorrido outros casos como este em nossas simulações, a maioria dos resultados encontrados estão coerentes com as probabilidades calculadas por Kattan *et al*. Isto é, o estágio que tem maior probabilidade de ocorrência, também tem o maior grau de pertinência e os valores estão próximos.

3.4 Considerações Finais

Saber em qual estágio um paciente com câncer de próstata se encontra é extremamente importante, pois a cirurgia apresenta um alto índice de cura se a doença estiver confinada na glândula.

Nesse capítulo apresentamos uma alternativa, que é a construção de um sistema baseado em regras *fuzzy* - SBRF - para prognosticar o estágio patológico de pacientes com câncer de próstata.

A vantagem, com relação aos nomogramas, está no fato de que a transição gradual entre as faixas da tabela, e entre os estágios, é levada em consideração. Nesse sentido, o SBRF é mais coerente com a realidade.

Suponha, por exemplo, pacientes com mesmo estado clínico, mesmo escore de Gleason e níveis de PSA iguais a 10, 10,1 e 20 ng/ml. Diferentemente do que ocorre se utilizarmos os nomogramas, os estágios do câncer para os pacientes com níveis de PSA 10 e 10,1 ng/ml são considerados, pelo SBRF, mais próximos e para os pacientes com níveis de PSA 10,1 e 20 ng/ml, mais distantes.

Um outro aspecto importante é que a utilização da teoria de conjuntos *fuzzy* permite incluir a subjetividade da linguagem, existente nas variáveis. Com isso, o SBRF tornou-se mais abrangente e o médico pode avaliar a variável *estado clínico* numa escala contínua entre 0 e 1. Cada valor do intervalo corresponde a um grau de pertinência, que pode ser considerado como a certeza de seu prognóstico.

Podemos citar ainda, a relativa facilidade de interpretação e reprodução desse modelo preditivo, quando comparado com os métodos estatísticos tradicionais.

Contudo, mesmo diante de todas essas diferenças citadas, entre os modelos estatísticos e o modelo *fuzzy*, os resultados encontrados nas simulações mostraram-se semelhantes aos resultados presentes nos nomogramas de Kattan *et al* [49].

Segundo o médico especialista que contactamos, prof. Dr. Ubirajara Ferreira - da Faculdade de Ciências Médicas e Hospital das Clínicas/UNICAMP - em relação ao estágio da doença, esses nomogramas revelam um certo pessimismo, quando observada a realidade clínica dos pacientes brasileiros. Além disso, ao avaliar o desempenho do SBRF que construímos, constatou a seme-

lhança nas respostas.

No próximo capítulo, discutiremos outras mudanças que fizemos nessa modelagem *fuzzy*, a fim de obter resultados mais otimistas e, dessa forma, aproximá-los um pouco mais da realidade vivida pelos pacientes.

Capítulo 4

Buscando um melhor desempenho do Modelo *Fuzzy*

4.1 Introdução

O processo de tomada de decisões médicas é complexo e exige que se considere a qualidade dos dados clínicos disponíveis, bem como a efetividade de vários tratamentos, os custos para o paciente, entre outros. Diante disso, torna-se evidente a importância da avaliação, de forma objetiva, do poder discriminatório dos testes de diagnóstico, em relação à doença.

O sistema baseado em regras *fuzzy* - SBRF - descrito no capítulo anterior, tem como objetivo prever o estágio patológico do câncer de próstata. De acordo com a evolução da doença, esse estágio pode ser: localizado na glândula, localmente avançado e metastático, com envolvimento das vesículas seminais e/ou linfonodos pélvicos. Portanto, o SBRF é um teste de diagnóstico, que discrimina os estágios da doença.

Como já mencionamos, foram feitas simulações do modelo e os resultados encontrados mostraram-se um tanto quanto pessimistas - assim como são

consideradas as probabilidades constantes nos nomogramas - quando observada, pelos especialistas, a realidade clínica dos pacientes brasileiros.

O objetivo desse capítulo é implementar outras modificações no modelo *fuzzy* que descrevemos, visando melhorar sua performance, isto é, obter resultados mais otimistas e, portanto, mais realísticos.

4.2 As variáveis de Entrada e Saída do Sistema

No modelo *fuzzy* que construímos, as variáveis de entrada - *Estado Clínico*, *Nível de PSA* e *Escore de Gleason* e a variável de saída - *Estágio da Doença* - foram consideradas como variáveis lingüísticas e seus valores como conjuntos *fuzzy*, em seus respectivos domínios.

Os termos lingüísticos atribuídos à estas variáveis foram mantidos. Assim, o *Estado Clínico* foi classificado em **T1**, **T2a**, **T2b**, **T2c** e **T3a**. O *Nível de PSA* foi considerado **Normal** - até 4 ng/ml - **Levemente Elevado** - entre 4 e 10 ng/ml - **Moderadamente Elevado** - entre 10 e 20 ng/ml - e **Altamente Elevado** - acima de 20 ng/ml. O *Escore de Gleason* foi denominado **Bem Diferenciado** (graus 2, 3 e 4), **Médio Diferenciado** (grau 5), **Médio-Baixo Diferenciado** (grau 6), **Pouco Diferenciado** (grau 7) e **Indiferenciado** (graus 8, 9 e 10). Finalmente, o *Estágio da Doença* foi classificado em **Localizado**, **Localmente Avançado** e **Metastático**.

As funções de pertinência, para as variáveis de entrada, estão representadas nas figuras 3.3, 3.4 e 3.5.

A variável de saída *Estágio da Doença* é qualitativa, e como já comentamos, foi escolhida uma escala de 0 a 1 para indicar a extensão do câncer. Inicialmente, numa espécie de “chute”, construímos suas funções de pertinência na forma triangular (ver figura 3.6). Posteriormente, essas funções foram

adaptadas e discutiremos isso mais adiante.

Fizemos uma primeira modificação nas funções de pertinência triangulares, que descrevem os conjuntos *fuzzy* da saída do sistema. Mantivemos a forma triangular, porém considerando uma intersecção entre os três estágios, pois pode ocorrer numa situação real. A figura 4.1 mostra a representação dessas funções.

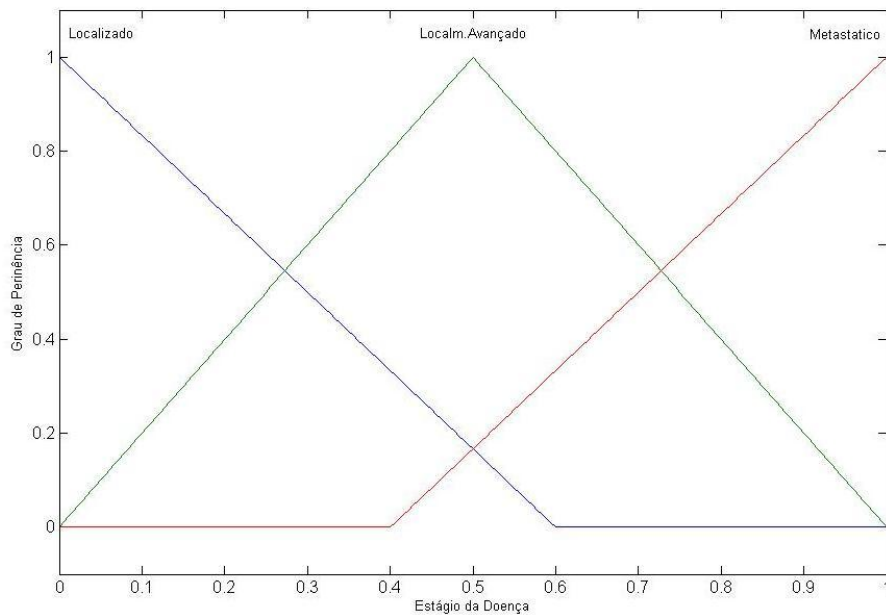


Figura 4.1: Funções de pertinência dos conjuntos *fuzzy*, assumidos pela variável lingüística *Estágio da Doença*.

4.3 A Base de Regras e algumas Simulações

As informações contidas nos nomogramas - de Stephenson e Kattan (2006) [49] - foram utilizadas na elaboração da base de regras *fuzzy*. Fizemos todas as diferentes combinações entre estado clínico, nível de PSA e grau de Gleason,

levando em conta os termos lingüísticos, atribuídos à cada uma das variáveis, e as probabilidades para os diferentes estágios do câncer de próstata. Vimos na tabela 3.1 algumas das regras construídas.

Essas regras foram ponderadas quando a diferença de probabilidade, de estar em dois estágios distintos, foi menor do que 20%. Caso contrário, para definir a regra foi considerada a maior das probabilidades. Na seção 3.3 descrevemos esse procedimento.

Na tentativa de aperfeiçoar o desempenho do modelo, decidimos reconstruir a base de regras considerando todas as probabilidades - presentes nos nomogramas - para os diferentes estágios da doença, referentes à todas as combinações entre as variáveis de entrada. A seguir apresentamos um exemplo do novo procedimento adotado.

Para um paciente com estado clínico T2a, nível de PSA < 4 e escore de Gleason 6, constam nos nomogramas as seguintes probabilidades: 64% para câncer localizado, 34% para doença localmente avançada e 2% para doença metastática. Com essas informações construímos as regras números 31, 32 e 33 da tabela 4.1.

- 31) “Se *Estado Clínico* é T2a, *Nível de PSA* é Normal e *Escore de Gleason* é Médio-Baixo Diferenciado, então *Estágio da Doença* é Localizado (0,64)”.
- 32) “Se *Estado Clínico* é T2a, *Nível de PSA* é Normal e *Escore de Gleason* é Médio-Baixo Diferenciado, então *Estágio da Doença* é Localmente Avançado (0,34)”.
- 33) “Se *Estado Clínico* é T2a, *Nível de PSA* é Normal e *Escore de Gleason* é Médio-Baixo Diferenciado, então *Estágio da Doença* é Metastático (0,02)”.

N	EST.CLÍN.	PSA	GLEASON	ESTÁGIO	PESO
01	T1	Normal	Bem Dif.	Localizado	0,80
02	T1	Normal	Bem Dif.	Loc. Avan.	0,19
03	T1	Normal	Bem Dif.	Metastático	0,01
31	T2a	Normal	Méd-B Dif.	Localizado	0,64
32	T2a	Normal	Méd-B Dif.	Loc. Avan.	0,34
33	T2a	Normal	Méd-B. Dif.	Metastático	0,02
124	T2c	Lev. Elev.	Pouco Dif.	Localizado	0,25
125	T2c	Lev. Elev.	Pouco Dif.	Loc. Avan.	0,48
126	T2c	Lev. Elev.	Pouco Dif.	Metastático	0,27

Tabela 4.1: Algumas das 285 regras da nova Base de Regras. A última coluna é formada pelos pesos utilizados para ponderá-las.

As probabilidades de ocorrência de cada estágio da doença constituíram os pesos (graus de ativação), que ponderaram as regras, da nova base de regras.

Definida essa nova base de regras, mantivemos o processo de inferência pelo Método de Mamdani e a defuzzificação pelo Método do Centro de Gravidade.

Fizemos algumas simulações desse novo modelo, utilizando inicialmente as funções de pertinência, da saída do sistema, na forma triangular (ver figura 4.1). Os dados pré-cirúrgicos considerados são de pacientes do Hospital das Clínicas da UNICAMP e os resultados encontram-se na tabela 4.2.

Assim, por exemplo, para paciente com tumor palpável (T2c), nível de PSA moderadamente elevado (18,2) e escore de Gleason pouco diferenciado (7), os graus de pertinência aos conjuntos *fuzzy* localizado, localmente avançado e metastático são, respectivamente, 0,10, 0,92 e 0,23. Esse resultado

Est.Clin.	PSA	Gleason	Saída	Localiz.	Loc. Avan.	Metast.
T1	12	5	0,399	0,33	0,80	0,00
T3a	23,7	7	0,632	0,00	0,74	0,39
T2c	18,2	7	0,538	0,10	0,92	0,23
T1	7,5	5	0,350	0,42	0,70	0,00
T2b	15,4	9	0,590	0,02	0,83	0,31

Tabela 4.2: Pertinência aos conjuntos *fuzzy* que representam os estágios da doença: Localizado na Próstata, Localmente Avançado e Metastático.

indica que a possibilidade de que o câncer de próstata esteja localmente avançado é maior.

Se verificarmos os nomogramas, as probabilidades correspondentes às pertinências que mencionamos são, respectivamente, 0,15, 0,45 e 0,40. Podemos observar que a diferença de probabilidades entre os estágios localmente avançado e metastático é pequena.

Nesse caso, os resultados do SBRF mostraram-se ligeiramente menos pessimistas, no sentido de que a possibilidade de a doença estar localmente avançada foi bem maior, do que a de haver formação de metástases.

No entanto, se analisarmos os quadros clínicos dos pacientes e os respectivos resultados das simulações, utilizando as funções de pertinência da saída na forma triangular, veremos que ainda são respostas mais próximas das apresentadas nos nomogramas, do que da realidade clínica desses pacientes.

Diante disso, fizemos uma nova adaptação das funções de pertinência, que descrevem a saída do sistema, a partir das probabilidades de ocorrência de cada estágio. Esse é o assunto da próxima seção.

4.4 Adaptação das Funções de Pertinência

A probabilidade de um paciente, com determinado quadro clínico - estado clínico, nível de PSA e grau de Gleason - estar em um dos estágios do câncer de próstata: localizado, localmente avançado e metastático, encontra-se nos nomogramas de Stephenson e Kattan (2006) [49].

No capítulo anterior, mostramos uma forma de adaptar as funções de pertinência, que descrevem a saída do sistema baseado em regras *fuzzy* (SBRF), por meio de ajustes de curvas - método dos mínimos quadrados. Nessa seção, adotaremos um procedimento análogo, porém com algumas modificações.

Considerando uma amostra de indivíduos e combinando os diferentes estados clínicos, níveis de PSA e escores de Gleason, obtemos, pelo SBRF, um número real α , pertencente ao intervalo $[0, 1]$, para cada combinação. Como já mencionamos, esse número α encontrado representa uma subpopulação de pacientes, com mesmo quadro clínico e, portanto, possui uma distribuição de probabilidade (ver figura 3.7).

Para o estágio *Localizado*, consideramos a mesma curva que foi ajustada, anteriormente, às probabilidades de o paciente estar nesse estágio. Isso ocorreu pois, para esse estágio da doença, os resultados foram satisfatórios.

Uma curva foi ajustada aos pontos que representam a probabilidade, fornecida pelos nomogramas, de o paciente estar em cada um dos estágios, *Localmente Avançado* e *Metastático*. Essas funções ajustadas, para esses dois estágios, foram normalizadas e concentradas ($p = 3$).

Várias amostras de indivíduos, com diferentes quadros clínicos, foram simuladas e novas curvas foram ajustadas às probabilidades de ocorrência de cada um desses dois estágios. As funções obtidas foram normalizadas e concentradas ($p = 3$). O processo foi repetido até não haver diferença significativa entre a saída nova e a anterior.

Na figura 4.2 estão representadas as funções de pertinência aos conjuntos *fuzzy*, da saída do sistema, obtidas por meio dos novos ajustes de curvas.

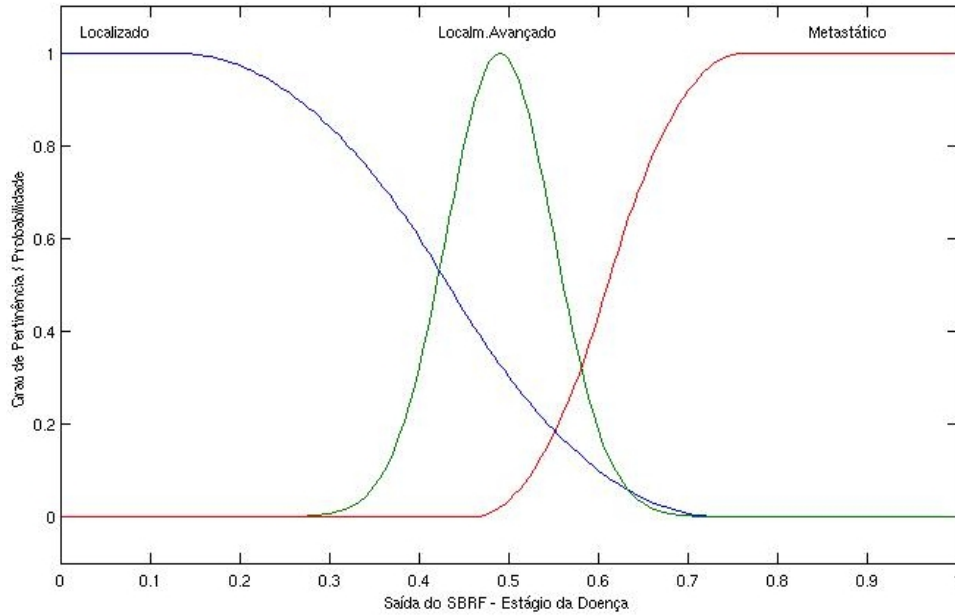


Figura 4.2: Funções de pertinência dos conjuntos *fuzzy*, assumidos pela variável lingüística *Estágio da Doença*.

4.5 Novas Simulações

Utilizando as funções de pertinência da saída do sistema, encontradas através dos novos ajustes de curvas, outras simulações de pacientes foram feitas, considerando os dados do Hospital das Clínicas da UNICAMP. As tabelas 4.3 e 4.4 mostram os resultados que obtivemos.

Seja, por exemplo, um paciente com estado clínico T2a, nível de PSA 4,6 e grau de Gleason 6. As possibilidades, determinadas pelo SBRF, de que o câncer de próstata esteja localizado, localmente avançado e metastático são, respectivamente, 0,59, 0,34 e 0. Transformando esses valores em probabili-

dades - via expressão (2.13) - obtemos 64%, 36% e 0%, respectivamente.

		SBRF		Kattan <i>et al</i>
Paciente	Estágio	Possib.	Probab.	Nomograma
Est. Clínico T2a	Localizado	0,59	64%	51%
PSA 4,6	Loc. Avançado	0,34	36%	44%
Gleason 6	Metastático	0	0%	5%
Est. Clínico T2b	Localizado	0,23	20%	13%
PSA 13,3	Loc. Avançado	0,81	71%	51%
Gleason 7	Metastático	0,10	9%	36%
Est. Clínico T2a	Localizado	0,45	37%	38%
PSA 11	Loc. Avançado	0,76	63%	52%
Gleason 6	Metastático	0	0%	9%
Est. Clínico T3a	Localizado	0,15	18%	3%
PSA 15,6	Loc. Avançado	0,43	51%	26%
Gleason 8	Metastático	0,26	31%	71%
Est. Clínico T1	Localizado	0,48	41%	49%
PSA 5,5	Loc. Avançado	0,68	59%	36%
Gleason 7	Metastático	0	0%	15%
Est. Clínico T2c	Localizado	0,28	22%	15%
PSA 11,3	Loc. Avançado	0,96	74%	45%
Gleason 7	Metastático	0,05	4%	40%
Est. Clínico T2a	Localizado	0,24	20%	5%
PSA 30	Loc. Avançado	0,85	72%	38%
Gleason 8	Metastático	0,09	8%	57%

Tabela 4.3: Comparação dos resultados, obtidos pelo novo SBRF, com as probabilidades de Kattan et al [49]. A coluna *Probab.* é a transformação da possibilidade em probabilidade, via equação (2.13).

		SBRF		Kattan <i>et al</i>
Paciente	Estágio	Possib.	Probab.	Nomograma
Est. Clínico T3a	Localizado	0,11	15%	6%
PSA 17,2	Loc. Avançado	0,32	43%	40%
Gleason 7	Metastático	0,32	42%	54%
Est. Clínico T2a	Localizado	0,34	26%	33%
PSA 6,98	Loc. Avançado	0,97	74%	52%
Gleason 7	Metastático	0,01	0%	14%
Est. Clínico T2a	Localizado	0,57	63%	51%
PSA 4,47	Loc. Avançado	0,34	37%	44%
Gleason 6	Metastático	0	0%	5%
Est. Clínico T1	Localizado	0,46	42%	33%
PSA 11,9	Loc. Avançado	0,64	58%	49%
Gleason 6	Metastático	0	0%	17%
Est. Clínico T2c	Localizado	0,19	18%	15%
PSA 14,5	Loc. Avançado	0,72	69%	45%
Gleason 7	Metastático	0,14	13%	40%

Tabela 4.4: Comparação dos resultados, obtidos pelo novo SBRF, com as probabilidades de Kattan et al [49].

Já nos nomogramas, referentes a esse quadro clínico, constam respectivamente, as seguintes probabilidades: 51%, 44% e 5%.

Podemos notar que as probabilidades, para o estágio localmente avançado, foram idênticas. Entretanto, a possibilidade de a doença estar localizada na glândula se elevou, enquanto que o risco de haver metástase foi reduzido à 0.

Para um outro paciente com estado clínico T2a, nível de PSA 30 e escore de Gleason 8, as possibilidades, segundo o SBRF, são 0,24 (doença localizada), 0,85 (doença localmente avançada) e 0,06 (metástase). Nesse caso,

convertidos à probabilidades, esses valores são: 20%, 72% e 8%, respectivamente. Nos nomogramas encontramos 5%, 38% e 57%.

Os resultados do SBRF, para este paciente, foram mais otimistas, na medida em que, para os estágios localizado e localmente avançado, forneceu possibilidades maiores, ao contrário do que acontece nos nomogramas. Em contrapartida, houve uma significativa redução da possibilidade de a doença estar espalhada pelo organismo.

De modo geral, podemos afirmar que após a implementação de todas as modificações, descritas nesse capítulo, os resultados das simulações se mostraram mais otimistas, em relação ao estágio do câncer de próstata, do que as probabilidades presentes nos nomogramas de Kattan *et al* [49].

4.6 Considerações Finais

No capítulo anterior, apresentamos uma alternativa para a predição do estágio patológico do câncer de próstata. Trata-se de um modelo matemático que, apoiado na Teoria de Conjuntos *Fuzzy*, inclui as incertezas e subjetividades envolvidas no problema.

Apesar de todas as vantagens que a modelagem *fuzzy* possui, para lidar com incertezas, como já discutimos anteriormente, nossos primeiros resultados foram semelhantes às probabilidades calculadas, por Kattan *et al*, através de métodos estatísticos tradicionais.

Na avaliação do especialista que nos auxiliou, prof. Dr. Ubirajara Ferreira - FCM e HC da UNICAMP - os nomogramas mostram-se pessimistas, em relação ao estágio da doença, quando observada a realidade clínica dos pacientes.

Propusemos então, neste capítulo, algumas modificações na modelagem

fuzzy, visando melhorar a performance do modelo. Efetuadas tais mudanças, os resultados das simulações se tornaram mais otimistas, do que os nomogramas, em relação ao estadiamento patológico da doença.

A avaliação do desempenho do novo modelo, feita pelo especialista que citamos, constatou que as respostas que obtivemos, no que diz respeito ao estágio do câncer de próstata, estão mais otimistas. Portanto, foram consideradas mais coerentes com a realidade clínica enfrentada pelos pacientes.

Capítulo 5

Conclusão e Trabalhos Futuros

5.1 Conclusão

O câncer de próstata também se sobressai como um importante problema de saúde pública no mundo. No Brasil, é o tumor maligno mais freqüente entre os homens. Se for descoberto precocemente, quando ainda não ultrapassou os limites da glândula, é potencialmente curável, por meio de cirurgia ou braquiterapia.

A predição correta, do estadiamento patológico do câncer de próstata, é fundamental, para que o especialista possa indicar a terapia adequada. O médico dispõe das informações dadas pelo exame clínico (toque retal), pela dosagem sérica do PSA e o grau de diferenciação do tumor (escore de Gleason). Quando necessário, utiliza também as probabilidades, constantes nos nomogramas de Kattan *et al* [49], que combinam esses dados.

Ao longo dessa dissertação, apresentamos uma opção para prever o estágio patológico da doença. O modelo que descrevemos consiste num sistema baseado em regras *fuzzy*, que combina os dados pré-cirúrgicos - estado clínico, nível de PSA e escore de Gleason - levando em conta um conjunto de

regras, de natureza lingüística, elaborado a partir das informações presentes nos nomogramas.

A saída do sistema é um número real, entre 0 e 1, e a ele corresponde um grau de pertinência aos estágios: doença localizada, localmente avançada e metastática.

Fizemos atualizações e algumas modificações, no modelo de Castanho (2005). Nas simulações utilizamos dados reais, de alguns pacientes do Hospital das Clínicas da UNICAMP, que foram diagnosticados como portadores da doença, no período de 1992 à 2006.

Além da relativa facilidade de interpretação e reprodução dessa modelagem, quando comparada aos métodos estatísticos tradicionais, enumeramos outras vantagens de se trabalhar com a teoria de Conjuntos *Fuzzy*. Esse modelo é mais abrangente, no sentido de que o especialista pode avaliar, numa escala contínua, a variável *Estado Clínico* e a saída *Estágio da Doença*. Isso permite uma transição gradual entre os estágios e não uma transição abrupta, como ocorre se observarmos os nomogramas.

Contudo, os resultados que obtivemos no capítulo 3, mostraram-se semelhantes aos encontrados nos nomogramas. De acordo com o especialista que contactamos, citado nos capítulos anteriores, quando observada a realidade dos pacientes, na clínica, esses resultados apresentam-se pessimistas, em relação ao estágio do câncer de próstata.

Diante disso, propusemos, no último capítulo, novas mudanças, na base de regras *fuzzy* e na adaptação das funções de pertinência, que representam a saída do sistema.

Implementadas essas novas modificações, foi possível melhorar o desempenho do modelo. Foram feitas novas simulações e os resultados encontrados mostraram-se mais otimistas, em relação ao estágio da doença, do que os

resultados anteriores e os nomogramas. Portanto, de acordo com o especialista, as respostas desse novo modelo estão mais próximas da realidade clínica enfrentada pelos pacientes.

5.2 Trabalhos Futuros

Uma proposta de trabalho futuro é avaliar o modelo que descrevemos nessa dissertação, através da Curva ROC (Receiver Operating Characteristics). Para isso será necessário utilizar dados de pacientes cujo real estágio patológico do câncer de próstata é conhecido.

Outro projeto futuro é desenvolver uma interface gráfica para o modelo *fuzzy*, facilitando o seu acesso por parte dos especialistas.

Uma outra proposta consiste na construção de um modelo de predição do estágio patológico do câncer de próstata, considerando o PSA ajustado à idade. Estudos recentes, como os de Chun e colaboradores (2006) [15], determinaram valores médios de PSA, para diferentes faixas etárias.

A modelagem, levando em conta esses valores médios de PSA, provavelmente vai exigir a construção de mais de um sistema baseado em regras *fuzzy*. Além disso, será necessário elaborar todas as regras, juntamente com um especialista, tendo em vista a falta de nomogramas que combinem o PSA ajustado à idade e as demais variáveis.

Pode-se também tentar aliar a esse sistema construído, para predizer o estágio da doença, um novo sistema *fuzzy* que forneça informações relacionadas à sobrevida dos pacientes, dependendo da terapia escolhida para o seu tratamento. Ou ainda, pode-se investigar a evolução do estágio, no decorrer do tempo.

No entanto, para concretizar um trabalho como esse, serão necessários

novos estudos, os quais deverão abordar sistemas dinâmicos e, possivelmente, redes neurais artificiais.

Referências Bibliográficas

- [1] K. P. Adlassnig, *A fuzzy logical model of computer-assisted medical diagnosis*. Methods of Information in Medicine **19**, no. 3, pp. 141-148, 1980.
- [2] ———, *Fuzzy set theory in medical diagnosis*. IEEE Transactions on Systems, Man and Cybernetics **16**, no. 2, pp. 260-265, 1986.
- [3] M. Amendola, A. L. Souza, L. C. Barros, *Manual do uso da teoria dos conjuntos fuzzy no MATLAB 6.5*. FEAGRI e IMECC/UNICAMP, 2005.
- [4] M. C. R. Alavanja, C. Samanic, M. Dosemeci, J. Lubin, R. Tarone, C. F. Lynch, C. Knott, K. Thomas, J. A. Hoppin, J. Barker, J. Coble, D. Sandler, A. Blair, *Use of Agricultural Pesticides and Prostate Cancer Risk in the Agricultural Health Study Cohort*. American Journal of Epidemiology **157**, no. 9, pp. 800-814, 2003.
- [5] L. C. Barros, R. C. Bassanezi, *Tópicos de Lógica Fuzzy e Biomatemática*. Coleção IMECC Textos Didáticos vol. 5, UNICAMP/IMECC. Campinas - SP, 2006.
- [6] ———, *Introdução à Teoria Fuzzy: Aplicações em Biomatemática*. Congresso Latino Americano de Biomatemática, X ALAB - V ELAEM, IMECC/UNICAMP. Campinas - SP, 2001.

- [7] L. C. Barros, M. B. F. Leite, R. C. Bassanezi, *The SI epidemiological models with a fuzzy transmission parameter*. Computers and Mathematics with Applications **45**, pp. 1619-1628, 2003.
- [8] A. Billis, *Patologia Cirúrgica da Próstata*. Impressão Digital do Brasil Gráfica e Editora. Campinas - SP, 2003.
- [9] M. Bilobrovec, *Sistema Especialista em Lógica Fuzzy para o Controle, Gerenciamento e Manutenção da qualidade em Processo de Aeração de Grãos*. Dissertação de Mestrado, Programa de Pós-Graduação em Engenharia de Produção - Universidade Tecnológica Federal do Paraná. Ponta Grossa, 2005.
- [10] H. B. Carter, A. W. Partin, *Diagnosis and Staging of Prostate Cancer*. In. Campbell's Urology, W. B. Saunders Company, pp. 2519-2537, 1997.
- [11] M. J. P. Castanho, *Construção e Avaliação de um Modelo Matemático para prever a evolução do Câncer de Próstata e descrever seu crescimento utilizando a Teoria dos Conjuntos Fuzzy*. Tese de Doutorado, FEEC - Universidade Estadual de Campinas. Campinas - SP, 2005.
- [12] W. J. Catalona, D. S. Smith, D. K. Ornstein, *Prostate cancer detection in men with serum PSA concentrations of 2.6 to 4.0 ng/ml and benign prostate examination*. JAMA **277**, no. 18, pp. 1452-1455, 1997.
- [13] V. Cavalcanti, *Braquiterapia ainda é cara, mas ganha eficiência*. Agência Estado, 2006.
- [14] F. K. -H. Chun, P. I. Karakiewicz, A. Briganti, A. Gallina, M. W. Kattan, F. Montorsi, H. Huland, M. Graefen, *Prostate Cancer Nomograms: An Update*. European Association of Urology, pp. 1-13, 2006.

- [15] F. K. -H. Chun, P. Perrotte, A. Briganti, S. Benayoun, T. Lebeau, A. Ramirez, D. J. Lewinshtein, L. Valiquette, J. P. Guay, P. I. Karakiewicz, *Prostate specific-antigen distribution in asymptomatic Canadian men with no clinical evidence of prostate cancer*. British Journal of Urology **98**, pp. 50-53, 2006.
- [16] M. R. Civanlar, H. J. Trussel, *Constructing membership functions using statistical data*. In. Fuzzy Sets and Systems **18**, pp. 01-13, 1986.
- [17] J. H. Cohen, A. R. Kristal, J. L. Stanford, *Fruit and Vegetable Intakes and Prostate Cancer Risk*. Journal of the National Cancer Institute **92**, no. 1, pp. 61-68, 2000.
- [18] L. I. Dini, W. J. Koff, *Perfil do Câncer de Próstata no Hospital de Clínicas de Porto Alegre*. Revista da Associação Médica Brasileira **52**, no. 1, pp. 28-31, 2006.
- [19] D. Dubois, H. Prade, *Fuzzy Sets and Systems: Theory and Applications*. Academic Press, INC, vol. 144, 1980.
- [20] J. I. Epstein, *Pathology of adenocarcinoma of the prostate*. In. Campbell's Urology, W. B. Saunders Company, pp. 2497-2517, 1997.
- [21] W. R. Fair, W. D. Heston, C. Cordon-Cardo, *An overview of cancer biology*. In. Campbell's Urology, W. B. Saunders Company, pp. 2259-2282, 1997.
- [22] U. Ferreira, A. C. Nardi, *Câncer de Próstata*. In. N. R. J. Netto, *Urologia Prática*. Atheneu, 4 ed., São Paulo, 1999.
- [23] P. H. Gann, M. Han, *The Natural History of Clinically Localized Prostate Cancer*. JAMA **293**, no. 17, pp. 2149-2151, 2005.

- [24] E. Giovannucci, E. B. Rimm, Y. Liu, M. J. Stampfer, W. C. Willett, *A Prospective Study of Tomato Products, Lycopene, and Prostate Cancer Risk*. Journal of the National Cancer Institute **94**, no.5, pp. 391-398, 2002.
- [25] M. R. Guerra, C. V. M. Gallo, G. A. S. Mendonça, *Riscos de câncer no Brasil: tendências e estudos epidemiológicos mais recentes*. Revista Brasileira de Cancerologia **51**, no. 03, pp. 227-234, 2005.
- [26] Hospital A. C. Camargo. <http://www.hcanc.org.br/site.php> - último acesso em fevereiro de 2007.
- [27] INCA, *Instituto Nacional do Câncer*. www.inca.gov.br - último acesso em fevereiro de 2007.
- [28] R. S. M. Jafelice, *Modelagem Fuzzy para a dinâmica de transferência de soropositivos para HIV em doença plenamente manifesta*. Tese de Doutorado, FEEC - Universidade Estadual de Campinas. Campinas - SP, 2003.
- [29] D. Kalmanson, H. F. Stegall, *Cardiovascular investigations and fuzzy sets theory*. The American Journal of Cardiology **35**, pp. 80-84, 1975.
- [30] G. J. Klir, U. H. St. Clair, B. Yuan, *Fuzzy Set Theory: Foundations and Applications*. Prentice Hall PTR. USA, 1997.
- [31] E. Massad, R. Menezes, P. Silveira, N. Ortega, *Métodos Quantitativos em Medicina*. Editora Manole. São Paulo - SP, 2004.
- [32] M. Mizumoto, *Fuzzy Logic and Intelligent Systems*. In. Improvement of Fuzzy Control Methods **1**. Kluwer Academic Publishers, pp. 01-16, USA, 1995.

- [33] D. G. Moreno, *Câncer de Próstata*. In. Informações para Pacientes. Liga Urológica Acadêmica - LUA, Universidade Federal de São Paulo. São Paulo - SP, 2006.
- [34] H. Neuwirth, *Gleason Score*. In. Marin Urology. www.marinurology.com - último acesso em outubro de 2006.
- [35] J. E. Oesterling, S. J. Jacobsen, G. G. Klee, K. Petterson, T. Piironem, P. A. Abrahamsson, *Free, complexed and total serum prostate specific antigen: the establishment of appropriate reference ranges for their concentrations and ratios*. Journal Urology **154**, no. 3, pp. 1090-1095, 1995.
- [36] N. R. S. Ortega, *Aplicação da teoria de conjuntos fuzzy a problemas de biomedicina*. Tese de Doutorado, Instituto de Física da Universidade de São Paulo. São Paulo - SP, 2001.
- [37] A. W. Partin, M. W. Kattan, E. Subong, P. C. Walsh, K. Wojno, J. Oesterling, P. Scardino, J. D. Pearson, *Combination of prostate-specific antigen, clinical stage, and gleason score to predict pathological stage of localized prostate cancer*. JAMA **277**, no. 18, pp. 1445-1451, 1997.
- [38] A. W. Partin, L. A. Mangold, D. M. Lamm, P. C. Walsh, J. I. Epstein, J. D. Pearson, *Contemporary update of prostate cancer staging nomograms (Partin Tables) for the new millennium*. Journal Urology **58**, no. 6, pp. 843-848, 2001.
- [39] W. Pedrycz, F. Gomide, *An Introduction to Fuzzy Sets: Analysis and Design*. MIT Press. Cambridge, USA, 1998.
- [40] K. J. Pienta, *Etiology, Epidemiology, and Prevention of carcinoma of the prostate*. In. Campbell's Urology, W. B. Saunders Company, pp. 2489-2496, 1997.

- [41] M. Pradal, A. M. Menezes, R. D. Dio, J. C. Barbério, *Marcadores Tumoraes - Breve Revisão*. Revista da Sociedade Brasileira de Cancerologia, no. 5, pp. 20-29, 1999.
- [42] G. J. C. Santos, *Lógica Fuzzy*. Monografia, Departamento de Ciências Exatas e Tecnológicas - Universidade Estadual de Santa Cruz. Ilhéus - BA, 2003.
- [43] A. D. Sasse, *E-CÂNCER - Câncer e Oncologia*. www.andre.sasse.com - último acesso em fevereiro de 2007.
- [44] I. S. Shaw, M. G. Simões, *Controle e Modelagem Fuzzy*. Edgard Blücher Ltda, 1a ed. São Paulo - SP, 1999.
- [45] L. H. Sobin, Ch. Wittekind, *TNM: Classificação dos Tumores Malignos*. Tradução de Ana Lúcia Amaral Eisenberg. 6a ed, Instituto Nacional do Câncer/Ministério da Saúde. Rio de Janeiro, 2004.
- [46] Sociedade Brasileira de Urologia, *Biópsia de Próstata*. Projeto Diretrizes - Associação Médica Brasileira e Conselho Federal de Medicina, 2006.
- [47] C. A. Souto, G. N. Fonseca, G. F. Carvalhal, H. S. Barata, J. C. S. Souto, M. Berger, *Marcadores Tumoraes em Câncer de Próstata*. Official Journal of the Brazilian Society of Urology **30**, pp. 09-17, 2004.
- [48] M. Srougi, *Adenocarcinoma da Próstata*. Urologia, pp. 356-372, 1999.
- [49] A. J. Stephenson, M. W. Kattan, *Nomograms for prostate cancer*. Journal Urological Oncology **98**, pp. 39-46, 2006.
- [50] L. Wang, H. Hricak, M. W. Kattan, H. -N. Chen, P. T. Scardino, K. Ku-roiwa, *Prediction of Organ-confined Prostate Cancer: Incremental Value*

- of MR Imaging and MR Spectroscopic Imaging to Staging Nomograms.* Journal Radiology **238**, no. 2, pp. 597-603, 2006.
- [51] WHO, *World Health Organization*. www.who.int/cancer/en/ - último acesso em março de 2007.
- [52] H. Zhu, K. A. Roehl, J. A. V. Antenor, W. J. Catalona, *Biopsy of Men with PSA level of 2.6 to 4.0 ng/ml associated with favorable pathologic features and PSA progression rate: A Preliminary Analysis*. Journal Urology **66**, pp. 547-551, 2005.
- [53] H. J. Zimmermann, *Fuzzy Sets, Decision Making and Expert Systems*. Kluwer Academic Publishers, 1993.