



UNIVERSIDADE ESTADUAL DE  
CAMPINAS

Instituto de Matemática, Estatística e  
Computação Científica

CRISTINA SACILOTTO

**Uso da teoria de conjuntos *Fuzzy* para análise  
prognóstica do câncer de rim**

Campinas

2017

Cristina Sacilotto

**Uso da teoria de conjuntos *Fuzzy* para análise  
prognóstica do câncer de rim**

Dissertação apresentada ao Instituto de Matemática, Estatística e Computação Científica da Universidade Estadual de Campinas como parte dos requisitos exigidos para a obtenção do título de Mestra em Matemática Aplicada.

Orientador: Laércio Luís Vendite

Coorientadora: Graciele Paraguaia Silveira

Este exemplar corresponde à versão final da Dissertação defendida pela aluna Cristina Sacilotto e orientada pelo Prof. Dr. Laércio Luís Vendite.

Campinas

2017

**Agência(s) de fomento e nº(s) de processo(s):** CAPES

Ficha catalográfica  
Universidade Estadual de Campinas  
Biblioteca do Instituto de Matemática, Estatística e Computação Científica  
Ana Regina Machado - CRB 8/5467

Sa14u Sacilotto, Cristina, 1990-  
Uso da teoria de conjuntos fuzzy para análise prognóstica do câncer de rim / Cristina Sacilotto. – Campinas, SP : [s.n.], 2017.

Orientador: Laércio Luis Vendite.

Coorientador: Graciele Paraguaia Silveira.

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

1. Rins - Câncer. 2. Fuhrman, Classificação de. 3. Lógica fuzzy. 4. Modelos matemáticos. I. Vendite, Laércio Luis, 1954-. II. Silveira, Graciele Paraguaia, 1982-. III. Universidade Estadual de Campinas. Instituto de Matemática, Estatística e Computação Científica. IV. Título.

Informações para Biblioteca Digital

**Título em outro idioma:** Use of fuzzy set theory for prognostic analysis kidney cancer

**Palavras-chave em inglês:**

kidneys - Cancer

Fuhrman classification

Fuzzy logic

Mathematical models

**Área de concentração:** Matemática Aplicada

**Titulação:** Mestra em Matemática Aplicada

**Banca examinadora:**

Laércio Luis Vendite [Orientador]

Estevão Esmi Laureano

Ubirajara Ferreira

**Data de defesa:** 31-07-2017

**Programa de Pós-Graduação:** Matemática Aplicada

**Dissertação de Mestrado defendida em 31 de julho de 2017 e aprovada  
pela banca examinadora composta pelos Profs. Drs.**

**Prof(a). Dr(a). LAÉRCIO LUÍS VENDITE**

**Prof(a). Dr(a). ESTEVÃO ESMI LAUREANO**

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

As respectivas assinaturas dos membros encontram-se na Ata de defesa

*À Deus e aos meus queridos pais*

# Agradecimentos

À DEUS, por permitir realizar este trabalho e por estar sempre ao meu lado, iluminando meus caminhos.

Aos meus queridos pais Neusa e Luiz, pelo amor incondicional, exemplos de superação e persistência extraordinários, pela educação que me proporcionaram, pela dedicação e todo o esforço prestados, pelo apoio em todas as etapas de minha vida, pela confiança que sempre tiveram em mim e pelo incentivo em lutar sempre.

Ao professor Laércio Luis Vendite, pelo auxílio, pela paciência e orientação durante a realização deste trabalho.

À minha coorientadora professora Graciele P. Silveira, pela orientação, pela disponibilidade com que sempre me atendeu nas inúmeras vezes que necessitei de sua ajuda e pelas sugestões tão úteis e valiosas para a concretização desta dissertação.

Ao Dr. Ubirajara Ferreira, por ter proposto um tema tão interessante e relevante, pelas instruções durante a qualificação, pela atenção e pelo tempo que dedicou para discutir partes deste trabalho.

Ao professor Sidney Ragazzi, por ter nos ajudado com a análise estatística que foi muito importante para o trabalho.

Aos médicos Dr. Athanase Billis e Dra. Larissa Eloy, pela disposição sempre presente em sanar minhas dúvidas nas diversas vezes que precisei, pelas dicas e auxílio na análise dos resultados.

Aos médicos Dr. Wagner Eduardo Matheus e Dr. Eduardo Azevedo, pela forma gentil com que me atenderam, pela disponibilidade e pelos esclarecimentos prestados.

Aos funcionários do SAM do Hospital das Clínicas da UNICAMP, pela simpatia em atender e compreensão ao disponibilizar os prontuários.

Aos familiares, pelo apoio e incentivo durante a caminhada.

À minha avó, pelos conselhos e carinho ao longo da vida.

Ao Ronaldo por todo o seu amor, atenção e dedicação durante esse tempo juntos.

À minha amiga e prima Taís, por sua ajuda e companhia tão preciosas e de fundamental importância.

Aos meus amigos de graduação Givanildo, Marcia, Vinicius, Raphaela, Pablo e José Renato e de pós-graduação, Aline, Vinicius, Lisbeth, Pricila, Fernanda, Fidelis e Daniel

pela verdadeira amizade, pelos conselhos, dicas e ajuda das diferentes formas possíveis, pelos incentivos, pelo carinho e preocupação. Esse auxílio contribuiu decisivamente para a realização deste trabalho.

À CAPES, pelo apoio financeiro.

À todos que, direta ou indiretamente, contribuíram para a execução deste trabalho.

# Resumo

As neoplasias renais malignas têm importante significado clínico e são responsáveis por cerca de 2% dos tumores malignos em humanos sendo, mais frequentes os carcinomas de células renais (CCR). A discussão e compreensão dos fatores prognósticos do CCR é fundamental para estabelecer-se uma abordagem na condução desses tumores. O entendimento dos avanços no diagnóstico e tratamento bem como qual o impacto que podem gerar na sobrevida dos pacientes são pontos relevantes para a definição prognóstica em indivíduos com a doença.

Nesta dissertação estudamos a construção de dois modelos matemáticos, que foram desenvolvidos para prever o risco de desenvolvimento de metástase e o risco de morte pelos pacientes que apresentam a doença com o subtipo mais comum, chamado convencional (células claras).

O sistema de graduação do câncer renal, conhecido e utilizado mundialmente, chamado Classificação de **Fuhrman** consiste na diferenciabilidade do núcleo celular de uma célula cancerosa ao compará-la com uma sadia.

A intenção é analisar a relação entre determinados fatores prognósticos (principalmente a graduação de Fuhrman) e a sobrevida desses pacientes, tendo em vista que constatou-se, em casos reais, pacientes com graus de Fuhrman baixos (considerados satisfatórios) e apresentaram um prognóstico desfavorável.

Os modelos consistem em sistemas baseados em regras fuzzy. O modelo cuja saída é “risco de metástase” combina os dados pós-cirúrgicos - grau de Fuhrman, estadiamento, presença de necrose e tamanho do tumor. Já o modelo com saída “risco de morte” condiz os dados do primeiro modelo, com a troca, apenas, do fator tamanho do tumor pela presença de metástase. Ambos valem-se de um conjunto de regras, de natureza linguística. As variáveis de entrada foram escolhidas com base em referências na área, em consulta a especialistas e análise feita estatisticamente com dados dos pacientes do Hospital das Clínicas/UNICAMP.

Com dados colhidos de pacientes do Hospital foram feitas simulações a fim de verificar se os modelos são compatíveis com a realidade. Os resultados obtidos foram satisfatórios e se aproximaram dos casos reais vividos pelos pacientes, no seguinte sentido nos casos que apresentaram metástase e/ou morte pela doença foram obtidos altos riscos de ocorrer esses fatos, enquanto que nos casos que não ocorreram, os riscos se mostraram baixos. Vale observar que a graduação de Fuhrman utilizada mundialmente possui escala de 1 a 4 graus; contudo, especialistas do Hospital das Clínicas sugeriram considerar a graduação

combinada, que é baseada no fato de que, em alguns casos, o tumor de rim apresenta mais de um grau na mesma lesão, essa classificação considera a soma dos dois graus presentes no tumor. Além disso, a análise estatística contribuiu para determinar a associação entre a graduação de Fuhrman e o prognóstico dos pacientes, considerando tanto o Fuhrman único como o Fuhrman combinado, tais associações se mostraram de baixa a moderada, comprovando que ambas as graduações de Fuhrman não são decisivas para avaliar a evolução da doença. Mais ainda, os resultados estatísticos indicaram as variáveis mais determinantes para o fenômeno estudado, que foram presença de necrose e estadiamento; conseqüentemente colaborou também com o ajuste dos pesos das regras nos modelos.

Com isso espera-se auxiliar o especialista com o estado clínico dos indivíduos para programar condutas que sejam adequadas a cada caso.

**Palavras-chave:** Câncer de Rim; Classificação Fuhrman; Modelo Matemático; Lógica Fuzzy.

# Abstract

Malignant renal neoplasms have important clinical significance and are responsible for about 2% of malignant tumors in humans being renal cell carcinomas (RCC). The discussion and understanding of the CCR's prognostic factors is fundamental to establishing a approach in the management of these tumors. The understanding of the diagnosis and treatment as well as the impact they can have on the survival of patients are relevant points for the prognostic in individuals with the disease.

In this dissertation we study the construction of two mathematical models, that were developed to predict the risk of developing metastases and the risk of death by patients with common subtype, called conventional (clear cells).

The renal cancer graduation system, known and used worldwide, called Fuhrman Classification consists of differentiability of the cell nucleus of a cancerous cell when comparing it with a healthy one.

The intention is to analyze the relationship between certain prognostic factors (mainly the graduation of Fuhrman) and the survival of these patients, considering that in real cases, patients with low degrees of Fuhrman considered satisfactory and presented an unfavorable prognosis.

The models consist of systems based on fuzzy rules. The model whose output is "risk of metastasis" combines post-surgical data - degree of Fuhrman, staging, presence of necrosis and tumor size. Already the model with output "risk of death" matches the data of the first model, with the exchange, only, of tumor size factor for the presence of metastasis. Both they use a set of rules of a linguistic nature. The variables were chosen based on references in the area, in consultation with the specialists and together with analysis done statistically with data of the patients of Hospital das Clínicas / UNICAMP.

With data collected from patients at the Hospital, simulations were performed to verify if the models are compatible with reality. The results obtained were satisfactory and approximated the actual cases experienced by the patients, in the following cases, in the cases that presented metastasis and / or death by the disease, high risks of these events were obtained, whereas in the cases that did not occur, the risks were shown low. It is worth noting that Fuhrman's graduation used worldwide has a scale of 1 to 4 degrees; however, experts from the Hospital das Clínicas suggested considering the combined grade, which is based on the fact that, in some cases, the kidney tumor has more than one degree in the same lesion, this classification considers the sum of the two degrees present in the

tumor. Moreover, the statistical analysis contributed to determine the association between Fuhrman's graduation and the patients' prognosis, considering both the Fuhrman single and the Fuhrman combined, such associations were shown to be low to moderate, proving that both Fuhrman's graduations are not decisive to evaluate the evolution of the disease. Moreover, the statistical results indicated the variables most determinant for the studied phenomenon, which were presence of necrosis and staging; consequently also collaborated with the adjustment of the weights of the rules in the models.

With this it is expected to assist the specialist with the clinical status of the individuals to conduct pipelines that are appropriate to each case.

**Keywords:** Kidneys - Cancer; Fuhrman classification; Mathematical model; Fuzzy Logic.

# Lista de ilustrações

Figura 1 – Anatomia do Rim. . . . .	21
Figura 2 – Carcinoma de células renais de células claras. A: Tumor de grau nuclear 1 com núcleos redondos ou uniformes; nucléolos não discerníveis ou ausentes. B: Carcinoma de grau nuclear 2 com contornos nucleares ligeiramente irregulares e nucléolos discretos (visíveis a 400x). C: A neoplasia nuclear de grau 3 possui núcleos grandes e irregulares com nucléolos visíveis a 100x. D: Carcinoma nuclear de grau 4 com núcleos bizarros e nucléolos grandes e prominentes. . . . .	26
Figura 3 – Representação gráfica da proposição p: Se x é A então y é B. . . . .	36
Figura 4 – Saídas parciais no método de Mamdani. . . . .	40
Figura 5 – Saída final no método de Mamdani. . . . .	40
Figura 6 – Estrutura básica de um sistema baseado em regras fuzzy. . . . .	44
Figura 7 – Gráfico para análise do Grau de Fuhrman: Tempo(meses) x Número de pacientes vivos após nefrectomia. . . . .	59
Figura 8 – Estrutura básica do sistema baseado em regras <i>fuzzy</i> . . . . .	60
Figura 9 – Funções de pertinência dos conjuntos <i>fuzzy</i> , assumidos pela variável linguística Estadiamento. . . . .	60
Figura 10 – Funções de pertinência dos conjuntos <i>fuzzy</i> , assumidos pela variável linguística Grau de Fuhrman. . . . .	61
Figura 11 – Funções de pertinência dos conjuntos <i>fuzzy</i> , assumidos pela variável linguística Presença de Necrose. . . . .	62
Figura 12 – Funções de pertinência dos conjuntos <i>fuzzy</i> , assumidos pela variável linguística Tamanho do Tumor. . . . .	62
Figura 13 – Funções de pertinência dos conjuntos <i>fuzzy</i> , assumidos pela variável linguística Risco de Metástase. . . . .	63
Figura 14 – Estrutura básica do sistema baseado em regras <i>fuzzy</i> . . . . .	67
Figura 15 – Funções de pertinência dos conjuntos <i>fuzzy</i> , assumidos pela variável linguística Metástase. . . . .	68
Figura 16 – Funções de pertinência dos conjuntos <i>fuzzy</i> , assumidos pela variável linguística Risco de Morte. . . . .	68

# Lista de tabelas

Tabela 1 – Graduação Histológica segundo Fuhrman et al (FUHRMAN; LASKY; LIMAS, 1982). . . . .	24
Tabela 2 – TNM 2009 - Classificação Clínica do CCR. . . . .	28
Tabela 3 – Distribuição de 148 carcinomas de células renais de acordo com a área de mais alto grau e estadiamento. . . . .	52
Tabela 4 – Distribuição de 148 carcinomas de células renais de acordo com o sistema de números combinados e estadiamento. . . . .	53
Tabela 5 – Algumas das 504 regras, da Base de Regras construída. Na 1 <sup>a</sup> coluna consta o $n^o$ da regra e, na última, o peso utilizado para ponderá-la. . .	65
Tabela 6 – Alguns resultados obtidos pelo SBRF. . . . .	71
Tabela 7 – Algumas das 504 regras, da Base de Regras construída. Na 1 <sup>a</sup> coluna consta o $n^o$ da regra e, na última, o peso utilizado para ponderá-la. . .	72
Tabela 8 – Alguns resultados obtidos pelo SBRF. . . . .	73

# Lista de abreviaturas e siglas

CCR	Carcinoma de Células Renais
TNM	Tumour Node Metastasis
SBRF	Sistema baseado em Regras Fuzzy
OMS	Organização Mundial da Saúde
WHO	World Health Organization
INCA	Instituto Nacional de Câncer
ONG	Organização não Governamental
SBU	Sociedade Brasileira de Urologia
KTS	Kang-Takagi-Sugeno
CADIAG	Computer-Assisted Diagnosis in Internal Medicine
AIDS	Acquired Immunodeficiency Syndrome
MATLAB	Matrix Laboratory
HC	Hospital das Clínicas
SAM	Serviço de Arquivo Médico
IMECC	Instituto de Matemática, Estatística e Computação Científica
UNICAMP	Universidade Estadual de Campinas
ESTAD	Estadiamento
NEC	Necrose
TAM	Tamanho
META	Metástase
POSSIB	Possibilidade
PROBAB	Probabilidade

# Sumário

	<b>Introdução</b>	<b>16</b>
<b>1</b>	<b>CÂNCER</b>	<b>19</b>
1.1	Rim	20
1.2	<b>Câncer de Rim - Caso de Células Claras</b>	<b>21</b>
1.2.1	Diagnóstico	22
1.2.2	Tipos de tumores	22
1.2.3	Fatores Prognósticos	23
<b>2</b>	<b>ASPECTOS DA TEORIA FUZZY</b>	<b>29</b>
2.1	<b>Conjuntos <i>Fuzzy</i></b>	<b>29</b>
2.2	<b>Lógica <i>Fuzzy</i></b>	<b>32</b>
2.2.1	Variáveis Linguísticas	32
2.2.2	Relações <i>Fuzzy</i>	33
2.2.3	Regras <i>Fuzzy</i>	34
2.2.4	Inferência	37
2.3	<b>Sistema baseado em Regras <i>Fuzzy</i></b>	<b>43</b>
2.4	<b>Probabilidade e Possibilidade</b>	<b>45</b>
<b>3</b>	<b>MODELAGEM MATEMÁTICA FUZZY PARA ANÁLISE DE RISCO DE METÁSTASE E RISCO DE MORTE</b>	<b>50</b>
3.1	<b>Abordagem Estatística: Grau de Fuhrman x Surgimento de Metástase e Demais Variáveis x Surgimento de Metástase</b>	<b>51</b>
3.2	<b>Modelagem <i>Fuzzy</i></b>	<b>58</b>
3.2.1	Primeiro Modelo	59
3.2.1.1	Simulações	65
3.2.2	Segundo Modelo	66
3.2.2.1	Simulações	69
<b>4</b>	<b>CONCLUSÃO E TRABALHOS FUTUROS</b>	<b>74</b>
4.1	<b>Conclusão</b>	<b>74</b>
4.2	<b>Trabalhos Futuros</b>	<b>75</b>
	<b>REFERÊNCIAS</b>	<b>77</b>

# Introdução

As neoplasias renais malignas têm importante significado clínico e são responsáveis por cerca de 2% dos tumores malignos em humanos sendo, mais frequentes os carcinomas de células renais (CCR) (SOUZA, 2008). Elas são desafiadoras devido a seu comportamento, pois trata-se de uma doença silenciosa que apresenta os sintomas tardiamente, em geral, quando a doença já encontra-se espalhada pelo organismo. Podemos citar entre eles: hematúria (sangramento na urina) e dores lombares e abdominais.

A discussão e compreensão dos fatores prognósticos do CCR é fundamental para estabelecer-se uma abordagem na condução desses tumores (POMPEO, 2006). Para tanto, analisaremos alguns deles e sua importância ao relacioná-los com a sobrevida dos pacientes. Um dos fatores com considerável valor prognóstico é o Sistema de Gradação de Fuhrman (FUHRMAN; LASKY; LIMAS, 1982) (BECKER, 2016) o qual classifica o padrão nuclear celular da neoplasia em quatro graus. Devido à constatação, mediante alguns casos reais, da incompatibilidade entre o grau de Fuhrman e o que se espera do prognóstico do paciente, surgiu o questionamento por parte da comunidade médica sobre essa classificação. O presente estudo visa fazer uma análise univariada desse critério, mas também estudar outros fatores (análise multivariada), de forma que, ao combiná-los com a classificação de Fuhrman podemos obter um resultado mais preciso para o prognóstico da neoplasia e verificar a relação entre o grau de Fuhrman e a sobrevida.

Os fatores prognósticos relacionados ao câncer de rim apresentam várias controvérsias e imprecisões, pois estão ligados à biologia do tumor. Entre essas variáveis encontra-se a gradação do tumor, e qualquer sistema de gradação aceito até hoje na prática patológica sofre de algum grau de subjetividade (FUHRMAN; LASKY; LIMAS, 1982), isso é o que ocorre com o de Fuhrman.

Um modelo matemático é uma representação ou interpretação simplificada da realidade. Em medicina, existem alguns procedimentos e características das variáveis estabelecidas. As doenças geralmente são descritas através de termos linguísticos, que são vagos, e que muitas vezes constituem variáveis qualitativas em medicina. O tratamento dessas variáveis, utilizando métodos quantitativos, apresenta dificuldades. (SILVEIRA, 2007) (SAVERGNINI, 2009).

A Teoria dos Conjuntos Fuzzy, introduzida por Lotfi Asker Zadeh, em 1965, surgiu a partir de situações duvidosas, nas quais a propriedade que define o conjunto é incerta. A ideia de Zadeh foi flexibilizar a pertinência de elementos aos conjuntos, criando o conceito de grau de pertinência, que descreve com melhor exatidão conceitos subjetivos (BARROS; BASSANEZI, 2010). Essa Teoria tem por objetivo associar a precisão da

matemática e as imprecisões do mundo real; deste modo, os dados incertos e a opinião do especialista são levados em consideração e incorporados aos modelos matemáticos.

Tendo em vista o crescimento da incidência de câncer no Brasil ultimamente, em decorrência do aumento da expectativa de vida e de uma maior exposição da população a fatores de risco, sua prevenção e tratamento têm sido uma preocupação da ciência médica.

Os modelos matemáticos em medicina, são classificados, segundo alguns autores, como modelos estocásticos, isto é, modelos que utilizam teoria de probabilidade. Existem também os modelos determinísticos, nos quais a evolução das variáveis é obtida pela estrutura do modelo, dadas as condições iniciais e os valores dos parâmetros (SILVEIRA, 2007).

Segre (SEGRE, 2010) é um exemplo de trabalho que faz uso de análise estatística a fim de obter um nomograma para o câncer renal que pode refletir o comportamento do CCR em relação à recidiva do tumor. Uma das conclusões deste projeto é a importância do tamanho da lesão como variável anatomopatológica. Esse resultado influenciou a construção de um dos modelos expostos nesta dissertação. Na literatura matemática, notam-se vários trabalhos que utilizam teoria de conjuntos *fuzzy* para analisar eventos biológicos, em particular, neoplasias. Os modelos matemáticos desenvolvidos têm contribuído amplamente para uma maior compreensão dos fenômenos médicos, diagnósticos, prognósticos e aumentando a eficiência dos tratamentos. Trata-se de uma linguagem do pensamento lógico, linguagem que está sendo cada vez mais necessária na área de saúde. É o que ocorre com os trabalhos sobre Câncer de Próstata e de Bexiga, os quais serviram de referências a esta pesquisa (SILVEIRA, 2007) (SAVERGNINI, 2009).

Assim, a subjetividade das informações no estudo do câncer renal nos motivou a utilizar a Teoria de Conjuntos Fuzzy para tratar desse problema. Diante das incertezas das informações, cremos que o uso dessa teoria para o presente estudo é apropriado.

O objetivo deste trabalho é a construção de dois modelos matemáticos *fuzzy* para prever os riscos de metástase e morte para indivíduos com tumores renais e analisar a relação entre a graduação de Fuhrman e o prognóstico de pacientes com essa neoplasia.

No primeiro capítulo são apresentadas informações gerais sobre o câncer, em particular sobre o câncer de rim. Além disso, algumas informações sobre o estadiamento, diagnóstico e padrão histológico, também são apresentados neste capítulo.

No Capítulo 2 são introduzidos alguns conceitos básicos sobre a Teoria dos Conjuntos *Fuzzy*, lógica *fuzzy* e sistema baseado em regras *fuzzy* (SBRF) que serão utilizados neste trabalho. Definições e algumas propriedades das medidas e distribuições de probabilidade e possibilidade, que serão úteis na modelagem, também serão apresentadas.

O Capítulo 3 é destinado à construção de dois modelos baseados em regras

*fuzzy*: um para prever o risco de metástase e outro para prever o risco de morte para pacientes com tumores renais. Tal previsão, necessária para a decisão sobre qual terapia ser aplicada, é dada de forma subjetiva. A saída de cada sistema é um número real, que associada à variável de saída, informa a possibilidade do risco de metástase e de morte. Os resultados de simulações bem como as suas probabilidades também serão apresentados. Além disso, nesse capítulo retratamos um estudo estatístico para analisar as correlações entre os fatores prognósticos e a sobrevivência dos pacientes. O estudo também envolveu a pesquisa com duas classificações de Fuhrman: grau único (padrão utilizado na área médica, citado anteriormente) e grau combinado: graduação que considera a soma dos graus presentes no tumor (muitas vezes, ocorre encontrar dois graus de Fuhrman na extensão tumoral), essa classificação é sugerida pelo médico patologista Prof. Dr. Athanase Billis do Hospital das Clínicas - UNICAMP. Com a finalidade de realizar esses estudos, utilizamos um banco de dados de pacientes do Hospital das Clínicas - UNICAMP.

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

# 1 Câncer

Tendo em vista o crescimento da incidência de câncer nos últimos anos, sua prevenção e tratamento tem sido uma grande preocupação da ciência médica (CASTANHO, 2005). Segundo a Organização Mundial da Saúde (OMS), o câncer é a segunda principal causa de morte em todo o mundo, com 8,2 milhões de mortes relacionadas ao câncer em todo o mundo em 2012, uma estimativa de 13% de todas as mortes mundiais. O número anual de novos casos deverá aumentar de 14,1 milhões em 2012 para 21,6 milhões em 2030. Cerca de 75% das mortes por câncer ocorrem em países de baixa e média renda, onde o número de casos de câncer está aumentando mais rapidamente (WORLD HEALTH ORGANIZATION - WHO, 2016).

Com base no documento World cancer report 2014 da International Agency for Research on Cancer (Iarc), da Organização Mundial da Saúde (OMS), é inquestionável que o câncer é um problema de saúde pública, especialmente entre os países em desenvolvimento, onde é esperado que, nas próximas décadas, o impacto do câncer na população corresponda a 80% dos mais de 20 milhões de casos novos estimados para 2025. No Brasil, as doenças cardiovasculares, o câncer, as causas externas e o diabetes representam mais da metade do total de causa de óbitos. A estimativa para o Brasil, biênio 2016-2017, aponta a ocorrência de cerca de 600 mil casos novos de câncer (INSTITUTO NACIONAL DE CÂNCER - INCA, 2017).

Câncer é a doença que se caracteriza pelo crescimento desordenado (maligno) de células que invadem os tecidos e órgãos, podendo espalhar-se (metástase) para outras regiões do corpo. Dividindo-se rapidamente, estas células tendem a ser muito agressivas e incontroláveis, determinando a formação de tumores (acúmulo de células cancerosas) ou neoplasias malignas (INSTITUTO NACIONAL DE CÂNCER - INCA, 2016).

É muito difícil precisar a cura para o câncer. Muitas vezes, ela não é possível, e o objetivo passa a ser a melhoria da qualidade de vida do paciente sem necessariamente, estender seu tempo de vida. Para tanto, conhecer o comportamento dessa doença, ajuda a controlar a sua evolução e conseqüentemente, definir tratamentos mais adequados e importantes para a sobrevivência dos pacientes.

Em 30 de maio de 2017, a 70<sup>a</sup> Assembléia Mundial da Saúde aprovou um projeto de resolução sobre prevenção e controle do câncer com 18 patrocinadores e mais de 40 Estados membros e 11 ONGs que falam em apoio da resolução. Durante a sua deliberação, houve um amplo consenso de que o câncer é uma preocupação crescente em saúde pública que exige maior atenção, priorização e financiamento (WORLD HEALTH ORGANIZATION - WHO, 2016).

A incidência de carcinoma de células renais (CCR) está aumentando globalmente. Atualmente, a doença representa aproximadamente 2,0% de todos os novos casos de câncer e mais de 100.000 mortes em todo o mundo. As melhorias no diagnóstico por imagem podem ter contribuído para a crescente incidência de CCR nas últimas três décadas, uma vez que a detecção de pequenos tumores têm aumentado. No entanto, pelo menos nos Estados Unidos, tanto a incidência de CCR em estágio tardio como as taxas de mortalidade por causa da doença também estão aumentando, o que implica que as melhorias na capacidade de diagnosticar o CCR estão sendo superadas pela crescente prevalência de alguns dos fatores de risco para este tumor (NARDI, 2010).

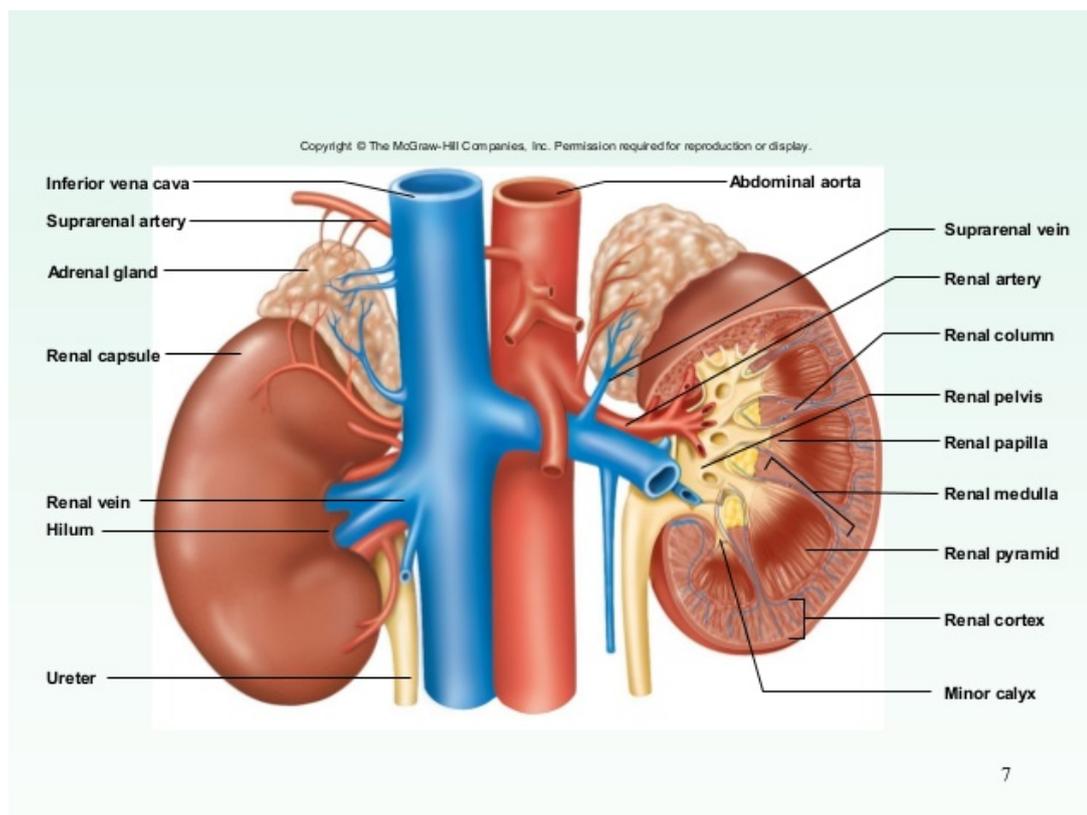
Em 2002, a estimativa global foi de 20.800 novos casos e 10.200 mortos, representando cerca de 1,9% dos tumores em geral. A maior incidência ocorreu na América do Norte, Austrália / Nova Zelândia e Europa, com taxas mais baixas em África, Ásia e Pacífico (ORNELLAS, 2012). O carcinoma das células renais (CCR) tem maior incidência nos países ocidentais. Na Europa há anualmente um aumento geral na incidência de cerca de 2% (exceto na Dinamarca e Suécia) (LJUNGBERG, 2009).

No Brasil, atualmente, os dados sobre as características gerais de pacientes com neoplasia renal em escala nacional são escassos, uma vez que a doença não está entre os 10 tipos de tumores mais frequentes relatados anualmente pelo Instituto Nacional do Câncer. A Sociedade Brasileira de Urologia (SBU), iniciou um projeto destinado a realizar um estudo epidemiológico sobre câncer renal em nosso país, o objetivo deste estudo foi descrever as características demográficas, clínicas e patológicas do CCR. Os dados de pesquisas regionais sugerem que CCR representa aproximadamente 1,2% de todos os casos de câncer no país (NARDI, 2010).

## 1.1 Rim

O rim tem a forma de um “grão de feijão”. Ele apresenta um polo inferior e um polo superior, e entre os dois, voltado para a região medial encontra-se o hilo renal, que apresenta o acesso para o seio renal e é utilizado como abertura de entrada e saída para os vasos sanguíneos, além de entrada para os nervos e saída para o ureter. A glândula suprarrenal está situada com uma base achatada sobre o polo superior do rim. O hilo também é local de entrada dos vasos sanguíneos (PAULSEN, 2016). Todas essas estruturas podem ser observadas na Figura 1, assim como a veia cava que transporta o sangue com gás carbônico, a aorta que transporta o sangue oxigenado e as demais estruturas que compõem o rim.

Figura 1 – Anatomia do Rim.



Fonte: <http://www.slideshare.net/MichaelWalls1/chapter-20-urinary-system-30957017>.

## 1.2 Câncer de Rim - Caso de Células Claras

Os rins são órgãos responsáveis pelo equilíbrio de água e sais do corpo, além de exercer uma função importante na eliminação de substâncias metabolizadas pelo organismo. Como ele é um órgão que se localiza mais profundamente na barriga, sintomas mais intensos e possibilidade de palpação do câncer só ocorrem quando a doença está mais avançada. Muitos casos são descobertos já com metástases e os sintomas podem ser decorrentes mais da metástase do que do próprio tumor de rim (HOSPITAL DO CÂNCER DE BARRETOS, 2016). Os sintomas, quando apresentados, são hematúria (sangue na urina) e dores lombares e abdominais.

Os avanços de diagnóstico e tratamento aliados com uma melhor compreensão sobre a doença levam à importantes alterações de paradigmas relacionados a abordagem de pacientes com câncer. O entendimento dessas mudanças e qual o impacto que elas podem gerar na sobrevivência dos pacientes são pontos relevantes para a definição prognóstica em uma população de indivíduos com a doença. Assim, torna-se necessário definirem-se quais mudanças são de maior relevância para que se possa integrá-las e entendê-las a fundo (SEGRE, 2010).

### 1.2.1 Diagnóstico

O diagnóstico do câncer de rim se dá por diversas formas, entre elas: ultrassonografia, ressonância nuclear magnética, tomografia computadorizada, cintilografia óssea e biópsia. Geralmente, o ultrassom é utilizado em estágios mais precoces da doença, em que visualiza-se a presença de nódulos ou massa renal; a ressonância é mais utilizada para verificar alterações vasculares e cistos renais, por exemplo; já a tomografia é útil para determinar o estadiamento da doença (extensão para outros órgãos) e verificar possíveis metástases pulmonares; para metástases ósseas, utiliza-se cintilografia; e a biópsia é realizada antes da retirada do tumor.

### 1.2.2 Tipos de tumores

Existem dois tipos principais de câncer de rim primário: tumores corticais renais e tumores de células de transição (urotelial). Estes tipos são diferentes na forma como eles se desenvolvem e mudam ao longo do tempo, bem como no tratamento.

Os tumores uroteliais começam na pelve renal, quando a urina deixa os túbulos renais. Os fatores como a localização e o tamanho de um tumor de células de transição são levados em conta para decidir sobre a melhor abordagem para tratamento desse tipo de câncer ([MEMORIAL SLOAN KETTERING CANCER CENTER, 2017](#)).

Até 90% dos tumores renais são tumores corticais renais. Estes tumores começam na parte principal do rim, que contém estruturas chamadas túbulos renais. Os rins possuem mais de um milhão de néfrons, unidades funcionais básicas do rim, que filtram sangue e são responsáveis pela formação da urina ([MAGALHÃES, 2017](#)). Cada néfron está associado a um túbulo renal. Os túbulos juntam-se para formar dutos colectores. Estes dutos depositam a urina na pelve renal.

O câncer que se desenvolve nos túbulos renais é chamado de tumor cortical renal. Existem vários tipos, e eles não necessariamente afetam as pessoas de maneiras previsíveis. Alguns tipos são malignos (cancerosos). Mas outros tipos - como cistos renais, oncocitomas e angiomiolipomas - são benignos (não cancerosos).

O carcinoma de células renais (CCR) é o tumor maligno mais comum do rim. O CCR inclui quatro subtipos geneticamente diferentes: convencional (células claras) representa 80 a 90% dos tumores renais, papilar representa 10 a 15% deles, CCR cromofóbico representa 4 a 5%, e carcinoma do tubo colector representa 1% dos tumores renais ([LJUNGBERG, 2009](#)).

Tumores papilares podem começar como um ou mais tumores em um ou ambos os rins. Existem dois tipos de tumores papilares. Tipo 1 é mais comum e geralmente cresce lentamente. Tipo 2 é geralmente muito mais agressivo e imprevisível. Os tumores papilares foram associados a certas síndromes geneticamente herdadas, tais como carcinoma renal

papilar hereditário, leiomiomatose hereditária e carcinoma de células renais. Muitas vezes, é necessário fazer testes genéticos se o médico suspeitar de uma dessas síndromes.

Tumores cromófobos são considerados uma forma menos agressiva de câncer de rim. Eles podem atingir um tamanho muito grande antes de se espalharem.

Os tumores dos ductos coletores são muito raros e mais comuns em adultos jovens. Eles são agressivos e não respondem ao tratamento convencional.

Neste trabalho, abordaremos o tipo células claras devido a sua maior incidência. Uma curiosidade desse tipo de célula é que quando vistas sob um microscópio, as células que formam o CCR parecem muito pálidas ou claras ([AMERICAN CANCER SOCIETY, 2016](#)).

### 1.2.3 Fatores Prognósticos

Como se sabe o câncer possui determinados parâmetros relacionados com a evolução da doença, que são chamados **fatores prognósticos**; o caso do câncer de rim apresenta diversas variáveis que determinam o comportamento desse tipo de tumor.

Cheville et al ([CHEVILLE, 2003](#)) constataram que o grau nuclear, o tamanho do tumor e o estágio TNM (Tumour Node Metastasis) estão associados com a mortalidade nos três tipos de carcinomas de células renais mais frequentes (células convencionais, papilar e cromóforo). Além disso, nesse mesmo estudo, a presença de necrose foi associada a uma maior taxa de mortalidade nos CCR convencionais e CCR cromóforo.

A necrose é a morte celular. As células tumorais passam por mutações (involuntárias) e começam a se dividir e multiplicar, aumentando o tamanho do tumor. O sangue, responsável pela “alimentação” do mesmo, não consegue suprir os nutrientes para essas células, o que ocasiona sua morte. Assim algumas áreas ficam necrosadas (mortas), indicando aumento do tamanho da lesão e conseqüentemente, um prognóstico desfavorável.

A invasão vascular é o envolvimento de veias renais e vasos do hilo. Ela é considerada fator de risco no desenvolvimento, diagnóstico e evolução do CCR ([SEGRE, 2010](#)) ([ORNELLAS, 2012](#)).

A diferenciação sarcomatóide é quando o tumor tem formato mais fino, apresentando caráter mais agressivo. Sabe-se que tumores que apresentam essa característica evoluem com pior prognóstico ([POMPEO, 2006](#)).

Além disso, o tumor pode comprometer outras partes do corpo próximas ao rim: o tecido adiposo perirrenal (que se encontra ao redor do rim) ([ORNELLAS, 2012](#)) e a glândula suprarrenal (localizada logo acima). Essas invasões sugerem também, um pior prognóstico ([POMPEO, 2006](#)).

O tamanho do tumor indica um prognóstico desfavorável também. No sentido

que, quanto maior o tamanho, mais chances o tumor tem de se espalhar e determinar um prognóstico ruim. Pode ser encontrado em literatura a correlação entre o tamanho e sobrevida livre de recidiva (MEDEIROS; GELB; WEISS, 1988) (ORNELLAS, 2012). A recidiva compreende a reincidência do tumor (podendo ser local ou à distância; quando ocorre à distância, é chamada metástase). Segundo Segre (2010), o tamanho do tumor é a variável anatomopatológica mais importante.

A graduação histológica consiste na diferenciabilidade do núcleo celular de uma célula cancerosa ao compará-la com uma sadia. O sistema de classificação mais utilizado é o de Fuhrman (FUHRMAN; LASKY; LIMAS, 1982). A graduação histológica de Fuhrman iniciou-se com o trabalho realizado pelos médicos Dra. Susan A. Fuhrman, Dr. Larry C. Lasky e Dra. Catherine Limas, em 1982. Por meio de visualização ao microscópio, classifica-se o padrão nuclear celular da neoplasia em quatro graus, considerando-se o tamanho do núcleo, a presença de irregularidades nucleares, assim como a presença de proeminência do nucléolo (Tabela 1). Podemos notar essas diferenças na Figura 2. Levando em conta a sobrevida, os tumores de grau 1 são considerados com prognóstico favorável, os de grau 4 com prognóstico pior e os de graus 2 e 3, com intermediário. Desta forma, pode-se observar que uma determinada característica histológica, no caso o grau nuclear, parece ser o critério morfológico mais significativo para o prognóstico.

Tabela 1 – Graduação Histológica segundo Fuhrman et al (FUHRMAN; LASKY; LIMAS, 1982).

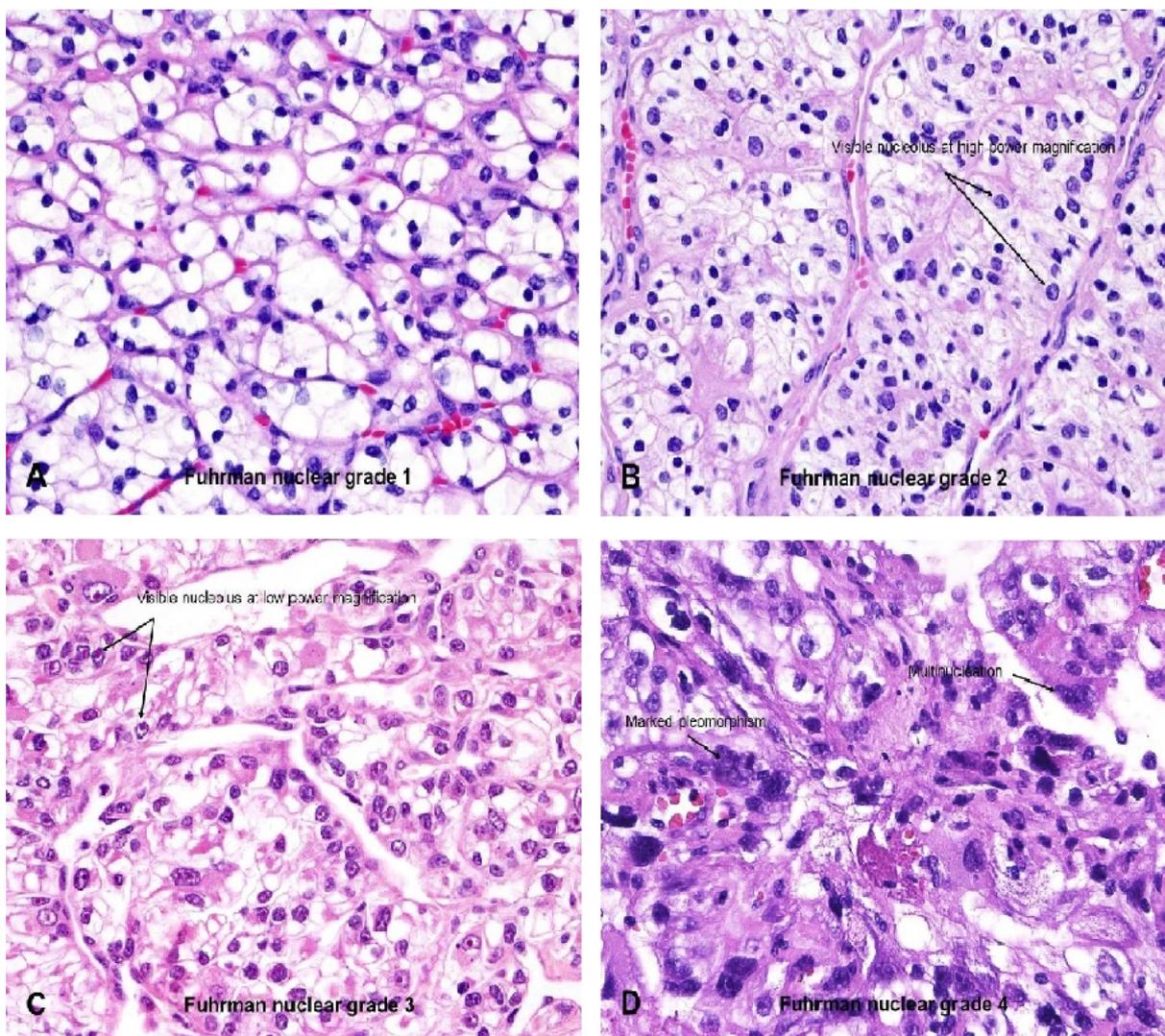
Grau	Características
I	Núcleo arredondado e uniforme, com cerca de $10\mu\text{m}$ de diâmetro e nucléolo ausente/pequeno.
II	Núcleo levemente irregular, com diâmetro de $15\mu\text{m}$ e nucléolo visível, mas pequeno.
III	Núcleo moderadamente irregular, com diâmetro de $20\mu\text{m}$ e nucléolo grande.
IV	Núcleo acentuadamente irregular/ pleomórfico e formas multilobulares, com cromatina agrupada e diâmetro maior que $20\mu\text{m}$ .

Fonte: Fuhrman, Lasky e Limas (1982), apud Pompeo (2006).

No trabalho sobre Fuhrman (FUHRMAN; LASKY; LIMAS, 1982), nota-se uma diferença considerável na taxa metastática entre o grau 1 e os demais graus, quando o tumor está confinado ao rim sem envolvimento da gordura perirrenal. Os tumores de grau 1 são claramente menos malignos, contudo quando atingem um tamanho grande, eles podem penetrar a cápsula e eventualmente metastizar. A maior diferença nas taxas metastática e de sobrevida, quando o tumor está confinado ao rim, foi observada entre os graus 1 e 2, sugerindo que até mesmo um desvio modesto da morfologia nuclear normal reflete uma profunda mudança no comportamento biológico. Vale ressaltar que, o maior

inconveniente ao determinar o grau pela biópsia deriva do fato que mais de um grau são encontrados no mesmo tumor. Uma vez que o grau mais alto presente em um tumor determina o prognóstico, e como a amostragem por biópsia pode produzir uma classificação erroneamente inferior, a classificação pré-operatória deve ser interpretada com cautela. Qualquer sistema de graduação aceito na prática da patologia sofre de algum grau de subjetividade, mas não necessariamente impede seu uso. Além disso, a graduação de Fuhrman se caracteriza por pobre reprodutibilidade e falta de uniformidade (POMPEO, 2006).

Figura 2 – Carcinoma de células renais de células claras. A: Tumor de grau nuclear 1 com núcleos redondos ou uniformes; nucléolos não discerníveis ou ausentes. B: Carcinoma de grau nuclear 2 com contornos nucleares ligeiramente irregulares e nucléolos discretos (visíveis a 400x). C: A neoplasia nuclear de grau 3 possui núcleos grandes e irregulares com nucléolos visíveis a 100x. D: Carcinoma nuclear de grau 4 com núcleos bizarros e nucléolos grandes e proeminentes.



Fonte: <https://www.auanet.org/education/modules/pathology/kidney-carcinomas/fuhrman-grade.cfm>.

Além dessa série de parâmetros considerados fatores prognósticos importantes, acredita-se que o estágio da doença tem valor prognóstico e pode influenciar a sobrevivência dos pacientes. Na literatura, encontramos essa variável como fator prognóstico adverso e associa-se com a evolução da doença ou mesmo a morte (POMPEO, 2006) (ORNELLAS, 2012). Recomenda-se a classificação UICC TNM 2009 (Tumour Node Metastasis) para o estadiamento do CCR, que está descrito na Tabela 2 (FERREIRA; SASSE, 2015) (LJUNGBERG, 2009).

Existem diferentes sistemas de estadiamento, entretanto o mais comum e útil para a maioria dos tipos de cânceres é o sistema TNM. O sistema de classificação TNM descreve a extensão do tumor e o envolvimento de órgãos. Ele é utilizado pelos médicos para avaliar o grau da doença. Esse sistema descreve três peças chave de informação:

1. **T:** indica o tamanho do tumor primário e se cresceu nas áreas próximas.
2. **N:** descreve a extensão da disseminação para os gânglios linfáticos (regionais) ou linfonodos nas proximidades. Os linfonodos são pequenas coleções de células do sistema imunológico às quais os cânceres geralmente se espalham primeiro.
3. **M:** indica se o câncer se espalhou (metastatizou) para outras partes do corpo. (Os lugares mais comuns de disseminação são os pulmões, ossos, fígado, cérebro e linfonodos distantes.)

Diante de vários fatores prognósticos considerados relevantes para a formação de conceito médico, a interpretação do estado do paciente, bem como de diagnóstico e tomada de decisão terapêutica, torna-se importante desenvolver modelos matemáticos que combine tais parâmetros, para prever o risco de metástase e de morte em pacientes com essa neoplasia. A ferramenta escolhida, para a construção desses modelos, foi a teoria dos conjuntos *fuzzy*, devido a sua capacidade de lidar com as incertezas e subjetividades envolvidas no problema.

Tabela 2 – TNM 2009 - Classificação Clínica do CCR.

<b>T</b>	<b>Tumor Primário.</b>
TX	Tumor primário não pode ser avaliado.
T0	Não há evidência de tumor primário.
T1	Tumor limitado ao rim.
T1a	Menor que 4 cm.
T1b	Maior que 4 cm, menor que 7 cm.
T2	Tumor limitado ao rim, com mais de 7 cm.
T2a	Tumor com mais de 7 cm até 10 cm.
T2b	Tumores com mais de 10 cm.
T3	Tumor invade a veia, a suprarrenal ou tecidos periféricos, porém não ultrapassa a fáscia de Gerota.
T3a	Invasão de adrenal, gordura perirrenal ou gordura do seio renal.
T3b	Invasão de veias intrarrenais, renal ou cava inferior infradiafragmática.
T3c	Extensão para veia cava inferior supradiafragmática, ou invade a parede da veia cava.
T4	Tumor ultrapassa a fáscia de Gerota.
<b>N</b>	<b>Linfonodos Regionais.</b>
NX	Os linfonodos regionais não podem ser avaliados.
N0	Ausência de metástase em linfonodo regional.
N1	Metástase em um único linfonodo regional.
N2	Metástase em mais de um linfonodo regional.
<b>M</b>	<b>Metástase à Distância.</b>
MX	A presença de metástase à distância não pode ser avaliada.
M0	Ausência de metástase à distância.
M1	Metástase à distância.

Fonte: [Ljungberg \(2009\)](#) e [Ferreira e Sasse \(2015\)](#).

## 2 Aspectos da Teoria Fuzzy

Como foi dito por Lotfi A. Zadeh (BARROS; BASSANEZI, 2010),

*“Há e haverá muitas tarefas que os homens podem cumprir com facilidade, que vão além da capacidade de qualquer computador, qualquer máquina e qualquer sistema lógico que podemos conceber nos dias de hoje.”*

No cotidiano, as ações humanas controlam os mais variados sistemas do mundo real por meio de informações imprecisas. O controle e a execução de tarefas devem seguir uma sequência de “ordens” linguísticas, traduzidas por um conjunto de regras, capazes de serem decodificadas pelo controlador.

Foram os primeiros trabalhos de Zadeh sobre teoria dos conjuntos fuzzy que deram origem ao que hoje é chamado de lógica fuzzy. 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. Para modelar fenômenos do mundo real tem-se que lidar com 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.

O objetivo deste capítulo é apresentar conceitos fundamentais da teoria *fuzzy* necessários para o desenvolvimento da dissertação.

### 2.1 Conjuntos *Fuzzy*

**Definição 2.1.** *Seja  $U$  um conjunto e  $A$  um subconjunto de  $U$ . A função característica de  $A$ ,  $\chi_A : U \rightarrow \{0, 1\}$  é dada por*

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

Desse modo,  $\chi_A$  é uma função cujo domínio é  $U$  e a imagem está contida no conjunto  $\{0, 1\}$ , onde  $\chi_A(x) = 1$  indica que  $x$  é um elemento de  $A$  e  $\chi_A(x) = 0$  indica que  $x$  não está em  $A$ . Podemos generalizar essa noção 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.** Um subconjunto fuzzy  $F$ , do universo  $U$  (clássico) é caracterizado por uma função  $\varphi_F : U \rightarrow [0, 1]$  chamada função de pertinência do subconjunto fuzzy  $F$ , onde  $\varphi_F(x)$  indica o grau de pertinência do elemento  $x$  em  $F$ . Podemos dizer que o subconjunto fuzzy  $F$  de  $U$  é dado por um conjunto (clássico) de pares ordenados:

$$F = \{(x, \varphi_F(x)) | x \in U\}.$$

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

Segundo Klir, Clair e Yuan (1997), o grau de pertinência  $\varphi_F(x)$ , de um elemento  $x$  em  $F$ , pode ser interpretado como o grau de compatibilidade do atributo associado com  $F$  e o elemento  $x$ . Seguem abaixo algumas definições importantes.

**Definição 2.3.** O suporte de um conjunto fuzzy  $F$  é o conjunto clássico de todos os elementos  $x \in U$ , cuja função de pertinência tem valor diferente de zero:

$$\text{supp}(F) = \{x \in U | \varphi_F(x) > 0\}.$$

**Definição 2.4.** O  $\alpha$  - cut ou  $\alpha$  - nível de um conjunto fuzzy  $F$ , denotado por  $F_\alpha$ , é um conjunto que consiste de elementos do universo  $U$  cujos valores de pertinência são iguais ou excedem um certo nível limite  $\alpha$ , onde  $\alpha \in [0, 1]$ :

$$F_\alpha = \{x \in U | \varphi_F(x) \geq \alpha\}.$$

**Definição 2.5.** O núcleo (core) de um conjunto fuzzy  $F$  é o conjunto clássico de  $U$ :

$$\text{core}(F) = \{x \in U | \varphi_F(x) = 1\}.$$

**Definição 2.6.** A altura de um conjunto fuzzy  $F$  é o maior grau de pertinência obtido, por qualquer elemento no conjunto, ou seja:

$$\text{hgt}(F) = \sup_{x \in U} \varphi_F(x).$$

**Definição 2.7.**  $F$  é dito normal se sua função de pertinência atinge o valor 1, isto é:

$$\sup_{x \in X} \varphi_F(x) = 1.$$

Se essa propriedade não é assegurada, chamamos de conjunto fuzzy subnormal.

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

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

$$Norm[\varphi_F(x)] = \frac{\varphi_F(x)}{hgt(F)}$$

*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:

$$Con[\varphi_F(x)] = (\varphi_F)^p(x)$$

onde  $p > 1$ .

**Definição 2.8.** Para qualquer conjunto *fuzzy*  $F$ , definido num universo finito  $U$ , a cardinalidade é definida por

$$|F| = \sum_{x \in U} \varphi_F(x).$$

ou alternativamente, como a integral seguinte:

$$|F| = \int_U \varphi_F(x) dx$$

Quando  $U$  não é finito,  $|F|$  nem sempre existe. Entretanto, se  $F$  tem suporte finito

$$|F| = \sum_{x \in \text{supp}U} \varphi_F(x).$$

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

**Definição 2.9.** O complementar  $\bar{F}$  de um conjunto *fuzzy*  $F$  com respeito ao universo  $U$ , é definido para todo  $x \in U$  por

$$\varphi_{\bar{F}}(x) = 1 - \varphi_F(x).$$

**Definição 2.10.** Dados dois conjuntos *fuzzy*  $A$  e  $B$ , a interseção  $A \cap B$  e a união  $A \cup B$  são definidas, para todo  $x \in U$ , por

$$\varphi_{A \cap B}(x) = \varphi_A(x) \wedge \varphi_B(x) \quad \varphi_{A \cup B}(x) = \varphi_A(x) \vee \varphi_B(x),$$

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

## 2.2 Lógica Fuzzy

Lógica é a área do conhecimento que estuda as consequências e 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* é uma lógica com suas próprias metodologias. Ela suporta os modos de raciocínio que são aproximados, ao invés de exatos. Ela trata de incertezas no sentido que, quanto menos incertas forem as premissas, menos incertas serão as conclusões. Desta forma, podemos dizer que a lógica clássica é uma espécie de limite da lógica *fuzzy* quando as incertezas tendem a zero (BARROS; BASSANEZI, 2010), sendo assim interpretado como um processo de composição nebulosa. O principal sucesso da lógica *fuzzy* está no fato de ter um caráter prático, já que possibilita conclusões a partir de proposições incertas.

Para Bilobrovec (2005), 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 necessariamente limites bruscos, mas proporciona graus de pertinência de elementos a uma determinada categoria.

### 2.2.1 Variáveis Linguísticas

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

As variáveis e suas classificações linguísticas 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: pelos intuitivos, ajustes de curvas, interpolações e até mesmo redes neurais (BARROS; BASSANEZI, 2010).

Os rótulos que descrevem uma variável linguística são chamados termos primários, como por exemplo, *baixo*, *médio*, *alto*, entre outros dados por conjuntos *fuzzy*. Podem ainda ser aplicados modificadores linguísticos, a esses termos primários, como *muito*, *pouco*, *mais ou menos*, dentre outros. Conectivos são operadores lógicos que combinam afirmações linguísticas em proposições. Os termos primários também podem ser relacionados por meio da negação *não*, ou através dos conectivos *e* e *ou*.

A importância de se utilizar variáveis *fuzzy*, ou variáveis linguísticas, está no fato de facilitarem a transição gradual entre estados, assim 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.

### 2.2.2 Relações Fuzzy

A compreensão de muitos fenômenos exige o estudo de relações ou interações entre os elementos de vários conjuntos. Segundo Barros e Bassanezi (2010), 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*. Seguem abaixo as definições de *t-norma* e *s-norma*, juntamente com outras que se fazem importantes para o trabalho.

**Definição 2.11.** *Uma norma triangular ou t-norma é uma operação binária  $t : [0, 1] \times [0, 1] \rightarrow [0, 1]$  que satisfaz as seguintes propriedades:  $\forall x, y \text{ e } z \in [0, 1]$ ,*

- *Comutativa:  $x t y = y t x$*
- *Associativa:  $x t (y t z) = (x t y) t z$*
- *Monotonicidade: Se  $y \leq z$  então  $x t y \leq x t z$*
- *Condições de Fronteira:  $x t 1 = x$  e  $x t 0 = 0$*

**Definição 2.12.** *Uma s-norma é uma operação binária  $s : [0, 1] \times [0, 1] \rightarrow [0, 1]$  que satisfaz as seguintes propriedades:  $\forall x, y \text{ e } z \in [0, 1]$ ,*

- *Comutativa:  $x s y = y s x$*
- *Associativa:  $x s (y s z) = (x s y) s z$*

- *Monotonicidade:* Se  $y \leq z$  então  $x \text{ s } y \leq x \text{ s } z$
- *Condições de Fronteira:*  $x \text{ s } 0 = x$  e  $x \text{ s } 1 = 1$

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

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

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

**Definição 2.13.** *Uma relação fuzzy  $\mathcal{R}$ , em  $A \times B$ , é um conjunto fuzzy caracterizado por uma função de pertinência  $\varphi_{\mathcal{R}}$ , que associa a cada par ordenado  $(x, y)$ , um grau de pertinência em  $\mathcal{R}$ ,*

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

Generalizando, se  $U$  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  $U$  é um subconjunto fuzzy de  $U$ , caracterizado por uma função de pertinência  $\varphi_{\mathcal{R}}(u_1, \dots, u_n)$ , com  $u_i \in U_i$ ,  $i = 1, \dots, n$ .

**Definição 2.14.** *O produto cartesiano  $\mathcal{R}$ , dos subconjuntos fuzzy,  $A_1, A_2, \dots, A_n$  de  $U_1, U_2, \dots, U_n$  é a relação fuzzy cuja função de pertinência é*

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

### 2.2.3 Regras Fuzzy

As regras fuzzy podem ser formalizadas através de coleções de relações fuzzy. Elas naturalmente fornecem uma maneira de construir modelos de sistemas envolvendo conhecimento de domínio, experiência e dados experimentais. A interpretação de regras fuzzy como relações fuzzy permite o uso de cálculo relacional para processar informações e realizar computação eficiente com regras fuzzy em inúmeras aplicações (GOMIDE; PEDRYCZ, 2007).

O papel da regra é descrever algumas partes do conhecimento sobre um problema a ser representado. A própria regra expressa uma certa relação entre as variáveis de entrada e saída (descritores). Ainda mais, vale destacar o fato de que a regra apresenta alguma relação interessante e transparente pertinente ao problema.

Silveira (2007) ressalta que uma proposição é um tipo particular de sentença, a saber, aquela que forma ou nega um predicado de um sujeito. Como já foi mencionado

anteriormente, a principal diferença entre proposições clássicas e proposições *fuzzy* é que, numa proposição *fuzzy* a verdade ou falsidade é uma questão de grau.

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 linguística 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  $U_1, U_2, \dots, U_n$  e  $x_1, x_2, \dots, x_n$  são elementos do domínio.

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

Por exemplo:

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

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

onde  $A_1, A_2, \dots, A_n$  são conjuntos *fuzzy* em  $U_1, U_2, \dots, U_n$  e  $B_1, B_2, \dots, B_m$  são conjuntos *fuzzy* em  $V_1, V_2, \dots, V_m$  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  $\mathcal{P}$  e  $\mathcal{Q}$  em  $U_1 \times U_2 \times \dots \times U_n \times V_1 \times V_2 \times \dots \times V_m$ , ou seja:

$$\mathcal{P}(x_1, \dots, x_n, y_1, \dots, y_m) = f(\mathcal{P}_a(x_1, \dots, x_n), \mathcal{P}_c(y_1, \dots, y_m)),$$

$$\mathcal{Q}(x_1, \dots, x_n, y_1, \dots, y_m) = f(\mathcal{Q}_a(x_1, \dots, x_n), \mathcal{Q}_c(y_1, \dots, y_m)),$$

onde  $f$  pode ser uma implicação, conjunção (t-normas) ou disjunção (s-normas) e  $\mathcal{P}_a$ ,  $\mathcal{Q}_a$ ,  $\mathcal{P}_c$  e  $\mathcal{Q}_c$  são as relações induzidas pelos antecedentes e consequentes das regras  $p$  e  $q$ . Logo, as regras podem ser expressas como proposições:

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

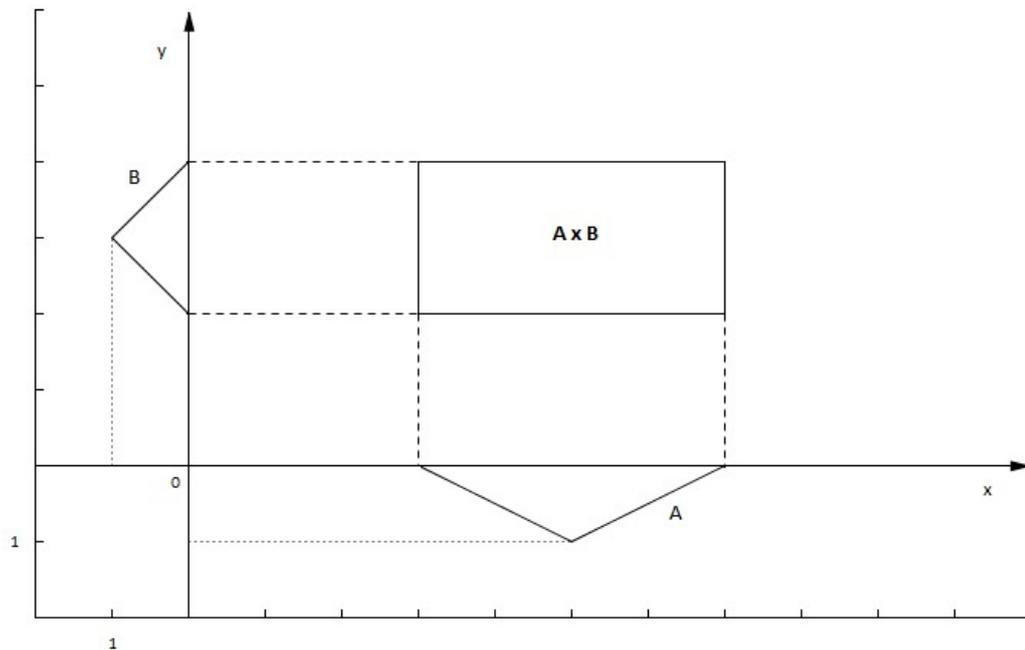
$$q : (x_1, x_2, \dots, x_n, y_1, y_2, \dots, y_m) \in \mathcal{Q}.$$

Uma regra pode ainda, ser ponderada, isto é ter a forma: “Se (*antecedente*), então (*consequente*) [*peso*]”, onde *peso* é um número real, no intervalo  $[0, 1]$ , que descreve a confiança da regra (GOMEZ; DASGUPTA, 2001). 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:

$$\mathcal{P}(x_1, \dots, x_n, y_1, \dots, y_m) = f(\mathcal{P}_a(x_1, \dots, x_n) * \textit{peso}, \mathcal{P}_c(y_1, \dots, y_m)).$$

A Figura 3 mostra uma representação gráfica da regra: “Se  $x$  é  $A$  então  $y$  é  $B$ ”.

Figura 3 – Representação gráfica da proposição  $p$ : Se  $x$  é  $A$  então  $y$  é  $B$ .



Fonte: Figura construída pela autora.

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

- Conjunção fuzzy - é uma função  $f_t : U \times V \rightarrow [0, 1]$ , definida por:

$$f_t(\varphi_A(x), \varphi_B(y)) = \varphi_A(x) \wedge \varphi_B(y), \quad \forall (x, y) \in U \times V.$$

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

- Disjunção fuzzy - é uma função  $f_s : U \times V \rightarrow [0, 1]$ , definida por:

$$f_s(\varphi_A(x), \varphi_B(y)) = \varphi_A(x) \text{ s } \varphi_B(y), \quad \forall (x, y) \in U \times V.$$

- Implicação fuzzy - é uma função  $f_i : U \times V \rightarrow [0, 1]$ , que obedece as seguintes condições:

- Monotônica no segundo argumento:

$$\varphi_B(y_1) \leq \varphi_B(y_2) \Rightarrow f_i(\varphi_A(x), \varphi_B(y_1)) \leq f_i(\varphi_A(x), \varphi_B(y_2));$$

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

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.

## 2.2.4 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; BARROS, 2005).

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

$$\varphi_{B'}(y) = \sup_{x \in X} \min[\varphi_{A'}(x), \varphi_{\mathcal{R}}(x, y)], \quad \forall y \in Y. \quad (2.1)$$

$$\text{Notação: } B' = \mathcal{R} \circ A' = \mathcal{R}(A')$$

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

$$\varphi_{\mathcal{R}}(x, y) = f[\varphi_A(x), \varphi_B(y)], \quad (2.2)$$

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

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

Regra: Se  $x \text{ é } A$  então  $y \text{ é } B$

Fato:  $x \text{ é } A'$

---

Conclusão:  $y \text{ é } B'$

Esse procedimento é chamado *modus ponens* generalizado. No esquema  $B'$  é calculado pela [Equação 2.1](#) e  $\mathcal{R}$  é determinado pela [Equação 2.2](#). 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 mecanismo 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 via uma conjunção do antecedente e o consequente da regra e, num conjunto 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  $B_j$  são subconjuntos *fuzzy* nos universos de discurso,  $X_k$  e  $U_k$ , e  $k$  é o número de conjuntos *fuzzy* no antecedente.

Regra 1: Se  $x_1 \text{ é } A_{11}, \dots, x_k \text{ é } A_{1k}$  então  $z_1 \text{ é } B_1$

ou

Regra 2: Se  $x_1 \text{ é } A_{21}, \dots, x_k \text{ é } A_{2k}$  então  $z_2 \text{ é } B_2$

ou

⋮ ⋮

ou

Regra n: Se  $x_1$  é  $A_{n1}$ ,  $\dots$ ,  $x_k$  é  $A_{nk}$  então  $z_n$  é  $B_n$

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

---

Conclusão:  $z$  é  $B = \mathcal{R}(A)$

O método de Mamdani propõe uma relação *fuzzy* binária  $\mathcal{M}$ , entre  $X$  e  $U$ , e baseia-se na regra de composição de inferência max-min. A seguir estão os passos para determinar  $B$ :

- Em cada regra  $R_j$  da base de regras, a proposição

“Se  $x$  é  $A_j$  então  $u$  é  $B_j$ ”

é modelada pela aplicação do mínimo ( $\wedge$ ).

Se a regra for ponderada, o peso  $w_j$  deve ser incluído no grau de ativação da regra,

- Para o conectivo lógico “e” usa-se a t-norma mínimo ( $\wedge$ ).
- Para conectar as regras *fuzzy* da base de regras, usa-se a s-norma máximo ( $\vee$ ) para o conectivo lógico “ou”.

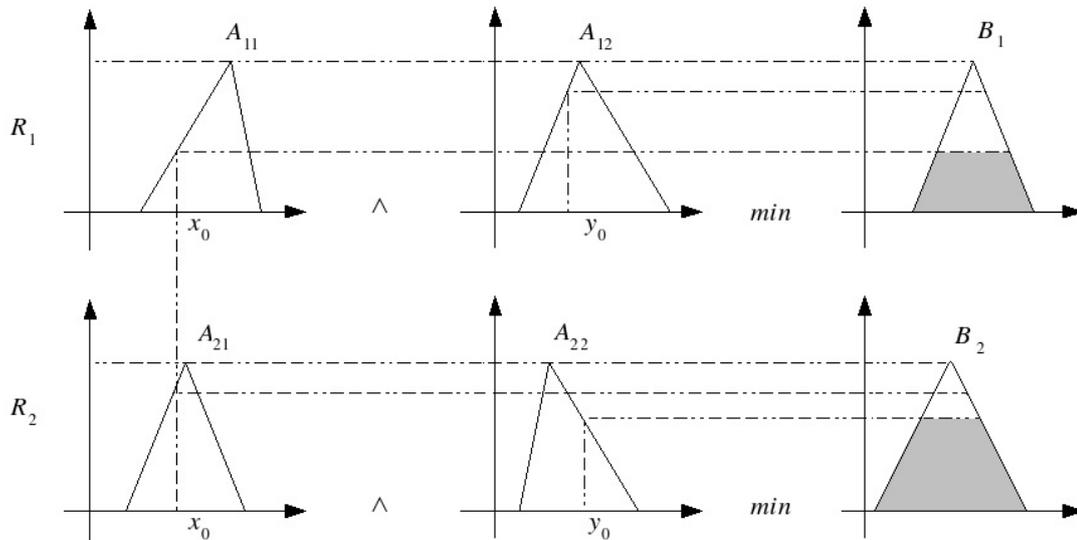
A relação fuzzy  $\mathcal{M}$  é o subconjunto *fuzzy* de  $X \times U$  cuja função de pertinência é

$$\varphi_{\mathcal{M}}(x, u) = \max_{1 \leq j \leq r} (\varphi_{R_j}(x, u)) = \max_{1 \leq j \leq r} [\varphi_{A_j}(x) \wedge \varphi_{B_j}(u)],$$

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

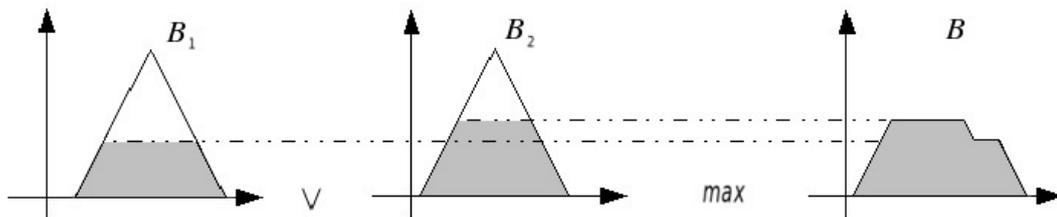
A [Figura 4](#) ilustra as saídas parciais, do método de inferência de Mamdani, para duas entradas reais,  $x_0$  e  $y_0$ , uma saída e duas regras.

Figura 4 – Saídas parciais no método de Mamdani.



Fonte: Silveira (2007).

Figura 5 – Saída final no método de Mamdani.



Fonte: Silveira (2007).

A saída final  $B$  é dada pela união das saídas parciais, como mostra a Figura 5.

O resultado do método de Mamdani, último gráfico da Figura 5, representa a função de pertinência  $\varphi_B$  do controle  $B$  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 (KTS)

Este método é similar ao método de Mamdani em muitos aspectos. A principal diferença entre a inferência *fuzzy* do tipo Mamdani e a do tipo KTS é que, nesse último, o consequente de cada regra é dado explicitamente por uma função das variáveis de entrada desta regra.

Este método foi desenvolvido por Sugeno e colaboradores para contornar o problema que envolve muitas áreas de atuação, como por exemplo a engenharia, em que frequentemente estão disponíveis informações funcionais sobre a estrutura dos sistemas, por exemplo conservação de energia.

O método de raciocínio, proposto por KTS, está associado a um conjunto de regras do tipo, “se ... então”, cujos consequentes são representados por funções (ORTEGA, 2001). Um conjunto de regras nesse modelo possui a seguinte forma:

$$\text{Regra}(i) : \text{Se } x_1 \text{ é } A_{i1}, \dots, x_k \text{ é } A_{ik} \text{ então } u_i = b_{i0} + b_{i1}x_1 + \dots + b_{ik}x_k,$$

onde  $i = 1, \dots, n$ ;  $n$  é o número de regras *fuzzy*;  $b_{ij}$ , com  $j = 0, \dots, k$ , são os parâmetros do consequente 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

$$u = \frac{\sum_{i=1}^n m_i z_i}{\sum_{i=1}^n m_i} = \frac{\sum_{i=1}^n m_i (b_{i0} + b_{i1}x_1 + \dots + b_{ik}x_k)}{\sum_{i=1}^n m_i}, \quad (2.3)$$

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

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

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

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

nas quais as saídas são combinadas da mesma forma que no caso linear, via expressão 2.3.

A grande vantagem desse modelo consiste no seu poder de representação, especialmente para descrever processos tecnológicos.

## Métodos de Defuzzificação

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. As diferenças são  $\sum \varphi_B(u_i)$  não precisa ser 1 e os pesos são os valores  $\varphi_B(u_i)$ , que indicam o grau de compatibilidade do valor  $u_i$ , com o conceito modelado pelo conjunto *fuzzy*  $B$ .

O centro de gravidade fornece a média da áreas de todas as figuras que representam os graus de pertinência de um conjunto *fuzzy* (BARROS; BASSANEZI, 2010). Esse método é o mais usual, embora seja complexo computacionalmente.

As equações 2.4 e 2.5 referem-se aos domínios discreto e contínuo, respectivamente.

$$G(B) = \frac{\sum_{i=0}^n u_i \varphi_B(u_i)}{\sum_{i=0}^n \varphi_B(u_i)} \quad (2.4)$$

$$G(B) = \frac{\int_{\mathbb{R}} u \varphi_B(u) du}{\int_{\mathbb{R}} \varphi_B(u) du} \quad (2.5)$$

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, por isso é considerado um procedimento radical. Neste caso temos:

$$C(B) = \frac{i + s}{2}$$

onde

$$i = \inf\{u \in \mathbb{R} : \varphi_B(u) = \max_u \varphi_B(u)\}$$

e

$$s = \sup\{u \in \mathbb{R} : \varphi_B(u) = \max_u \varphi_B(u)\}.$$

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

$$M(B) = \frac{\sum u_i}{i}$$

onde  $u_i$  são os elementos de maiores pertinências ao conjunto *fuzzy* B, isto é, para cada  $i$  tomamos

$$\varphi_B(u_i) = \max_u \varphi_B(u).$$

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

$$A(B) = \frac{\sum_{i=1}^n m_i \text{hgt}(B_i)}{\sum_{i=1}^n \text{hgt} B_i}$$

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 ao invés de um conjunto *fuzzy*.

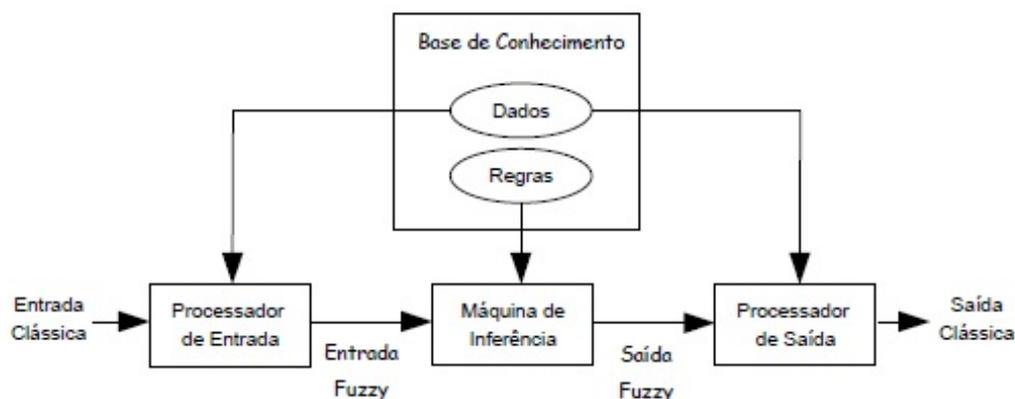
## 2.3 Sistema baseado em Regras Fuzzy

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*.

A lógica *fuzzy* é uma técnica que incorpora a forma humana de pensar, em um sistema de controle. Com o intuito de controlar tarefas, é possível traduzir termos linguísticos empregados por especialistas em fórmulas matemáticas, possibilitando a automação das mesmas. 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*.

Figura 6 – Estrutura básica de um sistema baseado em regras fuzzy.



Fonte: [Silveira \(2007\)](#).

A estrutura básica, de um sistema baseado em regras *fuzzy*, consiste em quatro módulos principais: um processador de entrada, uma base de conhecimento, uma máquina de inferência e um processador de saída, como está representado na [Figura 6](#).

## Processador de Entrada

Esse estágio é onde as entradas do sistema são modeladas por conjuntos *fuzzy*.

Identificadas as variáveis relevantes de entrada e saída do sistema e o intervalo de valores de cada variável, são atribuídos termos linguísticos que descrevem os estados destas variáveis. Estes termos são traduzidos - em um domínio apropriado - por uma função de pertinência, de 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.

## 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.

Este módulo pode ser considerado parte do “núcleo” do controlador *fuzzy*.

## Máquina de Inferência

O método de inferência determina a forma operacional do modelo linguístico. Ele propicia o cálculo da variável resposta (saída) a partir dos valores das variáveis de entrada, tratando-se de um mapeamento.

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 encaixa ao sistema que está sendo modelado. Os métodos mais utilizados são o método de Mamdani e KTS.

## 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 da saída de um modelo linguístico *fuzzy* de forma quantitativa, ou seja, ele nos fornece um valor numérico representativo que captura o significado essencial do conjunto *fuzzy* de saída do sistema (ORTEGA, 2001).

## 2.4 Probabilidade e Possibilidade

A teoria de conjuntos *fuzzy* é baseada no fato de que os conjuntos existentes no mundo real não possuem limites precisos. E, assim, torna-se necessário lidar com a existência de imprecisão, incerteza e definições qualitativas.

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, que são modeladas com base na teoria de probabilidade ou ainda, podem resultar da vagueza inerente aos próprios conceitos, as quais são modeladas com base na teoria *fuzzy*.

Nesse contexto, um conceito relevante é o de medida *fuzzy*. De acordo com Castanho (2005), uma medida *fuzzy* representa a incerteza com que se afirma que um dado elemento pertence a um conjunto clássico.

**Definição 2.15.** *Seja  $\Omega$  um conjunto não vazio e  $\mathcal{P}(\Omega)$  o conjunto das partes de  $\Omega$ . Uma*

medida fuzzy é uma função

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

que satisfaz as seguintes condições:

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

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

$$\Pi(X = x) = \pi_X(x) = \begin{cases} 1, & \text{se } x \in A \\ 0, & \text{se } x \notin A \end{cases} \quad (2.6)$$

Ortega (2001) afirma 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.

Ortega (2001) reafirma que, assim como ocorre com os conjuntos clássicos, quando uma variável linguística é definida, ela é restrita a um conjunto de valores. O que distingue 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 2.6, para um conjunto fuzzy, quando  $A$  é normal, é

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

Uma proposição *fuzzy* da forma, “ $X$  é  $A$ ”, onde  $X$  é uma variável que toma valores num universo de discurso  $U$  e  $A$  é um subconjunto *fuzzy* de  $U$ , induz a distribuição de possibilidade  $\Pi_X$ , 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  $U$ . Ou seja, se  $x$

$\in U$  e  $\varphi_A : U \rightarrow [0, 1]$  é a função de pertinência de  $A$ , então a possibilidade de que  $X = x$ , dada a proposição, “ $X$  é  $A$ ”, é

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

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

Dada uma distribuição de possibilidade  $\Pi_X$ , 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, segundo [Silveira \(2007\)](#).

**Definição 2.16.** *Uma medida de possibilidade  $\Pi$ , sobre  $\Omega$ , é uma função*

$$\Pi : \mathcal{P}(\Omega) \rightarrow [0, 1]$$

onde  $\mathcal{P}(\Omega)$  é o conjunto das partes de  $\Omega$  e tal que

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

**Definição 2.17.** *Uma medida de probabilidade  $P$  é uma função real que determina, para todo  $A$  em  $\Omega$ , uma probabilidade  $P(A)$ , tal que:*

1.  $P(A) \geq 0, \forall A \in \Omega$ ;
2.  $P(\Omega) = 1$ ;
3.  $P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$ , 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; CLAIR; YUAN, 1997). Isto é, dada uma medida de probabilidade  $P$  e uma medida de possibilidade  $\Pi$ , definidas em  $\mathcal{P}(\Omega)$ , 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), o grau de possibilidade de um evento é maior ou igual ao seu grau de probabilidade, ou seja,

$$P(A) \leq \Pi(A), \forall A \in \Omega,$$

onde  $P$  é a medida de probabilidade associada a  $p$ , e  $\Pi$  é a medida de possibilidade associada a  $\pi$ , respectivamente,

$$P(A) = \sum_{x \in A} p(x) \text{ e } \Pi(A) = \sup_{x \in A} [\pi(x)].$$

É essencial, para qualquer transformação probabilidade/possibilidade, obedecer o princípio de consistência acima. Klir, Clair e Yuan (1997), 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 linguísticos 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 sequências não crescentes. Então, as transformações mais simples,  $P \leftrightarrow \Pi$ , são baseadas na razão escalar  $\pi_i = p_i \alpha_i, \forall i$ , onde  $\alpha$  é uma constante positiva e são expressas pelas equações:

$$\pi_i = \frac{p_i}{p_1} \text{ e}$$

$$p_i = \frac{\pi_i}{\sum_{i=1}^n \pi_i}. \quad (2.7)$$

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 linguísticos, 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.

No próximo capítulo, apresentaremos dois modelos *fuzzy* para analisar o risco de desenvolvimento de metástase e risco de morte para o câncer renal. Exibimos, também, uma abordagem estatística com dados de pacientes do Hospital das Clínicas - UNICAMP que contribuiu para a construção dos modelos e analisar a relação entre grau de Fuhrman e o prognóstico dos indivíduos.

### 3 Modelagem Matemática Fuzzy para Análise de Risco de Metástase e Risco de Morte

A natureza imprecisa e, muitas vezes contraditória, da informação ou a sua falta é uma realidade e devem ser consideradas em toda decisão médica, pois podem ter consequências importantes, ou mesmo vitais, para o paciente. A teoria dos conjuntos *fuzzy* foi desenvolvida para lidar com o conceito de verdade parcial, ou seja, com valores de verdade entre o completamente verdadeiro e o completamente falso da lógica Booleana. Assim, dificilmente podemos considerar um indivíduo completamente doente, uma vez que algumas funções (ou a maioria delas) permanecem perfeitas ou por outro lado, não podemos nos considerar completamente saudáveis, pois quase sempre estamos resfriados, estressados ou mal alimentados (SAVERGNINI, 2009).

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 tanto complexo, baseado em conhecimento expresso linguisticamente e na presença de diversos tipos de incerteza e imprecisão.

Diante disso, a teoria dos conjuntos *fuzzy* é apropriada para desenvolver sistemas especialistas para auxiliar os médicos em suas decisões (CASTANHO, 2005).

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 Computação Inteligente (SILVEIRA, 2007). A informática médica acompanha esta evolução e o computador auxilia o médico nas mais diversas atividades. Segundo Ortega (2001), 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.

A teoria de conjuntos *fuzzy* tem sido utilizada frequentemente como uma alternativa no processo de tomada de decisão, por ser considerada adequada às reais necessidades clínicas e experimentais. Adlassnig (1980) 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 (ORTEGA, 2001) e evolução da AIDS (JAFELICE, 2003).

A intenção, nesse capítulo, é fazer modelagem *fuzzy*, analisando riscos de desenvolvimento de metástase e de morte, mostrando e discutindo informações obtidas na construção dos modelos, e apresentando algumas simulações de pacientes e os resultados obtidos. Ressaltamos que toda a implementação do modelo foi feita no *MATLAB*<sup>®</sup> 2015.

Para auxiliar na construção dos modelos foi feita uma análise estatística por meio de regressão logística. O estatístico professor Sidney Ragazzi nos ajudou nessa fase. O estudo contribuiu também para averiguar a associação entre o grau de Fuhrman e o prognóstico de pacientes. Os dados dos pacientes do Hospital das Clínicas - UNICAMP serviram de base para esse estudo. Os comentários e resultados encontram-se na próxima Subseção.

### 3.1 Abordagem Estatística: Grau de Fuhrman x Surgimento de Metástase e Demais Variáveis x Surgimento de Metástase

O aumento de incidência de neoplasias renais se elevou com a realização de exames para diagnóstico e de outras condições, desta forma, aumentando o diagnóstico de pequenos tumores.

Mediante contato com o médico urologista Prof. Dr. Ubirajara Ferreira da equipe médica do Hospital das Clínicas - UNICAMP, e diante de questionamentos por parte dele e de seus companheiros de equipe, surgiu a ideia de estudar e analisar a relação entre a graduação de Fuhrman e a evolução do câncer renal. Esse questionamento se baseou na ocorrência de grande quantidade de casos reais da doença que se verificou no Hospital retratados com graus baixos de Fuhrman, mas apresentaram metástases em seu seguimento (após a cirurgia) ou mesmo a tempo de cirurgia.

Juntamente com as primeiras pesquisas, entramos em contato com o médico patologista Prof. Dr. Athanase Billis, do Hospital das Clínicas (HC), da UNICAMP, que nos disponibilizou laudos de pacientes sobre tumores do subtipo células claras passados por nefrectomia<sup>1</sup>, a fim de fazer um levantamento e digitalizar (construir uma tabela) os dados que seriam utilizados para alimentar o banco de dados do Hospital. Para essa tarefa, contamos com o auxílio da médica patologista Dra. Larissa Eloy também do Hospital das Clínicas que nos orientou e contribuiu para sanar as dúvidas que surgiram. Tais informações foram úteis ao nosso trabalho.

Foram coletados dados sobre o tumor e na construção da tabela pedida pelo Dr.

---

<sup>1</sup> Nefrectomia é a remoção cirúrgica do rim.

Athanasios levou-se em conta duas classificações para o grau de Fuhrman: Fuhrman único e Fuhrman combinado. A classificação combinada foi motivada pelos estudos anteriores de tumor de bexiga, em que para essa neoplasia tal classificação se mostrou realista (BILLIS, ). O que ocorre em muitos casos de neoplasias renais, é a presença de dois graus de Fuhrman em um único tumor. Ou seja, o tumor apresenta em sua maior extensão um determinado grau, contudo em um pequeno foco pode apresentar um grau mais agressivo. O Fuhrman único é descrito pelo pior grau presente na extensão, mesmo que esteja presente em uma pequena parte da massa renal. O grau combinado leva em conta a soma dos graus encontrados, sugerindo ser uma maneira mais fidedigna de representar os graus existentes. Para sabermos qual dos graus era mais preciso na análise prognóstica dos pacientes, foi feito um estudo estatístico por meio de regressão logística.

Como as informações encontradas nos laudos eram características histopatológicas do tumor (tais como grau de Fuhrman, tamanho do tumor, se tem envolvimento de veias, linfonodos e/ou glândula suprarrenal, estadiamento, presença de necrose) e não eram suficientes para o modelo, uma vez que as informações sobre a evolução da doença (se permaneceu estável, ou se evoluiu para metástase ou se houve óbito pela neoplasia) não constavam nos laudos, buscamos esses dados que se encontravam nos prontuários dos pacientes. Nessa etapa, os médicos que nos ajudaram com acesso ao SAM - Serviço de Arquivo Médico do Hospital das Clínicas - UNICAMP foram Prof. Dr. Wagner Eduardo Matheus e o então residente, Dr. Eduardo Azevedo. Com esse levantamento de dados foram construídas as Tabelas 3 e 4.

Tabela 3 – Distribuição de 148 carcinomas de células renais de acordo com a área de mais alto grau e estadiamento.

Grau	Estadiamento									
	TX		T1		T2		T3		T4	
	n	%	n	%	n	%	n	%	n	%
1	3	30	27	41,54	4	15,38	5	10,87	-	-
2	6	60	30	46,15	15	57,69	19	41,3	1	100
3	1	10	8	12,31	5	19,23	19	41,3	-	-
4	-	-	-	-	2	7,69	3	6,52	-	-

Fonte: Tabela construída pela autora.

No fim da coleta de dados, conseguimos reunir as seguintes informações: idade e sexo dos pacientes; data da nefrectomia e tipo (radical ou parcial); lado do rim retirado, peso e tamanho do mesmo; se o tumor é composto por nódulo único ou múltiplo, e seu local no rim; grau de Fuhrman; tamanho do tumor; se apresenta invasão vascular, em veia cava, em seio renal, em tecido adiposo perirrenal, em ureter e/ou glândula suprarrenal; estadiamento; presença de necrose; diferenciação tumoral; se tem presença de êmbolos ou trombos neoplásicos; tipo de arranjo; margem cirúrgica de ressecção e linfonodos livres ou

Tabela 4 – Distribuição de 148 carcinomas de células renais de acordo com o sistema de números combinados e estadiamento.

Grau	FS	Estadiamento									
		TX		T1		T2		T3		T4	
		n	%	n	%	n	%	n	%	n	%
1+1	2	3	30	27	41,54	4	15,38	5	10,87	-	-
1+2	3	6	60	20	30,77	9	34,62	13	28,26	-	-
2+2	4	-	-	10	15,38	6	23,08	6	13,04	1	100
2+3	5	1	10	7	10,77	4	15,38	10	21,74	-	-
3+3	6	-	-	1	1,54	1	3,85	9	19,57	-	-
3+4	7	-	-	-	-	1	3,85	2	4,35	-	-
4+4	8	-	-	-	-	1	3,85	1	2,17	-	-

Fonte: Tabela construída pela autora.

comprometidos pela neoplasia; se apresentou sintomas e quais foram; se houve metástase a tempo de cirurgia ou durante o seguimento, e local da mesma; se houve recidiva, quando ocorreu e qual a conduta utilizada; ou se houve óbito e qual o tempo de seguimento de cada paciente em meses (tempo que o paciente seguiu com o tratamento após a cirurgia até a última consulta nos ambulatórios do Hospital das Clínicas). O estadiamento, assim como diversos outros parâmetros, se encontravam ausentes em determinados prontuários. Como esse fator em específico é considerado importante para o estudo dessa doença e os dados que são utilizados para fazer essa classificação se encontravam na planilha construída com dados dos laudos, conseguimos obter a classificação dos casos que faltavam essa informação.

De porte dos dados, contamos com a ajuda do estatístico professor Sidney Ragazzi que fez uma análise uni e multivariada, colaborando para definirmos as variáveis para o modelo e para os pesos das regras. Ele utilizou os dados de Fuhrman único e combinado, estadiamento, tamanho do tumor e presença de necrose. Essas variáveis foram escolhidas com base em referências na área e consultas a especialistas. A análise levou em conta o surgimento de metástase ou a morte pela doença sem considerar o tempo para a sua ocorrência, ou seja, todos os casos reais de metástase e/ou morte foram considerados no estudo independente do tempo de ocorrência dos fatos.

O estudo estatístico mostrou que o Fuhrman combinado apresentou os mesmos resultados quando comparamos com o Fuhrman único no que diz respeito ao desenvolvimento de metástase, isto é, não há um grau mais adequado para descrever a relação com metástase. O estudo mostrou também que existe uma associação entre os casos de Fuhrman combinado x metástase e Fuhrman único x metástase sendo que o grau dessas associações são de baixa a moderada, não podendo ser dito que é determinante ter um Fuhrman alto para se desenvolver metástase, ou seja, existe uma dependência entre essas variáveis, porém não é tão significativa. Vale a pena ressaltar que não há diferença entre

as duas classificações de Fuhrman com relação ao prognóstico dos pacientes.

Para embasar esses resultados serão apresentados a seguir os fundamentos estatísticos.

O Teste Qui-Quadrado (correção de Yates) foi utilizado para verificar se os eventos envolvidos possuem relação entre si. Para o trabalho foram feitos testes com Fuhrman único e Fuhrman combinado. As hipóteses testadas para Fuhrman único foram as seguintes:

$$\begin{cases} H_0 : \text{As variáveis Fuhrman Único e Metástase/Morte são independentes} \\ H_1 : \text{As variáveis Fuhrman Único e Metástase/Morte não são independentes.} \end{cases}$$

Como o valor encontrado para  $P - Valor = 0,000629006$  é menor do que  $\alpha = 0,05$  (5%), rejeita-se  $H_0$ . Ou seja, as variáveis Fuhrman Único e Metástase/Morte não são independentes.

Desta forma, podemos medir a associação entre essas variáveis. As medidas de associação são utilizadas com a finalidade de medir o nível de associação entre variáveis qualitativas nominais.

1. Coeficiente de contingência quadrático médio: usa o próprio valor da estatística  $Q_{obs}^2$  como medida de associação, mas como  $Q_{obs}^2$  depende de  $n$  (tamanho da amostra), implica que a medida é crescente em  $n$ . Para se obter uma medida não afetada pela dimensão da amostra, definimos este coeficiente

$$\Phi^2 = \frac{Q_{obs}^2}{n}$$

Desvantagens: Não é limitado superiormente.

2. Coeficiente de contingência proposto por Pearson (1904)

$$P = \sqrt{\frac{\frac{Q_{obs}^2}{n}}{1 + \frac{Q_{obs}^2}{n}}} = \sqrt{\frac{\frac{Q_{obs}^2}{n}}{n + Q_{obs}^2}}$$

Obtemos valores entre 0 e 1, atingindo o seu valor mínimo se, e somente se  $Q_{obs}^2 = 0$ , ou seja, no caso de completa independência. O valor máximo não é, em geral atingido. Mesmo em casos de completa dependência, o valor é modificado com o número de linhas e de colunas da tabela de dados.

A próxima medida soluciona este problema.

3. Coeficiente de Tschuprov

$$T = \sqrt{\frac{\frac{Q_{obs}^2}{n}}{(r-1)(c-1)}}$$

Tem valor 0 no caso de independência e consegue atingir o valor máximo 1, no caso de dependência completa, mas apenas se  $r = c$  (número de linhas igual ao número de colunas).

4. Coeficiente de Cramér

$$C = \sqrt{\frac{\frac{Q_{obs}^2}{n}}{\min\{r-1; c-1\}}}$$

O valor deste coeficiente fica próximo de 1 em caso de forte dependência e próximo de 0 em caso de independência. O coeficiente de Cramér não é limitado superiormente.

Para quaisquer dos coeficientes apresentados existem expressões para as suas variâncias. A grande desvantagem destes coeficientes é que não permitem interpretações probabilísticas como as que se fazem com o coeficiente de correlação, por exemplo, mas nos dão uma boa ideia do nível de associação entre as variáveis.

Quanto maior o valor do coeficiente, maior o nível de associação entre as variáveis.

A conclusão para a relação entre Fuhrman único e metástase foi a seguinte: Pelo coeficiente de contingência quadrático médio ( $\Phi^2 = 0,220278338$ ) temos uma associação baixa. Pelo coeficiente de contingência ( $P = 0,42487037$ ) temos uma associação moderada. Para as demais, Coeficiente de Tschuprov ( $T = 0,156446064$ ) e Coeficiente de Cramér ( $C = 0,073426113$ ), temos uma baixa associação.

Concluimos que há uma associação mais para baixa do que para moderada.

O mesmo teste foi usado para analisar o grau de Fuhrman combinado. O teste Qui-Quadrado testou as seguintes hipóteses:

$$\begin{cases} H_0 : \text{As variáveis Fuhrman Combinado e Metástase/Morte são independentes} \\ H_1 : \text{As variáveis Fuhrman Combinado e Metástase/Morte não são independentes.} \end{cases}$$

Como o  $P - Valor = 0,0000570487$  é menor do que  $\alpha = 0,05$  (5%), rejeita-se  $H_0$ . Ou seja, as variáveis Fuhrman Combinado e Metástase/Morte não são independentes.

Calculando as mesmas medidas de associação agora com relação ao Fuhrman combinado a conclusão foi a seguinte: Pelo coeficiente de contingência quadrático médio ( $\Phi^2 = 0,38486109$ ) temos uma associação moderada. Pelo coeficiente de contingência ( $P = 0,527167825$ ) temos uma associação mais acentuada do que a anterior. Para as demais, temos uma baixa associação: Coeficiente de Tschuprov ( $T = 0,14622302$ ) e Coeficiente de Cramer ( $C = 0,12828703$ ).

Concluimos que há uma associação de baixa à moderada.

Além disso, também foi feita análise através do método de regressão logística. Consideramos o odds ratio, que é uma medida para descrever a força da associação entre dois valores de dados binários. Ele é usado como uma estatística descritiva e desempenha um papel importante na regressão logística.

Mediante os dados coletados, as conclusões foram as seguintes: odds ratio estimado de 2,09 implica que as chances de metástase aparecer depois da cirurgia é cerca de 2,09 vezes maior do que não aparecer quando Fuhrman combinado aumenta de uma unidade. O odds ratio estimado de 2,07 implica que as chances de metástase estar presente a tempo de cirurgia é cerca de 2,07 vezes maior do que não ter metástase quando Fuhrman combinado aumenta de uma unidade.

Agora, se considerarmos o grau de Fuhrman único, temos o odds ratio estimado de 4,41 o que implica que as chances de metástase aparecer depois da cirurgia é cerca de 4,41 vezes maior do que não aparecer quando Fuhrman único aumenta de uma unidade. O odds ratio estimado de 3,96 implica que as chances de metástase estar presente a tempo de cirurgia é cerca de 3,96 vezes maior do que não ter metástase quando Fuhrman único aumenta de uma unidade.

Esses valores foram obtidos quando consideramos os casos de morte pela doença. Os valores, quando analisamos o surgimento de metástase, são inferiores à esses. Devido a isso, o grau de Fuhrman único se mostrou mais preciso e eficaz para o fenômeno estudado.

Em linhas gerais, as duas classificações para o Fuhrman (único e combinado) indicam a mesma chance de desenvolvimento de metástase, isto é, os dois graus corroboram os mesmos resultados quando analisamos o modelo com metástase. Já considerando o modelo com morte, o Fuhrman único evidencia mais os casos de metástase e morte, sugerindo que o Fuhrman único é mais compatível que o Fuhrman combinado.

Portanto, diante desses fatos, decidimos construir um modelo *fuzzy* considerando o grau de Fuhrman único.

Além do estudo com os graus, também foi feita análise estatística para outros fatores: estadiamento, presença de necrose e tamanho do tumor. A pesquisa revelou que todos os fatores possuem dependência com os casos de metástase e de morte, contudo a presença de necrose e o estadiamento foram os fatores principais que mais influenciam no prognóstico dos pacientes.

A seguir tem-se os resultados que justificam as observações comentadas.

Neste caso, o Teste Qui-Quadrado também foi útil. Para os demais fatores, estadiamento, presença de necrose e tamanho do tumor, foram encontrados os seguintes p-valores respectivamente,  $P - Valor = 0,000119$ ,  $P - Valor = 7,50013E^{-05}$  e  $P - Valor = 0,000290115$ . Como todos são inferiores a  $\alpha = 0,05$  (5%), pode-se observar a dependência de cada um com o fato de ocorrer metástase ou morte pela doença.

As medidas de associação encontradas foram as seguintes:

#### 1. Estadiamento

Pelo coeficiente de contingência quadrático médio ( $\Phi^2 = 0,368789$ ) temos uma associação baixa. Pelo coeficiente de contingência ( $P = 0,519064$ ) temos uma associação moderada. Para as demais, temos uma baixa associação: Coeficiente de Tschuprov ( $T = 0,143137$ ) e Coeficiente de Cramér ( $C = 0,12293$ ).

Concluimos que há uma associação de baixa para moderada.

#### 2. Necrose

Pelo coeficiente de contingência quadrático médio ( $\Phi^2 = 0,164459142$ ) temos uma associação baixa. Pelo coeficiente de contingência ( $P = 0,375808745$ ) temos uma associação moderada. Para as demais, temos uma baixa associação: Coeficiente de Tschuprov ( $T = 0,234136101$ ) e Coeficiente de Cramér ( $C = 0,164459142$ ).

Concluimos que há uma associação mais para baixa do que para moderada.

#### 3. Tamanho do Tumor

Pelo coeficiente de contingência quadrático médio ( $\Phi^2 = 0,235235636$ ) temos uma associação baixa. Pelo coeficiente de contingência ( $P = 0,43639187$ ) temos uma associação moderada. Para as demais, temos uma baixa associação: Coeficiente de Tschuprov ( $T = 0,161670322$ ) e Coeficiente de Cramér ( $C = 0,078411879$ ).

Concluimos que há uma associação mais para baixa do que para moderada.

O método de regressão logística revelou as informações a seguir.

O odds ratio estimado de 5,20 implica que as chances de metástase/morte ocorrer depois da cirurgia é mais de cinco vezes maior do que não ocorrer quando necrose aumenta de uma unidade (ou seja, passa de 0 para 1) mantendo os outros fatores constantes. E o odds ratio estimado de 1,84 implica que as chances de metástase/morte ocorrer depois da cirurgia é 1,8 vezes maior do que não ocorrer quando estadiamento aumenta de uma unidade, mantendo os outros fatores constantes. Enquanto que, os coeficientes de tamanho do tumor e grau de Fuhrman se mostraram não significativos.

O odds ratio estimado de 9,42 implica que as chances de metástase/morte ocorrer a tempo de cirurgia é mais de nove vezes maior do que não ocorrer quando necrose aumenta de uma unidade (ou seja, passa de 0 para 1) mantendo os outros fatores constantes. E o odds ratio estimado de 2,35 implica que as chances de metástase/morte ocorrer a tempo de cirurgia é 2,35 vezes maior do que não ocorrer quando estadiamento aumenta de uma unidade, mantendo os outros fatores constantes. Novamente, os coeficientes de tamanho do tumor e grau de Fuhrman se mostraram não significativos.

Assim, pode-se ver que houve confirmação de que os fatores considerados afetam o prognóstico e devem ser levados em conta para programar condutas aos pacientes.

Todas essas observações nos nortearam para o desenvolvimento dos modelos apresentados na próxima Subseção.

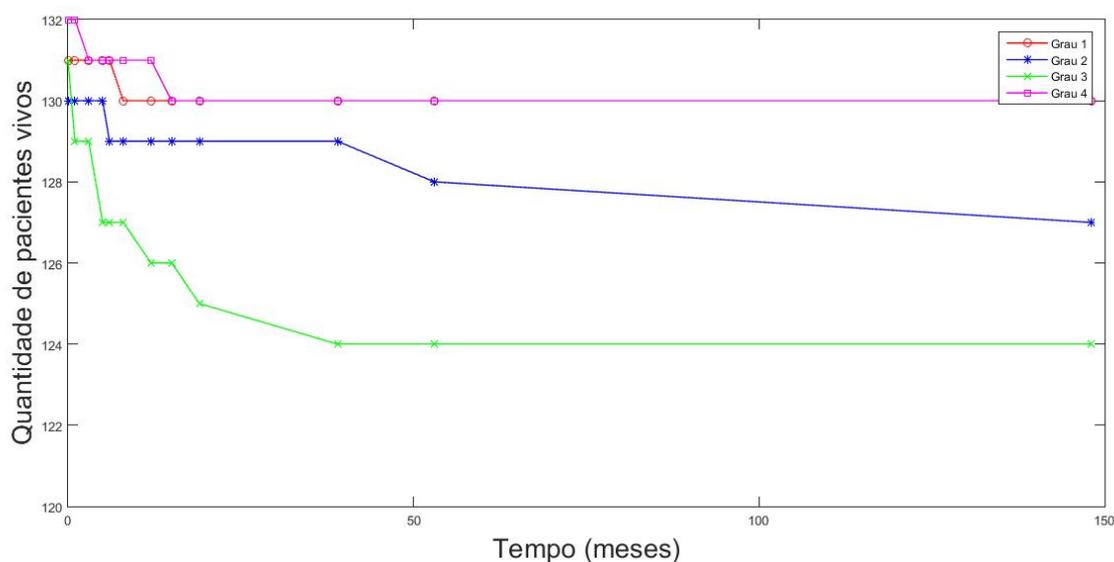
Encontramos na literatura alguns gráficos denominados curvas de sobrevida. Essas curvas nada mais são do que a representação gráfica da Função de sobrevida  $S(t)$  no eixo vertical vs. os tempos de sobrevida ( $t$ ) no eixo horizontal. Motivados com isso, construímos um gráfico com os dados de Fuhrman coletados no Hospital das Clínicas para analisar a sobrevivência dos pacientes ao longo do tempo de seguimento (o maior tempo encontrado nos dados). Podemos ver esse gráfico na [Figura 7](#). Vale lembrar que não é uma curva de sobrevida, é apenas uma representação gráfica comparando grau de Fuhrman e sobrevivência dos pacientes.

Ao observar o gráfico, nota-se que apesar do número de pacientes com grau 4 seja pequeno, o prognóstico para Fuhrman 4 é igual o de Fuhrman 1, no sentido de pacientes com grau de Fuhrman 1 vivem tanto quanto os de Fuhrman 4. Isso atesta a incompatibilidade dessa Classificação.

## 3.2 Modelagem Fuzzy

Tendo em vista a desconfiança médica a respeito da efetividade do prognóstico baseado no Grau de Fuhrman, surgiu a motivação para estudar o comportamento da

Figura 7 – Gráfico para análise do Grau de Fuhrman: Tempo(meses) x Número de pacientes vivos após nefrectomia.



doença segundo esse critério. Diante disso, trabalhamos não somente com a graduação de Fuhrman, mas também outras variáveis para os modelos que desenvolvemos.

Como já visto anteriormente, 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*).

### 3.2.1 Primeiro Modelo

Na modelagem fuzzy, as variáveis de entrada - Estadiamento, grau de Fuhrman, presença de necrose e tamanho do tumor - e a variável de saída - Risco de Metástase - foram consideradas como variáveis linguísticas e seus valores como conjuntos *fuzzy*, em seus respectivos domínios.

As variáveis quantitativas, contínuas por natureza, são fuzzificadas de maneira direta, enquanto que as variáveis qualitativas são arranjadas numa escala entre 0 e 1.

A Figura 8 ilustra o esquema do SBRF que utilizamos.

Para a variável linguística *Estadiamento* foram atribuídos sete estágios, traduzidos por conjuntos *fuzzy*, com base na Tabela TNM (conferir Tabela 2). Os estágios T3b e T3c foram considerados apenas como T3b", pois o que os difere é invasão da veia cava

Figura 8 – Estrutura básica do sistema baseado em regras *fuzzy*.

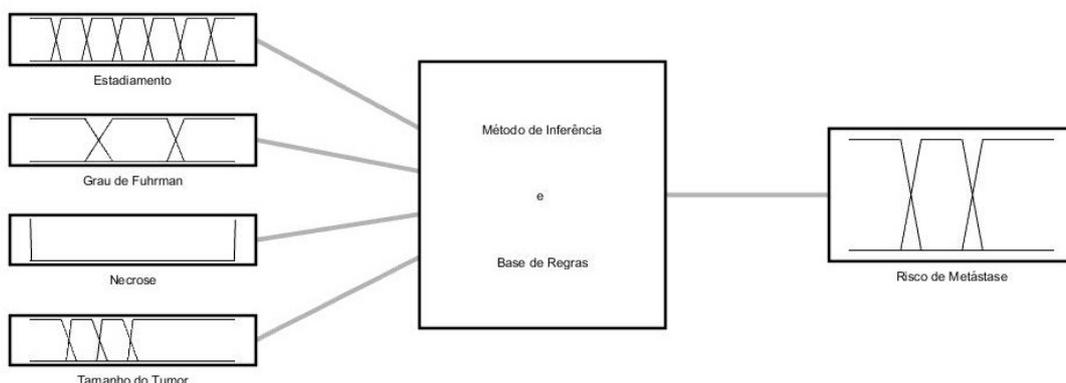
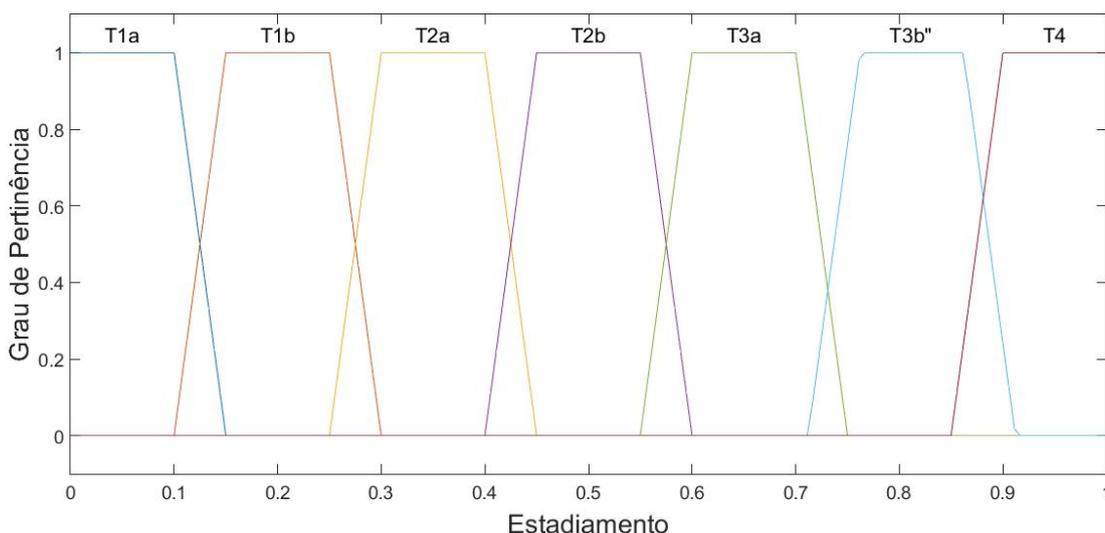


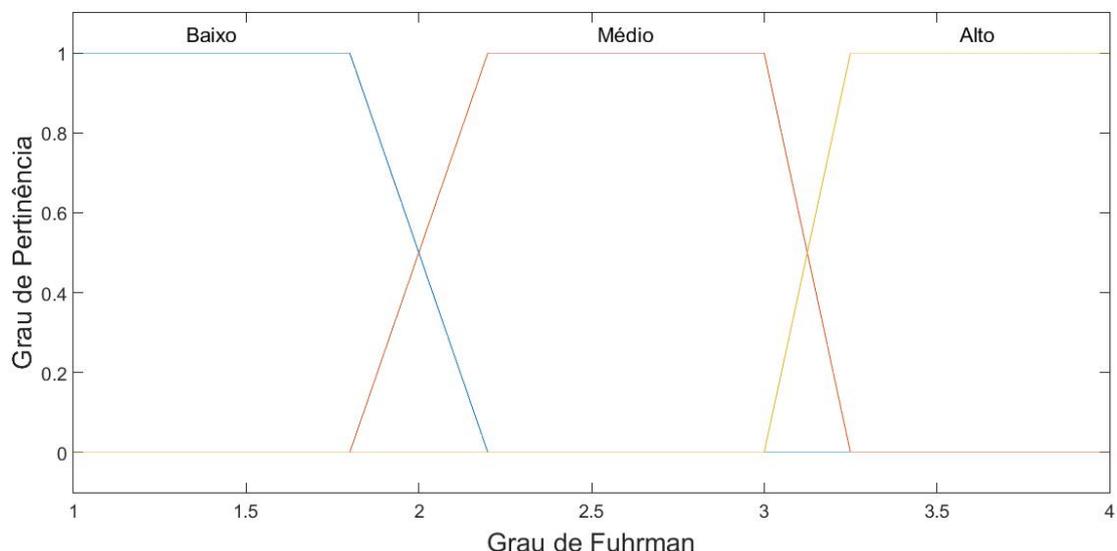
Figura 9 – Funções de pertinência dos conjuntos *fuzzy*, assumidos pela variável linguística Estadiamento.



abaixo do diafragma no estágio T3b e a invasão da mesma veia acima do diafragma no caso T3c. A Figura 9 mostra a representação gráfica dessa variável.

O *estadiamento clínico* relaciona-se com o tamanho do tumor e o envolvimento de órgãos, portanto, o domínio  $[0, 1]$  representa o percentual de comprometimento do rim, ou seja, 0 representa um rim sem tumor, enquanto 1 representa o rim tomado pelo tumor e espalhado para outros órgãos além do rim. Nos estágios T1 e T2, leva-se em conta o tamanho do tumor, a partir do estágio T3, o tumor já não se localiza apenas no rim, mas espalhou e pode ter comprometido veias, ou glândula suprarrenal, ou tecido adiposo perirrenal, ou outros órgãos.

Figura 10 – Funções de pertinência dos conjuntos *fuzzy*, assumidos pela variável linguística Grau de Fuhrman.



No sistema de graduação clássico aconselhado por Fuhrman (FUHRMAN; LASKY; LIMAS, 1982), o grau 1 indica um prognóstico favorável, os graus 2 e 3 sugerem um prognóstico intermediário e o grau 4 indica um prognóstico pior. Assim, para a variável *Grau de Fuhrman*, utilizamos os termos linguísticos Baixo, Médio e Alto, como ilustrado na Figura 10.

A variável *Presença de Necrose* foi classificada por funções triangulares, devido a dificuldade, por parte do patologista, em determinar qual o tamanho da região tumoral necrosada ao analisar o material recolhido. Desta forma, atribuímos 0 para a ausência de necrose e 1 no caso de presença, como podemos observar na Figura 11.

A última variável considerada foi *Tamanho do Tumor*, por ser amplamente encontrada em bibliografias como fator prognóstico importante (SEGRE, 2010) (POMPEO, 2006). Essa variável foi discriminada da seguinte maneira: Até 4 cm, de 4 a 7 cm, de 7 a 10 cm e maior que 10 cm. Para essas subdivisões, atribuímos os seguintes termos linguísticos, Baixo, Médio, Médio Grande e Grande, que podem ser vistos na Figura 12.

Para a variável de saída - *Risco de Metástase* - foram atribuídos os seguintes termos: Baixo, Médio e Alto. As funções de pertinência dos conjuntos *fuzzy* foram construídas de forma trapezoidal, como mostra a Figura 13. É interessante aos especialistas saber o valor preditivo da doença e, portanto, analisar o risco de desenvolver metástase é algo relevante.

A base de regras é construída, normalmente, de acordo com as informações de

Figura 11 – Funções de pertinência dos conjuntos *fuzzy*, assumidos pela variável linguística Presença de Necrose.

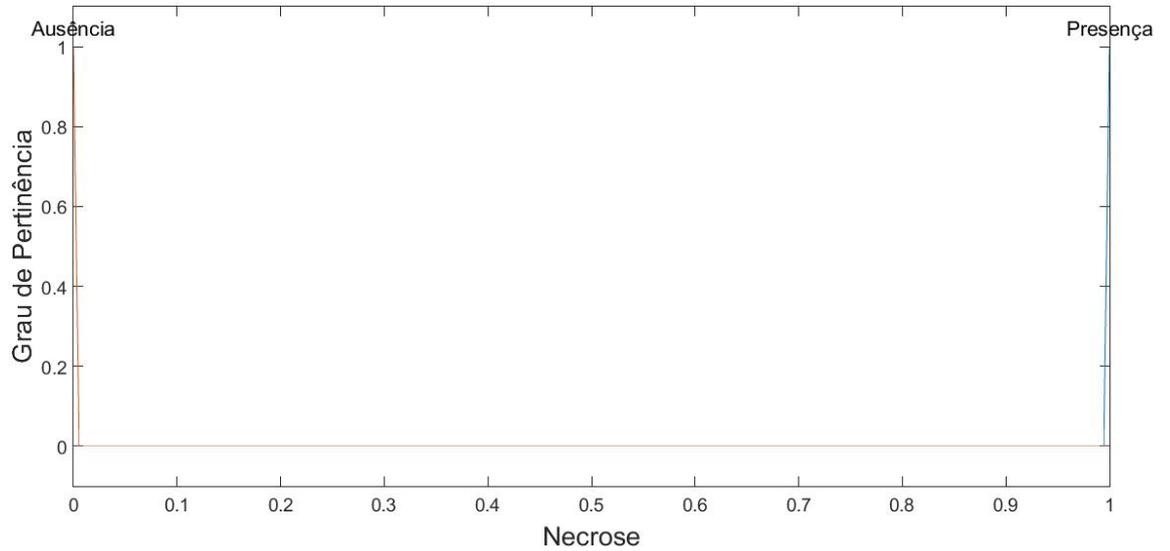


Figura 12 – Funções de pertinência dos conjuntos *fuzzy*, assumidos pela variável linguística Tamanho do Tumor.

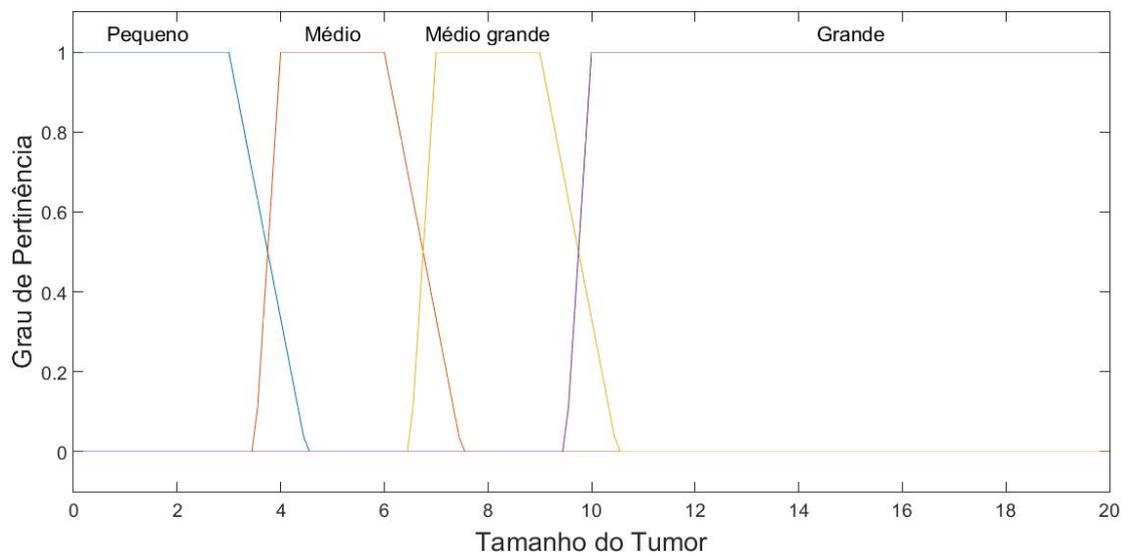
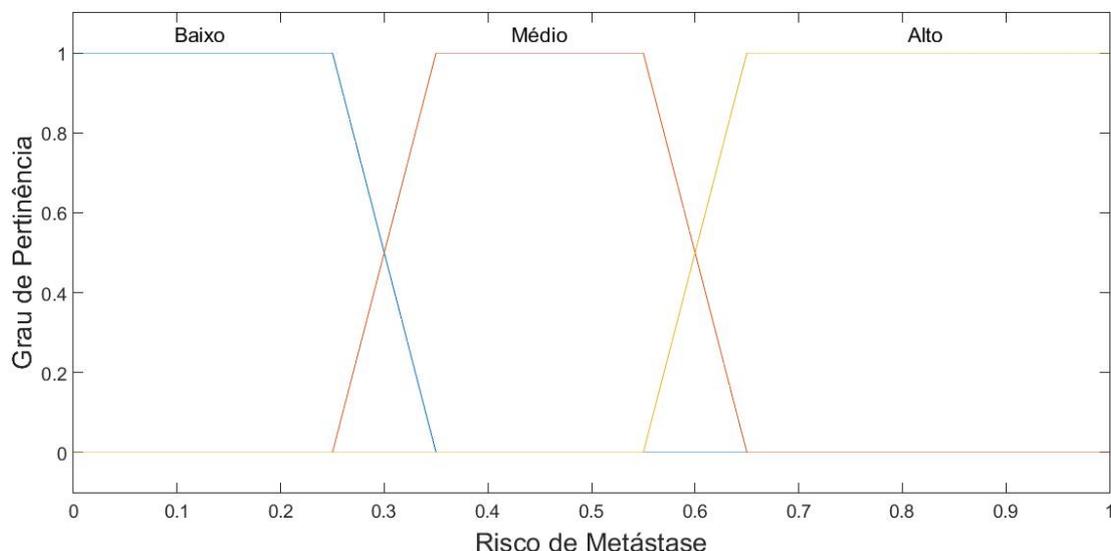


Figura 13 – Funções de pertinência dos conjuntos *fuzzy*, assumidos pela variável linguística Risco de Metástase.



um especialista. Este conjunto de regras *fuzzy* desempenha o papel da função matemática, para obter a saída do sistema.

Neste trabalho, a base de regras foi elaborada a partir de consultas ao especialista prof. Dr. Athanase Billis (HC) que sugeriu considerar, além da principal variável - *grau de Fuhrman* - o *estadiamento* e a *presença de necrose*. Com base em referências na área, decidimos considerar a variável *tamanho do tumor*. Vale observar que o trabalho de Segre (2010) trata da confecção de um nomograma para predição prognóstica em pacientes brasileiros com CCR, nele o estudo é feito considerando vários tipos celulares e é possível concluir que o tamanho tumoral é uma variável muito importante para a predição da doença. Além desse, outros trabalhos na literatura contribuíram para a escolha dessa variável (ORNELLAS, 2012) (MEDEIROS; GELB; WEISS, 1988) (POMPEO, 2006).

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

Suponha que um paciente apresenta um tumor com *Estadiamento* T1a, *Grau de Fuhrman* igual a 1, com *Necrose* presente e *Tamanho do tumor* igual a 3 cm. Com essas informações foi construída a regra número 13 da Tabela 5:

“Se *Estadiamento* é T1a e *Grau de Fuhrman* é Baixo e *Necrose* é Presente e *Tamanho do Tumor* é Pequeno então o *Risco de Metástase* é Baixo.”

De modo análogo, isto é, fazendo todas as diferentes combinações entre es-

tadiamento clínico, grau de Fuhrman, se há presença ou não de necrose e tamanho do tumor e levando em conta os termos linguísticos atribuídos à cada uma dessas variáveis, foi construída a base de regras, formada por 504 regras.

As regras foram ponderadas com base, principalmente, em dados obtidos dos pacientes do Hospital das Clínicas - UNICAMP e também em estudo estatístico que apontou estadiamento e necrose como fatores mais determinantes para a sobrevivência dos pacientes. Foi feita uma análise multivariada, através do método de regressão linear, com a ajuda do professor Sidney, ex-docente do IMECC/UNICAMP. Esta análise foi necessária devido a dificuldade dos especialistas ponderarem as regras. Assim, os resultados obtidos estão baseados na casuística de 132 pacientes do Hospital das Clínicas da UNICAMP. Esse número é a quantidade de pacientes em que todos os dados analisados estavam completos. Desta forma, procuramos atribuir pesos maiores às regras em que havia presença de necrose e conforme o aumento no estadiamento.

Como exemplo de ponderação das regras, suponham os casos:

Se Estadiamento é T3a e Grau de Fuhrman é Baixo e Necrose é ausente e Tamanho do Tumor é Médio então temos três situações para o Risco de Metástase: Baixo, Médio e Alto. Para as três regras desse exemplo foram atribuídos os seguintes pesos, respectivamente, 0.4, 0.3 e 0.3. Agora, se consideramos que Estadiamento é T3b e Grau de Fuhrman é Baixo e Necrose é Presente e Tamanho do Tumor é Médio então temos, novamente, três situações para o Risco de Metástase: Baixo, Médio e Alto. Para esse caso, os pesos atribuídos seguiram-se desta forma: 0.2 para risco Baixo, 0.3 para Médio e 0.5 para Alto. Isso revela que regras com pesos maiores são mais relevantes e podem ocorrer com frequência maior.

Assim, por exemplo, para um paciente cujo *Estadiamento Clínico* é T2a, *Grau de Fuhrman* é baixo (1), *Necrose* é presente e *Tamanho do tumor* é médio grande (8 cm), foram construídas as regras 163 e 164 (ver [Tabela 5](#)).

A regra 160 tem peso igual a 0,5; já as regras 161 e 162 têm pesos iguais a 0,2 e 0,3, respectivamente. Isso indica que com esses dados clínicos, a chance do risco do tumor desenvolver metástase ser baixo é maior do que as outras chances e, por sua vez, a chance do risco ser médio é ligeiramente menor do que ser alto.

Definida a base de regras, precisamos *traduzi-la* matematicamente, como vimos na Seção 2.3, por meio das técnicas da lógica *fuzzy*, com a finalidade 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

Tabela 5 – Algumas das 504 regras, da Base de Regras construída. Na 1<sup>a</sup> coluna consta o  $n^o$  da regra e, na última, o peso utilizado para ponderá-la.

N	ESTAD.	FUHRMAN	NEC.	TAM.	RISCO META.	PESO
1	T1a	Baixo	Ausência	Pequeno	Baixo	1
13	T1a	Baixo	Presença	Pequeno	Baixo	0.8
54	T1a	Alto	Ausência	Médio	Alto	0.2
101	T1b	Médio	Ausência	Médio	Médio	0.3
160	T2a	Baixo	Presença	Médio	Baixo	0.5
161	T2a	Baixo	Presença	Médio	Médio	0.2
162	T2a	Baixo	Presença	Médio	Alto	0.3
247	T2b	Médio	Ausência	Médio grande	Baixo	0.4
324	T3a	Médio	Ausência	Grande	Alto	0.7
431	T3b"	Alto	Presença	Grande	Médio	0.2
432	T3b"	Alto	Presença	Grande	Alto	0.7
468	T4	Médio	Ausência	Grande	Alto	0.8

pelo Método do Centro de Gravidade, para que, enfim, a saída obtida fosse um número real. Esse valor é estabelecido no intervalo  $[0, 1]$  e a ele corresponde uma pertinência aos conjuntos *fuzzy*: Baixo, Médio e Alto, que representam o risco da doença desenvolver metástase (ver Figura 13).

### 3.2.1.1 Simulações

A partir dos dados de pacientes com câncer de rim, é possível determinar o grau de compatibilidade com cada conjunto *fuzzy* da variável de saída *Risco de Metástase*.

Foram feitas simulações deste primeiro modelo com o objetivo de verificar a compatibilidade dos resultados gerados pelo sistema. Os dados considerados são de pacientes do Hospital das Clínicas da UNICAMP, obtidos em setembro e outubro de 2016.

Com o intuito de estabelecer uma relação entre pertinência ao conjunto e probabilidade, encontramos uma forma de descrever os resultados encontrados (Possibilidade) por meio de probabilidade de ocorrência em cada conjunto *fuzzy* que é a forma mais usual na área médica para relatar informações.

### Relação Possibilidade/Probabilidade

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

$$P(A) \leq \Pi(A), \forall A \in \Omega,$$

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

A [Tabela 6](#) contém alguns resultados, obtidos em termos de possibilidade e transformados em probabilidades, através da [Equação 2.7](#), vista na [Seção 2.4](#).

Assim, pela [Tabela 6](#), podemos ver que um paciente com tumor no estágio T3a, isto é, tumor invadindo a gordura do seio renal, a biópsia revela grau de Fuhrman médio (2), sem necrose e tamanho médio (6,5cm), tem pertinências aos conjuntos Médio e Alto para Risco de Metástase. Já outro paciente com tumor no mesmo estágio (T3a) e grau de Fuhrman (Médio), mas com presença necrose e tamanho maior (Grande - 11,5cm) possui maior pertinência ao conjunto Alto de Risco de Metástase.

De modo geral, os graus de pertinência estão de acordo com a realidade vivida pelos pacientes, no sentido que, quanto maior o nível do estadiamento e se há presença de necrose, maior o grau de pertinência ao conjunto de saída, indicando chances maiores para risco de metástase e consequentemente maiores probabilidades desse fato ocorrer.

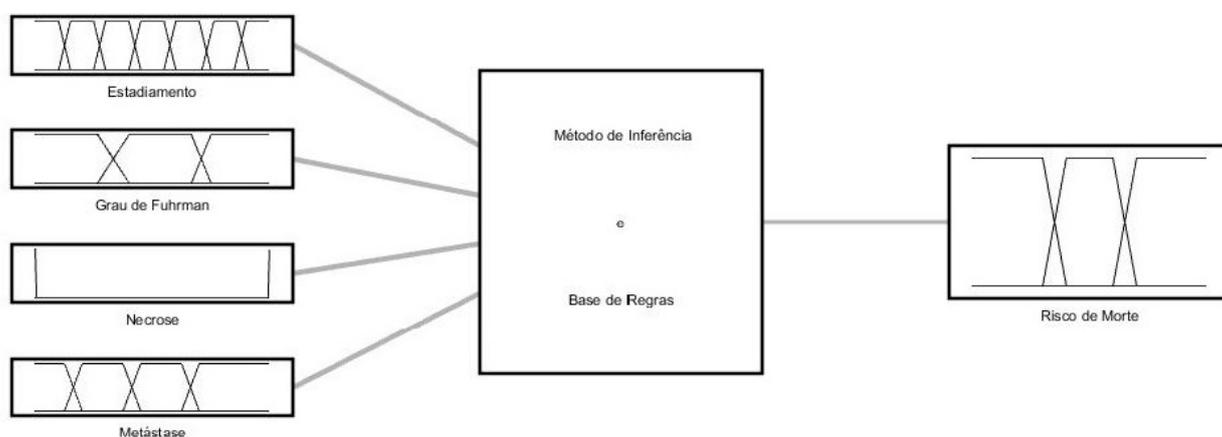
Portanto, os resultados foram analisados pelos especialistas e, mostraram-se um tanto quanto satisfatórios e coerentes com a realidade dos pacientes observados.

### 3.2.2 Segundo Modelo

No modelo, cuja saída é o *Risco de Morte*, as variáveis usadas como entrada são as mesmas que o modelo com *Risco de Metástase*, trocando apenas a entrada *Tamanho do Tumor* por *Metástase*. Desta forma, foram escolhidas como variáveis de entrada do sistema: *Estadiamento*, *Grau de Fuhrman*, *Presença de Necrose* e *Metástase* e como variável de saída: *Risco de Morte*.

A [Figura 14](#) ilustra o sistema baseado em regras *fuzzy* utilizado para este segundo modelo.

Figura 14 – Estrutura básica do sistema baseado em regras fuzzy.



Assim como no primeiro modelo, as variáveis de entrada e de saída do sistema foram consideradas como variáveis linguísticas e seus valores como conjuntos *fuzzy*, em seus respectivos domínios. As variáveis quantitativas, contínuas por natureza, são fuzzificadas de maneira direta, enquanto que as variáveis qualitativas são arranjadas numa escala entre 0 a 1.

Os termos linguísticos e as funções de pertinência atribuídos às variáveis *Estadiamento*, *Grau de Fuhrman* e *Presença de Necrose* foram os mesmos usados no modelo 1. Para analisar o risco de morte achamos conveniente considerar a presença de metástase, que é considerada fator prognóstico em algumas literaturas (POMPEO, 2006) (ORNELLAS, 2012). Podemos ver a classificação dessa variável na Figura 15.

A variável de saída do sistema *Risco de Morte* foi considerada como Baixo, Médio e Alto. O domínio desta variável é o intervalo  $[0, 1]$ . A Figura 16 ilustra as funções de pertinência dos conjuntos *fuzzy*.

O próximo passo é a construção da base de regras. Para isso, foram feitas todas as combinações possíveis entre as variáveis de entrada *Estadiamento*, *Grau de Fuhrman*, *Presença de Necrose* e *Metástase* e a variável de saída *Risco de Morte* levando em conta os termos linguísticos, atribuídos à cada uma destas variáveis.

A Tabela 7 traz parte da base de regras. Ela foi construída da mesma forma que a tabela para *Risco de Metástase*. Na primeira coluna, N indica o número da regra, seguindo com as variáveis de entrada e saída e, na última coluna, Peso indica o peso atribuído à cada regra, que foi obtido da mesma forma que o modelo anterior, baseados em dados de pacientes do Hospital das Clínicas (HC) e em estudo estatístico.

Figura 15 – Funções de pertinência dos conjuntos *fuzzy*, assumidos pela variável linguística Metástase.

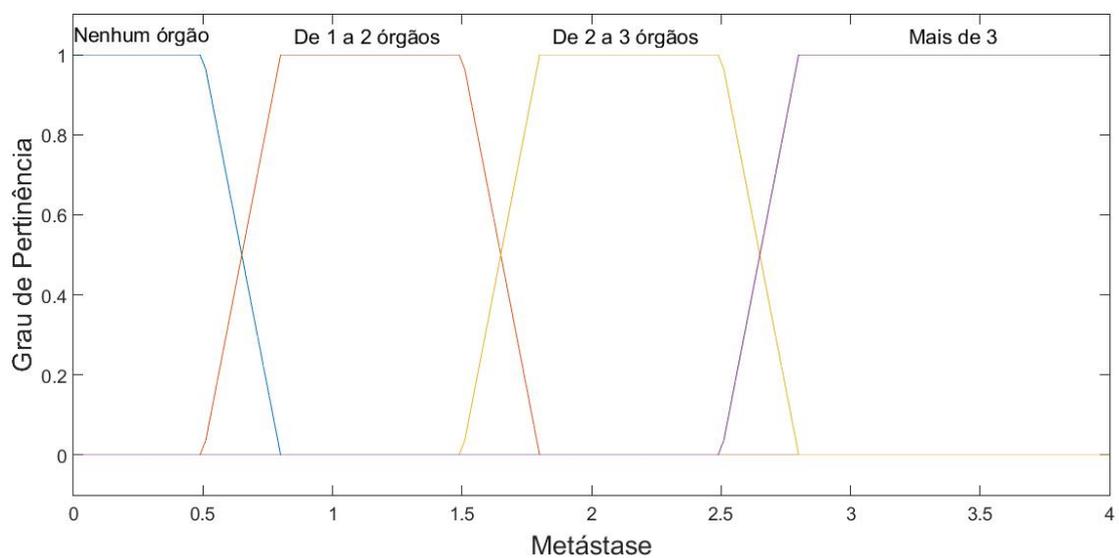
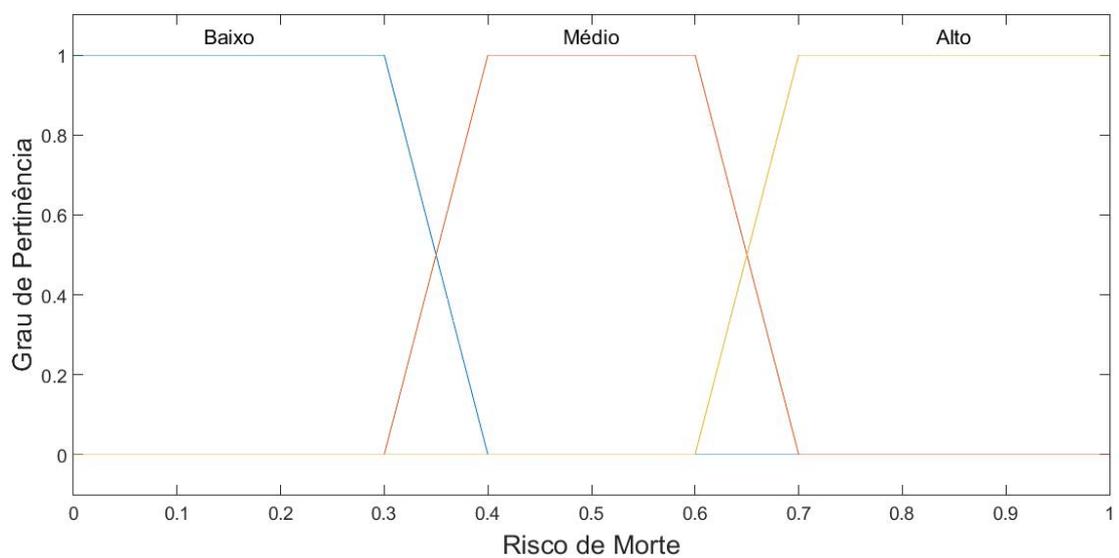


Figura 16 – Funções de pertinência dos conjuntos *fuzzy*, assumidos pela variável linguística Risco de Morte.



Como já comentamos, o número de regras é determinado pela quantidade de termos de cada variável em análise. Assim, como temos quatro variáveis de entrada - *Estadiamento*, *Grau de Fuhrman*, *Presença de Necrose* e *Metástase* - contendo respectivamente 7, 3, 2 e 4 termos linguísticos para cada variável e uma variável de saída com 3 termos linguísticos, foram criadas 504 regras.

Definida a base de regras, devemos traduzi-la matematicamente a fim de obtermos a saída do sistema. O método utilizado no módulo de inferência foi o Método de Mamdani e a defuzzificação foi feita pelo Método do Centro de Gravidade, assim como no modelo anterior.

### 3.2.2.1 Simulações

Com o intuito de verificar a confiabilidade dos resultados gerados pelo modelo, foram feitas simulações com os mesmos dados dos pacientes do Hospital das Clínicas da UNICAMP. Alguns resultados, em termos de possibilidade e probabilidade, encontram-se na [Tabela 8](#). A possibilidade foi obtida a partir do sistema *fuzzy* e a probabilidade foi obtida da transformação possibilidade-probabilidade, como utilizamos anteriormente.

Seja, por exemplo, um paciente com estadiamento T1b, grau de Fuhrman 3, com necrose e sem metástase. As possibilidades, determinadas pelo SBRF, de que o risco de morte para esse paciente seja baixo, médio e alto são, respectivamente, 0,658, 0,342 e 0. Transformando esses valores em probabilidades - via expressão (2.7) - obtemos 66%, 34% e 0%, respectivamente.

Observando a [Tabela 8](#), podemos notar que o risco de morte pela doença se agrava conforme há presença de necrose e com o aumento no estadiamento. A importância destas variáveis também foi observada durante a análise multivariada.

De modo geral, quando comparamos os resultados obtidos pelo modelo com os dados e prognósticos dos pacientes podemos afirmar que as simulações se mostraram satisfatórias e coerentes com a realidade.

Tabela 6 – Alguns resultados obtidos pelo SBRF.

PACIENTE	SAÍDA	RISCO DE METÁSTASE	POSSIB.	PROBAB.
Estad. T1a	0,1489	Baixo	1	100%
Furman baixo		Médio	0	0%
Necrose ausente		Alto	0	0%
Tam. 1,7 cm				
Estad. T1a	0,2616	Baixo	0,884	88%
Furman baixo		Médio	0,116	12%
Necrose presente		Alto	0,000	0%
Tam. 2,1 cm				
Estad. T1b	0,3110	Baixo	0,39	39%
Furman médio		Médio	0,61	61%
Necrose ausente		Alto	0,000	0%
Tam. 6,5 cm				
Estad. T2a	0,3544	Baixo	0	0%
Furman médio		Médio	1	100%
Necrose ausente		Alto	0	0%
Tam. 9 cm				
Estad. T2b	0,3245	Baixo	0,255	26%
Furman médio		Médio	0,745	74%
Necrose ausente		Alto	0	0%
Tam. 11 cm				
Estad. T3a	0,5756	Baixo	0	0%
Furman médio		Médio	0,744	74%
Necrose ausente		Alto	0,256	26%
Tam. 6,5 cm				
Estad. T3a	0,6336	Baixo	0	0%
Furman médio		Médio	0,164	16%
Necrose presente		Alto	0,836	84%
Tam. 11,5 cm				
Estad. T3b"	0,5975	Baixo	0	0%
Furman baixo		Médio	0,525	53%
Necrose ausente		Alto	0,475	47%
Tam. 5 cm				
Estad. T3b"	0,6880	Baixo	0	0%
Furman médio		Médio	0	0%
Necrose ausente		Alto	1	100%
Tam. 10 cm				
Estad. T4	0,6767	Baixo	0	0%
Furman baixo		Médio	0	0%
Necrose presente		Alto	1	100%
Tam. 9 cm				

Tabela 7 – Algumas das 504 regras, da Base de Regras construída. Na 1ª coluna consta o n<sup>o</sup> da regra e, na última, o peso utilizado para ponderá-la.

N	ESTAD.	FUHRMAN	NEC.	METÁSTASE	RISCO MORTE	PESO
4	T1a	Baixo	Ausência	De 1 a 2 órgãos	Baixo	0.8
109	T1b	Médio	Presença	Nenhum órgão	Baixo	0.6
110	T1b	Médio	Presença	Nenhum órgão	Médio	0.3
147	T2a	Baixo	Ausência	Nenhum órgão	Alto	0.1
272	T2b	Alto	Ausência	De 2 a 3 órgãos	Médio	0.4
298	T3a	Baixo	Ausência	Mais de 3 órgãos	Baixo	0.2
333	T3a	Médio	Presença	De 2 a 3 órgãos	Alto	0.7
424	T3b"	Alto	Presença	De 1 a 2 órgãos	Médio	0.1
425	T3b"	Alto	Presença	De 1 a 2 órgãos	Alto	0.9
467	T4	Médio	Ausência	Mais de 3 órgãos	Alto	0.8

Tabela 8 – Alguns resultados obtidos pelo SBRF.

PACIENTE	SAÍDA	RISCO DE MORTE	POSSIB.	PROBAB.
Estad. T1a	0,2072	Baixo	1	100%
Furman baixo		Médio	0	0%
Necrose ausente		Alto	0	0%
Meta. nenhum órgão				
Estad. T1b	0,2684	Baixo	1	100%
Furman baixo		Médio	0	0%
Necrose ausente		Alto	0	0%
Meta. nenhum órgão				
Estad. T1b	0,3342	Baixo	0,658	66%
Furman médio		Médio	0,342	34%
Necrose presente		Alto	0,000	0%
Meta. nenhum órgão				
Estad. T2a	0,5671	Baixo	0	0%
Furman alto		Médio	1	100%
Necrose presente		Alto	0	0%
Meta. de 2 a 3 órgãos				
Estad. T2b	0,2684	Baixo	1	100%
Furman médio		Médio	0	0%
Necrose ausente		Alto	0	0%
Meta. nenhum órgão				
Estad. T3a	0,6658	Baixo	0,000	0%
Furman médio		Médio	0,342	34%
Necrose ausente		Alto	0,658	66%
Meta. de 2 a 3 órgãos				
Estad. T3a	0,7316	Baixo	0	0%
Furman médio		Médio	0	0%
Necrose presente		Alto	1	100%
Meta. nenhum órgão				
Estad. T3b"	0,6985	Baixo	0	0%
Furman baixo		Médio	0,015	1%
Necrose ausente		Alto	0,985	99%
Meta. nenhum órgão				
Estad. T3b"	0,6526	Baixo	0	0%
Furman médio		Médio	0,474	47%
Necrose ausente		Alto	0,526	53%
Meta. mais de 3 órgãos				
Estad. T3b"	0,7928	Baixo	0	0%
Furman médio		Médio	0	0%
Necrose presente		Alto	1	100%
Meta. de 2 a 3 órgãos				

## 4 Conclusão e Trabalhos Futuros

### 4.1 Conclusão

O câncer é um importante problema de saúde pública no mundo. No Brasil, constitui a segunda causa de morte por doença, atrás somente das doenças do aparelho circulatório. Com relação ao tipo de câncer, o tumor de rim é o terceiro tumor mais comum do sistema urinário, precedido pelos câncer de próstata e bexiga.

O carcinoma de células renais (CCR) é uma doença heterogênea com prognóstico amplamente variável. Um conhecimento preciso do risco individual de progressão da doença e mortalidade após o tratamento é essencial para aconselhar pacientes, a tomada de decisões de tratamento e a seleção de regimes de seguimento adequados. Esse conhecimento depende de previsões com base na avaliação de fatores prognósticos clínicos e patológicos (VOLPE; PATARD, 2010).

Ao longo desta dissertação, foi apresentada uma opção para prever os riscos de metástase e morte para pacientes com tumores de rim. Cada modelo proposto consiste num sistema baseado em regras *fuzzy* que combina dados do tumor dos pacientes (estadiamento, grau de Fuhrman e presença de necrose e tamanho do tumor para o Modelo 1 e estadiamento, grau de Fuhrman, presença de necrose e presença de metástase para o segundo modelo) levando em conta um conjunto de regras, de natureza linguística, elaborado a partir da ajuda dos especialistas que contactamos, de pesquisa em literaturas na área, tais como Segre (2010) e Pompeo (2006), e em estudo estatístico com dados reais de pacientes do Hospital das Clínicas (HC) - UNICAMP. A construção dos dois modelos é baseada na teoria dos conjuntos *fuzzy*, a qual permite incluir a incerteza das informações disponíveis e realizar processos de raciocínio aproximado utilizando essas informações.

A saída dos dois sistemas é um número real entre 0 e 1, e a este número corresponde um grau de pertinência em cada conjunto para Risco de Metástase: Baixo, Médio e Alto, no primeiro modelo, e Risco de Morte: Baixo, Médio e Alto, para o segundo modelo.

Fizemos simulações utilizando os dados de pacientes do Hospital das Clínicas da UNICAMP que apresentavam informações completas de todas as variáveis envolvidas na modelagem. Os resultados das simulações bem como as suas probabilidades também foram apresentados.

Além disso, apresentamos um estudo estatístico para analisar as relações entre os fatores prognósticos e a sobrevida dos pacientes. Esse estudo foi feito pelo estatístico Sidney Ragazzi. O estudo também envolveu a pesquisa com duas classificações de Fuhrman: grau único e grau combinado (muitas vezes, ocorre encontrar dois graus de Fuhrman na extensão tumoral), essa última classificação é sugerida pelo médico patologista Prof. Dr. Athanase Billis do Hospital das Clínicas - UNICAMP. Os dados de pacientes citados anteriormente foram úteis para realizar essa análise.

Os resultados dessa análise contribuíram para a construção dos modelos *fuzzy*. Pois, ela revelou que há associação entre Fuhrman e metástase, porém esta associação é baixa a no máximo, moderada. Também mostrou que quando estamos considerando apenas metástase, as duas classificações para o Fuhrman (único e combinado) indicam a mesma chance de desenvolvimento de metástase. Já considerando o modelo com morte, o Fuhrman único evidencia mais os casos de metástase e morte, sugerindo que o Fuhrman único é mais compatível que o Fuhrman combinado. A segunda conclusão do estudo é o fato que entre as variáveis analisadas (estadiamento, grau de Fuhrman, necrose e tamanho do tumor), as que se mostraram mais significativas e de maior influência na sobrevida dos pacientes foram: presença de necrose e estadiamento. Este fato pode ser observado na Seção 3.1. Desta forma, foi possível ponderar as regras dos modelos, atribuindo pesos maiores às regras com presença de necrose e estadiamento maiores.

Com isso, pudemos construir os modelos e realizar as simulações com os dados reais coletados. Ao compararmos os resultados dos modelos com o prognóstico dos casos reais, as respostas dos modelos se mostraram próximas da realidade clínica enfrentada pelos pacientes com tumor de rim. Os especialistas analisaram os resultados e também consideraram satisfatórios.

## 4.2 Trabalhos Futuros

Uma proposta de trabalho futuro é a inclusão de novas variáveis de entrada como invasão vascular, idade, sexo e cor. Também poderíamos incluir como variáveis de entrada, se o paciente é ou não fumante e se é diabético ou não, pois esses fatores foram notados com frequência quando tivemos acesso aos prontuários dos pacientes.

Vale lembrar que quanto maior a quantidade de variáveis mais precisos serão os resultados. No entanto, um número maior de variáveis envolvidas acarreta em uma complexidade maior para análise do problema principalmente na atribuição de peso para as regras, tendo em vista a dificuldade dos especialistas em ponderá-las.

Para melhorar os resultados dos modelos também poderíamos trabalhar com

um maior banco de dados, que são úteis para a verificação da confiabilidade dos mesmos e obter resultados mais precisos.

Uma outra proposta interessante seria o desenvolvimento de um software para cada modelo, facilitando assim, o acesso por parte dos especialistas.

Além disso, podem ser feitos outros estudos sobre câncer renal envolvendo o tratamento da doença. Pode-se também tentar construir um novo sistema *fuzzy* que forneça informações relacionadas à sobrevida dos pacientes, dependendo da terapia escolhida para o seu tratamento e assim, analisar qual terapia é mais adequada a cada caso ou qual delas é mais eficiente para esse tipo de neoplasia. Trabalhar com imunoterapia é um estudo alternativo também.

## Referências

- ADLASSNIG, K. P. A fuzzy logical model of computer-assisted medical diagnosis. *Methods of Information in Medicine*, v. 19, n. 3, p. 141–148, 1980.
- AMENDOLA, M.; SOUZA, A. L. de; BARROS, L. C. *Manual do uso da teoria dos conjuntos fuzzy no MATLAB 6.5*. Campinas - SP, 2005.
- AMERICAN CANCER SOCIETY. *What is kidney cancer? Kidney Cancer (Adult) - Renal Cell Carcinoma*. New York, NY, USA, 2016. Disponível em: <<https://www.cancer.org/cancer/kidney-cancer/about/what-is-kidney-cancer.html>>. Acesso em: 20 out. 2016.
- BARROS, L. C.; BASSANEZI, R. C. *Tópicos de Lógica Fuzzy e Biomatemática*. Campinas - SP: UNICAMP/IMECC, 2010.
- BECKER, A. et al. Critical analyses of a simplified fuhrman grading scheme for prediction of cancer specific mortality in patients with clear cell renal cell carcinoma - impact on prognosis. *EJSO European Journal of Surgical Oncology*, v. 42, n. 3, p. 419–425, 2016.
- BILLIS, A. et. al. *Department of Pathology, School of Medicine, State University of Campinas (UNICAMP)*, Campinas - SP.
- BILOBROVEC, M. *Sistema Especialista em Lógica Fuzzy para o Controle, Gerenciamento e Manutenção da qualidade em Processo de Aeração de Grãos*. Dissertação (Dissertação de Mestrado) — Universidade Tecnológica Federal do Paraná, Ponta Grossa - PR, 2005.
- CASTANHO, M. J. de P. *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*. Dissertação (Mestrado) — FEEC - Universidade Estadual de Campinas, Campinas - SP, mar. 2005.
- CHEVILLE, J. C. et al. Comparisons of outcome and prognostic features among histologic subtypes of renal cell carcinoma. *The American Journal of Surgical Pathology*, v. 27, n. 5, p. 612–624, 2003.
- DUBOIS, D.; PRADE, H. *Fuzzy Sets and Systems: Theory and Applications*. New York, NY, USA: Academic Press, INC, 1980.
- FUHRMAN, S. A.; LASKY, L. C.; LIMAS, C. Prognostic significance of morphologic parameters in renal cell carcinoma. v. 6, n. 7, p. 655–663, 1982.
- GOMEZ, J.; DASGUPTA, D. Evolving fuzzy classifiers for intrusion detection. *Proceedings of the 2002 IEEE, Workshop on Information Assurance*, New York, NY, USA, 2001.
- GOMIDE, F.; PEDRYCZ, W. *Fuzzy Systems Engineering: Toward Human - Centric Computing*. USA: Wiley-IEEE Press, 2007.
- HOSPITAL DO CÂNCER DE BARRETOS. *O que é câncer de rim*. Barretos - SP, 2016. Disponível em: <<https://www.hcancerbarretos.com.br/cancer-de-rim>>. Acesso em: 19 mar. 2016.

INSTITUTO NACIONAL DE CÂNCER - INCA. *O que é câncer*. Rio de Janeiro - RJ, 2016. Disponível em: <<http://www2.inca.gov.br/wps/wcm/connect/cancer/site/oquee>>. Acesso em: 15 out. 2016.

\_\_\_\_\_. *ESTIMATIVA 2016 - Incidência de Câncer no Brasil*. Rio de Janeiro - RJ, 2017. Disponível em: <<http://www.inca.gov.br/estimativa/2016/index.asp?ID=2>>. Acesso em: 03 jul. 2017.

JAFELICE, R. S. da M. *Modelagem Fuzzy para a dinâmica de transferência de soropositivos para HIV em doença plenamente manifesta*. Tese (Doutorado) — FEEC - Universidade Estadual de Campinas, Campinas - SP, 2003.

KLIR, G. J.; CLAIR, U. H. S.; YUAN, B. *Fuzzy Set Theory: Foundations and Applications*. USA: Prentice Hall PTR, 1997.

LJUNGBERG, B. et al. *Orientações sobre carcinoma das células renais*. Associação Europeia de Urologia, 2009. 48-60 p. Disponível em: <<http://www.uroweb.org>>.

MAGALHÃES, L. *Anatomia e Histologia do Néfron*. 2017. Disponível em: <<https://www.todamateria.com.br/nefron/>>. Acesso em: 04 jul 2017.

MEDEIROS, J. L.; GELB, A. B.; WEISS, L. M. Renal cell carcinoma - prognostic significance of morphologic parameters in 121 cases. *PubMed*, v. 61, n. 8, p. 1639–1651, 1988.

MEMORIAL SLOAN KETTERING CANCER CENTER. *Kidney Cancer Types*. [S.l.], 2017. Disponível em: <<https://www.mskcc.org/cancer-care/types/kidney/types>>. Acesso em: 02 mai. 2017.

NARDI, A. C. et al. Epidemiologic characteristics of renal cell carcinoma in brazil. *International Brazilian Journal Of Urology*, v. 36, n. 2, p. 151–158, 2010.

ORNELLAS, A. A. et al. Prognostic factors in renal cell carcinoma: analysis of 227 patients treated at the brazilian national cancer institute. *International Brazilian Journal Of Urology*, v. 38, n. 2, p. 185–194, 2012.

ORTEGA, N. R. S. *Aplicação da teoria de conjuntos fuzzy a problemas de biomedicina*. Tese (Doutorado) — Instituto de Física da Universidade de São Paulo, São Paulo - SP, 2001.

PAULSEN, F. e. J. W. *Sobotta - Atlas da anatomia humana - Órgãos internos*. 23. ed. [S.l.: s.n.], 2016. 166-167,171 p.

POMPEO, A. C. L. et al. Câncer renal: Prognóstico. *Projeto Diretrizes, Sociedade Brasileira de Urologia, Associação Médica Brasileira e Conselho Federal de Medicina*, p. 1–13, 2006.

SAVERGNINI, K. D. *Modelagem Fuzzy para prever os riscos de recidiva e progressão de tumores superficiais de bexiga*. Dissertação (Dissertação de Mestrado) — Universidade Estadual de Campinas - UNICAMP, Campinas - SP, 2009.

SEGRE, C. C. *Nomograma para definição prognóstica em câncer renal. Estudo com pacientes brasileiros*. Tese (Doutorado) — Faculdade de Medicina - Universidade de São Paulo, São Paulo - SP, 2010.

SILVEIRA, G. P. *Aplicação da Teoria de Conjuntos Fuzzy na predição do estadiamento patológico do Câncer de Próstata*. Dissertação (Dissertação de Mestrado) — Universidade Estadual de Campinas - UNICAMP, Campinas - SP, 2007.

SOUZA, O. E. R. de. *Neoplasias renais em nefrectomias realizadas no período de 1990 a 2006 na Universidade Federal do Triângulo Mineiro*. Dissertação (Mestrado) — Universidade Federal do Triângulo Mineiro, Uberaba - MG, ago. 2008.

FERREIRA, U.; SASSE, A. D. (Ed.). *Diretrizes baseadas em evidências em tumores urológicos*. 7. ed. São Paulo - SP, 2015. 42-47 p.

VOLPE, A.; PATARD, J. J. Prognostic factors in renal cell carcinoma. *World Journal of Urology*, v. 28, n. 3, p. 319–327, 2010.

WORLD HEALTH ORGANIZATION - WHO. *What is cancer?* Geneva, 2016. Disponível em: <<http://www.who.int/cancer/en/>>. Acesso em: dez. 2016.