



**UNIVERSIDADE
FEDERAL RURAL
DE PERNAMBUCO**



Manoel Antônio de Freitas Neto

Um Modelo Fuzzy para Estimar o Risco de Desenvolver Câncer de Colo de Útero

Recife

2021

Manoel Antônio de Freitas Neto

Um Modelo Fuzzy para Estimar o Risco de Desenvolver Câncer de Colo de Útero

Monografia apresentada ao Curso de Bacharelado em Sistemas de Informação da Universidade Federal Rural de Pernambuco, como requisito parcial para obtenção do título de Bacharel em Sistemas de Informação.

Universidade Federal Rural de Pernambuco – UFRPE

Departamento de Estatística e Informática

Curso de Bacharelado em Sistemas de Informação

Orientador: Prof^{fa}. Dr^a. Silvana Bocanegra

Recife

2021

Dados Internacionais de Catalogação na Publicação
Universidade Federal Rural de Pernambuco
Sistema Integrado de Bibliotecas
Gerada automaticamente, mediante os dados fornecidos pelo(a) autor(a)

F866m de Freitas Neto, Manoel Antônio
Um Modelo Fuzzy para Estimar o Risco de Desenvolver Câncer de Colo de Útero / Manoel Antônio de
Freitas Neto. - 2021.
65 f. : il.

Orientadora: Silvana Bocanegra.
Inclui referências.

Trabalho de Conclusão de Curso (Graduação) - Universidade Federal Rural de Pernambuco,
Bacharelado em Sistemas da Informação, Recife, 2021.

1. Câncer de Colo de Útero. 2. Fatores de Risco. 3. Lógica Fuzzy. 4. Método de inferência de Mamdani.
5. HPV. I. Bocanegra, Silvana, orient. II. Título

CDD 004

*Dedico este trabalho à todas as pessoas que sempre estiveram ao meu lado,
torcendo pelo meu sucesso, em especial minha família, amigos e professores.
Dedico também à todos os cientistas que lutam para transformar o mundo em um
lugar melhor para todos. ...*

Agradecimentos

Agradeço principalmente à minha família que sempre me incentivou a buscar o estudo e sempre me deu condições de correr atrás dos meus objetivos. Sem o apoio dessas pessoas eu não conseguiria concluir esse trabalho com finalidade de aprofundá-lo no mestrado. Deixo aqui minhas lembranças aos meus pais Tânia e Amaury, que sempre deram suporte nos meus objetivos, meus tios Kátia e Silvio, que me ajudaram desde a escolha do curso até a formação, meus primos Pedro, Kênio, Lucas, Vitor, meus avós Baia e Veião e minhas tias. Minha madrinha Helade cuja profissão me inspirou a aprofundar meus conhecimentos na área de saúde, sempre me dando apoio e ajuda. Por fim à minha tia Azinete, que sempre esteve presente na minha vida e fez parte de toda construção que tive.

Aos meus professores da graduação, principalmente minha orientadora Silvana Bocanegra, que soube me guiar, dar suporte e tranquilizar em todos os passos dessa monografia. Ao professor Jones Albuquerque que me apresentou a esse universo epidemiológico, não determinístico e probabilístico. À Iolanda Matias por tantas horas de aprendizado e conversas no LIKA.

Agradeço também aos meus amigos da graduação que fizeram parte de cada pedaço da minha formação, que passei noites em claro fazendo projetos impossíveis, que brigamos e fizemos as pazes, que nos sacrificamos e tivemos momentos de descontração, vocês Ariana, Cristian e Angela, meus amigos porteiros Evele, Brunna, Filipe, Marta e Vinícius, a amizade de vocês vai além do mundo acadêmico.

Aproveito também para agradecer especialmente a minha amiga Giovanna Bloise, que tanto me inspirou e auxiliou na pesquisa, a amizade mais inesperada de todas.

Aos meus amigos do ensino médio Clarinha e Mateus, que tornaram esses anos de graduação leves e descontraídos.

Aos meus amigos Pedro e Paulo, que de uma tarde ao caso fizeram surgir uma amizade que ultrapassou a universidade. Muito obrigado por vocês existirem.

Aos meus amigos Olindenses Carol, Marcela, Ricardo ...Que estiveram para me ajudar, conversar e aconselhar, obrigado por dividir momentos bons e ruins.

E todas as pessoas que não foram citadas mas sabem que fizeram parte em algum momento ...

À todos vocês muito obrigado!

“Se vivêssemos num planeta onde nada nunca mudasse, pouco haveria para fazer. Não haveria nada a descobrir. Não haveria incentivo para a ciência. E se vivêssemos num mundo imprevisível, onde as coisas mudassem de modo aleatório ou de maneiras muito complexas, não seríamos capazes de ter uma ideia sobre as coisas. Mais uma vez, não haveria essa coisa chamada ciência. Mas vivemos num universo intermediário, onde as coisas mudam, porém de acordo com certos padrões, regras ou, como nós as chamamos, leis da natureza. Se atirmos um pedaço de pau para o ar, ele sempre cairá. Se o Sol se põe no Ocidente, ele sempre nasce de novo na manhã seguinte no Oriente. E assim fica possível ter uma ideia das coisas. Podemos fazer ciência, e com isso melhorar nossas vidas. ”

(Carl Sagan)

Resumo

O câncer de colo de útero é o quarto tipo de câncer que mais mata mulheres no mundo e no Brasil. Por ano, são quase 600.000 novas ocorrências onde mais da metade das pessoas morrem. Apesar dessa alta taxa de incidência, a chance de recuperação para casos identificados precocemente é igualmente alta. Diante disto, a proposta desse trabalho consiste no desenvolvimento de um modelo matemático computacional que possa ser útil para estimar o risco de uma pessoa desenvolver a doença e assim auxiliar os especialistas numa possível previsão e prevenção. Para construção do modelo foi utilizada a teoria de conjuntos de lógica fuzzy. Inicialmente foram estudados os principais fatores de risco associados ao surgimento do câncer e criadas variáveis linguísticas para subsidiar a construção do sistema de regras fuzzy. Ao selecionar as variáveis idade, idade da primeira relação sexual, histórico de câncer dos pais, presença de ISTs, tabagismo, quantidade de parceiros homens no último ano e vacinação contra HPV é possível estimar o risco do paciente desenvolver câncer de colo uterino. Ele pode ser classificado em cinco categorias: muito baixo, baixo, moderado, alto e muito alto. Além disso, um aplicativo foi desenvolvido para captação de dados e futura integração com o modelo para estimativa de risco.

Palavras-chave: Câncer de Colo de Útero, Fatores de Risco, HPV, Lógica Fuzzy, Método de inferência de Mamdani

Abstract

Cervical cancer is the fourth type of cancer that kills more women in the world and in Brazil. There are almost 600.000 cases per year in which more than half die. Despite this high incidence rate, the probability of recuperation in cases with early detection is equally high. In view of this, the proposal of this work is the development of a mathematical computational model that can be useful to estimate the risk of a person developing the disease and thus help the specialists in a possible prediction and prevention of this disease. To build the model the theory of fuzzy logical sets was used. Initially was studied the main risk factors associated to the development of the cancer and created linguistics variables to subsidize the building of the fuzzy rules system. By selecting the variables age, age of the first sexual relationship, parent's historical of cancer, presence of ISTs, smoking, amount of male partners in the last year and HPV vaccination is possible to estimate the risk of a patient developer cervical cancer. They can be classified into five categories: very low, low, moderate, high and very high. Besides that, a mobile application was developed for data capture and future integration with the model for risk estimation.

Keywords: Cervical Cancer, Risk Factors, HPV, Fuzzy Logic, Mamdani's Inference Method,

Lista de ilustrações

Figura 1 – Taxa de mortalidade de câncer de colo de útero nas Regiões do Brasil.	12
Figura 2 – Incidência e Mortalidade de Cancer de Colo de Utero.	16
Figura 3 – Prevalencia de HPV em meninas após programa de imunização na Inglaterra.	18
Figura 4 – Índice de incidencia de câncer de colo de útero para cada 100.000 habitantes.	22
Figura 5 – diferença entre lógica crisp(a) e lógica fuzzy(b)	32
Figura 6 – Funcionamento de uma variável linguística e termos linguísticos. . .	37
Figura 7 – Funções de pertinência trapezoidal(a e c) e triangular(b)	38
Figura 8 – Exemplo de funções de pertinência.	38
Figura 9 – Diagrama de um Modelo Fuzzy.	39
Figura 10 – Saídas de regras baseadas no método de Mamdani usando t-norma.	41
Figura 11 – Saída final fuzzy no método de Mamdani usando t-conorma.	41
Figura 12 – Esquema do banco de dados do aplicativo.	43
Figura 13 – Telas do aplicativo: As duas primeiras são telas de questionário, a última representa a tela de resultado como ficaria após o cálculo de risco.	44
Figura 14 – Diagrama de Classes do App.	45
Figura 15 – Resumo do modelo fuzzy para cálculo de risco de câncer de colo de útero.	46
Figura 16 – Representação gráfica da variável Idade.	47
Figura 17 – Representação gráfica da variável Idade da primeira relação sexual.	48
Figura 18 – Representação gráfica da variável Histórico Familiar.	49
Figura 19 – Representação gráfica da variável Presença de ISTs.	50
Figura 20 – Representação gráfica da variável Tabagismo.	51
Figura 21 – Representação gráfica da variável Quantidade de Parceiros.	52
Figura 22 – Representação gráfica da variável Vacina Contra HPV.	52
Figura 23 – Representação gráfica da variável Chance	53
Figura 24 – Resultado da variável chance para o paciente A.	55
Figura 25 – Resultados parciais das variáveis idade, idade da primeira relação sexual, histórico de câncer, presença de outras ISTs, tabagismo, quantidade de parceiros e vacina contra HPV do paciente A.	56
Figura 26 – Resultado da variável chance para o paciente B.	57
Figura 27 – Resultado da variável chance para o paciente C.	58
Figura 28 – Resultado da variável chance para o paciente D.	59
Figura 29 – Resultado da variável chance para o paciente E.	60

Lista de tabelas

Tabela 1 – Exemplo de grau de pertinência para o conjunto Universo.	33
Tabela 2 – Resumo das variáveis linguísticas e termos linguísticos.	53
Tabela 3 – Resumo dos resultados dos casos de teste.	60

Lista de abreviaturas e siglas

CC	Câncer Cervical
NIC	Neoplasias Intraepiteliais Cervicais
CDC	Centers for Disease Control and Prevention
OMS	Organização Mundial da Saúde
HPV	Human PapilomaVirus
ISTs	Infecções Sexualmente Transmissíveis
...	...

Sumário

	Lista de ilustrações	7
1	INTRODUÇÃO	12
1.1	O problema	12
1.2	Motivação	13
1.3	Objetivos	14
1.4	Contribuições Esperadas	14
1.5	Organização do Documento	14
2	O CÂNCER CERVICAL E FATORES DE RISCO	15
2.1	O Câncer e o Vírus	15
2.1.0.1	Diagnóstico Precoce	16
2.1.0.2	Triagem	17
2.1.0.3	Abordagem de Prevenção ao Câncer de Colo em Larga Escala	17
2.1.1	Vírus do Papiloma Humano - HPV	18
2.2	Fatores Relevantes para o Risco	20
2.2.1	Idade	20
2.2.2	Idade da Primeira relação sexual	20
2.2.3	Gestações, Partos e Abortos	21
2.2.4	Alimentação	21
2.2.5	Localização	21
2.2.6	Presença de ISTs	22
2.2.7	Tabagismo	22
2.2.8	Histórico Familiar	23
2.2.9	Diabetes, Hipertensão e Uso de Corticoides	23
2.2.10	Número de Parceiros	24
2.2.11	Uso de Preservativo	24
2.2.12	Citologia Oncótica	24
2.2.13	Colposcopia	25
2.2.14	Teste de HPV	25
2.2.15	Vacina Contra HPV	25
3	MODELO USADOS PARA ESTIMAR RISCO EM DOENÇAS	26
3.1	Trabalhos Relacionados	26
3.2	Análise de Risco para Câncer de Colo de Útero	29

4	TEORIA DE CONJUNTOS FUZZY	31
4.1	História da Teoria Fuzzy	31
4.2	Conceitos Básicos	31
4.3	Conjuntos Fuzzy	33
4.4	Operações Entre Conjuntos Fuzzy	34
4.5	Normas e Conormas Triangulares	35
4.6	Sistema Baseado em Regras Fuzzy	35
4.6.1	Regras de Inferência	35
4.6.2	Variáveis Linguísticas	36
4.6.3	Função de Pertinência	37
4.6.4	Sistemas Baseados em Regras Fuzzy (SBRF)	39
4.7	Método de Inferência de Mamdani	40
5	UM MODELO BASEADO EM LÓGICA FUZZY PARA ESTIMAR O RISCO DE DESENVOLVER CÂNCER DE COLO DE ÚTERO	42
5.1	O Aplicativo	42
5.1.1	Arquitetura do App	44
5.2	Modelo Fuzzy	45
5.2.1	O Modelo	46
5.2.2	Variáveis Antecedentes e Consequente	46
5.3	Regras de Inferência	54
5.4	Resultados Preliminares	54
6	CONSIDERAÇÕES FINAIS	61
	REFERÊNCIAS	62

1 Introdução

1.1 O problema

O Câncer de colo de útero é o quarto tipo de câncer mais comum entre as mulheres no mundo, ficando atrás apenas do câncer de mama, câncer de tireóide e câncer colorretal, excluindo os tumores de pele não melanoma. Essa enfermidade é responsável por mais de 570.000 novos casos em todo o planeta, levando aproximadamente 331.000 mulheres a óbito, segundo o Instituto do câncer (INCA - Brasil).

No Brasil, a estatística para 2020 prevê que 12,6 a cada 100.000 pessoas irão desenvolver câncer de colo uterino, isso corresponde a 16.590 casos novos. O histórico da neoplasia coloca-a em quarto lugar na responsabilidade de mortes de pacientes oncológicos no país. A Figura 1 mostra o gráfico da taxa de mortalidade de Câncer de colo por região do Brasil, de 1980 a 2019.

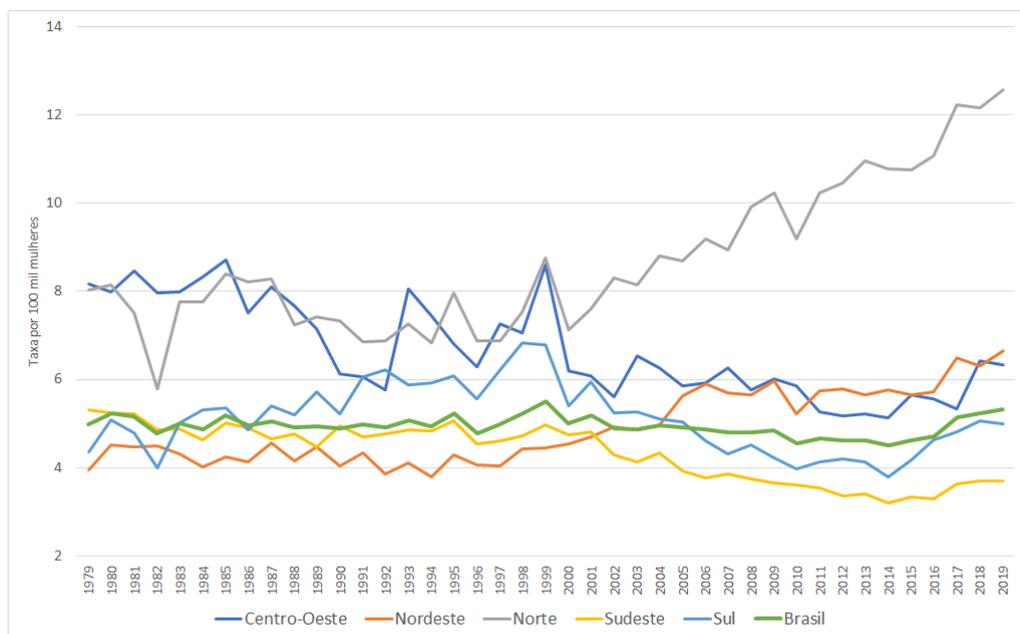


Figura 1 – Taxa de mortalidade de câncer de colo de útero nas Regiões do Brasil. (Instituto Nacional de Câncer, 2021)

Segundo o INCA, a idade média das pessoas ao receber o diagnóstico é de 50 anos, porque muitas delas deixam de se prevenir após chegar a uma idade mais avançada, 20% dos casos são diagnosticados com pacientes acima de 65 anos. Por mais que os dados sejam alarmantes, essa neoplasia tem um alto índice de cura e prevenção quando identificada precocemente, chegando perto de 100%, e 80% nos casos

considerados ambulatoriais, ou seja, que são tratados em consultório e não precisam de intervenção cirúrgica. O HPV é a principal causa do surgimento de câncer de colo, mas além disso, há diversos outros fatores que podem ou não estar relacionados à presença do vírus, por exemplo, tabagismo, idade de início da vida sexual, qualidade da alimentação, prevenção etc. Eles são considerados pelos cientistas da área como fatores de risco e prevenção que exercem influência no desenvolvimento do câncer.

Várias estratégias são usadas para calcular risco de desenvolvimento de doenças, por exemplo: comparações estatísticas (ROTHBERG et al., 2018), autômatos celulares (CABRERA-BECERRIL et al., 2017) e modelos compartimentais com equações diferenciais (SALDAÑA; KOROBENIKOV; BARRADAS, 2019), explicados no Capítulo 3. Na década de 1960, um matemático chamado Lotfali Askar-Zadeh teorizou uma lógica matemática capaz de representar conjuntos de uma forma diferente do habitual, ela foi batizada de lógica fuzzy (LAGHETTO, 2016). Fuzzy em inglês significa nebuloso. A lógica de conjuntos fuzzy quebra com o paradigma de que um elemento pertencente a um conjunto universo apenas se relaciona com outros conjuntos de forma completamente verdadeira ou completamente falsa. A proposta de Zadeh foi introduzir um conceito que classificasse elementos através de graus de pertinência ao invés das opções binárias já conhecidas. Após a publicação do seu artigo, a teoria de Zadeh incentivou diversos outros estudiosos de áreas adjacentes como engenharia, física e arquitetura a estudar e desenvolver trabalhos utilizando a lógica fuzzy. Os modelos foram tão difundidos e aceitos que não demorou até chegar no campo da medicina, por volta do final da década de 1960. Hoje em dia, a teoria de conjuntos fuzzy também é implementada para criar sistemas baseados em regras fuzzy capazes de calcular o risco de uma pessoa desenvolver doenças, desde viroses até cânceres. (SENEM; MARIN, 2002)

1.2 Motivação

Diante disso, fica exposto que o câncer de colo de útero é considerado uma patologia com alto poder de mortalidade e relevância entre os cânceres devido a sua taxa de incidência, mas também com grande chance de recuperação quando diagnosticado e tratado previamente. O número de vidas que pode ser poupado, além da quantidade de recursos que são utilizados em tratamentos custosos e sofridos são os principais motivadores desse trabalho. Melhorar a qualidade de vida e chance de recuperação dos pacientes, evitando assim, um processo doloroso de tratamento de um câncer tão invasivo. Dessa forma, verifica-se a necessidade do desenvolvimento de métodos que possam ser usados para calcular o risco de surgimento da doença e auxiliar os profissionais nas estratégias de prevenção. A teoria de lógica fuzzy vem sendo utilizada na medicina desde a década de 1970 (LAGHETTO, 2016) dado a possibilidade de

trabalhar com incertezas e respostas nebulosas, sendo observada no estudo de cálculo de risco de outras doenças como câncer de mama (YILMAZ; AYAN, 2013), câncer de pulmão (YILMAZ; AYAN, 2013), câncer de próstata (SENEM; MARIN, 2002), covid-19 (SOUSA; XIMENES; BRITO, 2020) etc. A teoria é bastante aproveitada na área de saúde porque consegue se adaptar à respostas vagas e trabalhar com níveis de pertencimento ao invés de exatidão.

1.3 Objetivos

O objetivo desse trabalho é construir um modelo matemático-computacional para estimar o risco de um indivíduo desenvolver câncer de colo de útero. Esse modelo toma como base um subconjunto das variáveis mais envolvidas no processo de desenvolvimento da neoplasia para construir um modelo fuzzy que calcula o risco de desenvolvimento da doença utilizando regras de inferência. Essas informações pessoais podem ser captadas através de uma aplicação para smartphone desenvolvida no decorrer do estudo que futuramente pode ser integrada ao modelo de estimativa.

1.4 Contribuições Esperadas

Espera-se que esse modelo possa ser aprimorado e validado com dados reais obtidos em clínicas e consultas ginecológicas, através do auxílio de profissionais da área. A partir disso, é possível aplicar esse sistema de forma geográfica, identificando regiões onde há indivíduos com alta chance de desenvolver o câncer, podendo atuar na prevenção, alertando pessoas suscetíveis e economizando recursos financeiros do tratamento tardio.

1.5 Organização do Documento

Esse trabalho é composto por mais cinco capítulos. No Capítulo 2 é abordado o câncer de colo de útero, causas, tratamentos, programas de imunização, fatores de risco e prevenção, além de uma breve explanação sobre o HPV. No Capítulo 3 são apresentados os trabalhos que foram usados como ideia para o estudo, pesquisas relacionadas, métodos alternativos para calcular risco de outras doenças e modelos fuzzy. No Capítulo 4 é explicada a teoria de conjuntos fuzzy e quais são os artifícios necessários para construir um sistema baseado em regras fuzzy. No Capítulo 5 é explicada a construção da aplicação móvel de coleta de dados e o desenvolvimento do modelo fuzzy que estima o risco de desenvolver câncer de colo de útero. No Capítulo 6 encontram-se as conclusões iniciais desse trabalho, suas limitações e possíveis desenvolvimentos de trabalhos futuros. Finalizando, tem-se as referências bibliográficas.

2 O Câncer Cervical e Fatores de Risco

Neste capítulo serão apresentados dados sobre a doença, estratégias de controle e prevenção e os principais fatores associados ao desenvolvimento do câncer. Esse estudo auxiliará no entendimento da enfermidade e servirá de base para construção do modelo.

2.1 O Câncer e o Vírus

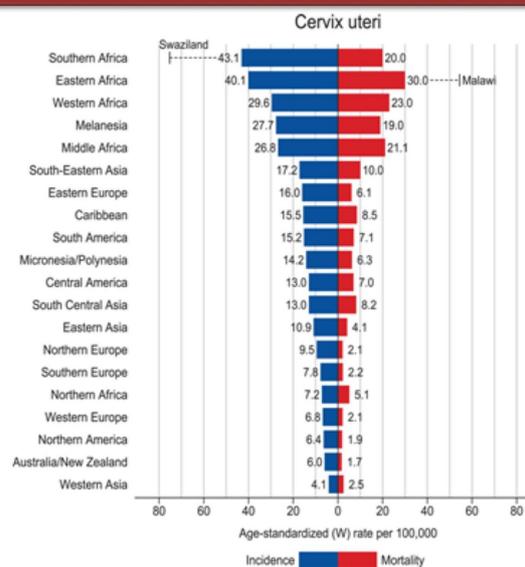
O câncer de colo de útero se mantém como o quarto tipo de câncer mais frequente na população feminina, representando 6,6% de todos os tipos de cânceres que afetam as mulheres, segundo a organização mundial de saúde (OMS). Aproximadamente 90% das mortes causadas por câncer de colo acontecem em países de baixa e média renda, atualmente, o Brasil é classificado como país de alto índice de desenvolvimento, segundo o a ONU ([United Nations Development Programme, 2020](#)). Especificamente no Brasil, ao longo dos anos, com a implementação dos planos de controle de câncer de colo de útero, que se deu a partir de 1940, no governo de Getúlio Vargas, com a introdução da citologia e da colposcopia, a taxa de mortalidade teve uma queda significativa. Mas, infelizmente, esse número ainda se mantém alto nas primeiras décadas do século XXI, no Brasil, tendo um aumento na região Norte ([Instituto Nacional de Câncer, 2021](#)). Em 2014, foram implementados serviços para diagnóstico e tratamento de lesões precursoras do câncer, com o objetivo de pré diagnosticar e evitar o desenvolvimento. Também houve a execução da campanha de vacinação contra o vírus do papiloma humano (HPV), através do programa nacional de imunização (PNI).

A alta taxa de mortalidade por câncer cervical, ao redor do mundo, pode ser reduzida com abordagens abrangentes que incluem prevenção, diagnóstico prévio, programas efetivos de triagem e tratamento das lesões precursoras. A medida mais efetiva para prevenir a infecção é a vacinação contra os tipos mais graves de HPV, causadores das lesões, que pode reduzir significativamente a chance de desenvolvimento do câncer. A Figura 2 apresenta dados estatísticos de incidência e mortalidade de câncer de colo nos países onde há maior número de registros.

Global cervical cancer statistics 2018: Estimates of incidence and mortality

Incidences

2nd most commonly prevailing cancers among women world wide
Majority High-Risk HPV-16 HPV-18 Infections
Slow growing, easily detectable at pre-cancerous stage



CA: A Cancer Journal for Clinicians, Volume: 68, Issue: 6, Pages: 394-424, First published: 12 September 2018, DOI: (10.3322/caac.21492)

Figura 2 – Incidência e Mortalidade de Cancer de Colo de Utero.
([Medical Xpress](#), 2018)

2.1.0.1 Diagnóstico Precoce

O diagnóstico precoce das lesões e do câncer propriamente dito, em estágio inicial, é importante pois além de aumentar as chances de recuperação do paciente, diminui os custos do tratamento em países onde o sistema de triagem não é eficaz. Infelizmente, a realidade em lugares que não investem recursos no combate é que os casos de câncer de colo uterino são identificados, muitas vezes, em estágios mais avançados, onde o tratamento não é tão eficaz, resultando em uma alta taxa de mortalidade. ([World Health Organization](#), 2021)

Os sintomas do câncer de colo de útero em situação avançada incluem: sangramento vaginal irregular, intermenstrual (entre períodos) ou anormal após a relação sexual; dor nas costas, pernas ou região pélvica; fadiga, perda de peso, perda de apetite, desconforto vaginal ou corrimento odoroso ([Healthline](#), 2018). Caso o câncer esteja em estágio mais avançado, sintomas mais graves podem aparecer. A OMS disponibiliza um pacote de orientação para intervenções na atuação primária, chamado de PEN (Package of essential noncommunicable), de pacientes com casos suspeitos de câncer de colo uterino, em regiões com poucos recursos disponíveis.

2.1.0.2 Triagem

A triagem é um processo que visa classificar os pacientes após exames de detecção de lesões pré-cancerígenas. Pessoas que são positivas para anormalidades, precisam ficar em observação para possível diagnóstico e tratamento, a fim de evitar o desenvolvimento do câncer. Em geral, esse processo necessita de recursos e investimento, porque os pacientes precisam passar por procedimentos de teste e acompanhamento, por isso a triagem não é encontrada ou é muito deficiente em países com uma fraca estrutura relacionada à prevenção de câncer. O objetivo da triagem é atingir o máximo de indivíduos possível, e principalmente alcançar aqueles que estão com risco de desenvolver a patologia. Por mais que seja um processo complicado de ser implementado, o custo benefício, para o local e para a vida da população, a longo prazo, é muito maior do que tratar o câncer quando já desenvolvido.

2.1.0.3 Abordagem de Prevenção ao Câncer de Colo em Larga Escala

A OMS recomenda que seja feita uma abordagem muito abrangente no que diz respeito às técnicas disponíveis para prevenir, controlar e tratar o câncer de colo uterino. Essas abordagens podem e devem ocorrer ao longo da vida de todas as pessoas suscetíveis, sendo cada uma indicada para cada estágio etário. Políticas como: educação sexual, mobilização do governo, vacinação contra HPV, triagem dos pacientes, tratamento e cuidados paliativos são de enorme eficácia no tratamento do câncer cervical. Segundo a própria OMS, a maioria das mortes por câncer de colo de útero podem ser evitadas se essas intervenções citadas forem implementadas e estiverem disponíveis para todas as pessoas, principalmente a vacinação contra o Papiloma Vírus.

Com o objetivo de combater o aumento dos casos de câncer, com a catalisação de ações nacionais e globais, as nações unidas criaram um programa global, em conjunto com agências mundiais, chamado UN Joint Global Programme on Cervical Cancer Prevention and Control. O programa foca em educar e direcionar a prevenção e tratamentos de câncer de colo em 3 pilares: Imunização de meninas contra o HPV, triagem e tratamento de lesões precursoras e diagnosticar e tratar cânceres já estabelecidos (incluindo cuidados paliativos) e estender isso para todas as pessoas. O programa foi iniciado em 6 países de média e baixa renda, por causa da taxa de mortalidade e incidência de casos nesses locais, são eles: Bolívia, Marrocos, República da Tanzânia, Quirguistão, Myanmar e Mongólia. A Figura 3 apresenta o resultado do programa de imunização feito na Inglaterra com meninas de 12 e 13 anos, administrando a vacina contra os subtipos 16 e 18 de HPV, comprovando a eficácia da prevenção. ([World Health Organization, 2016](#))

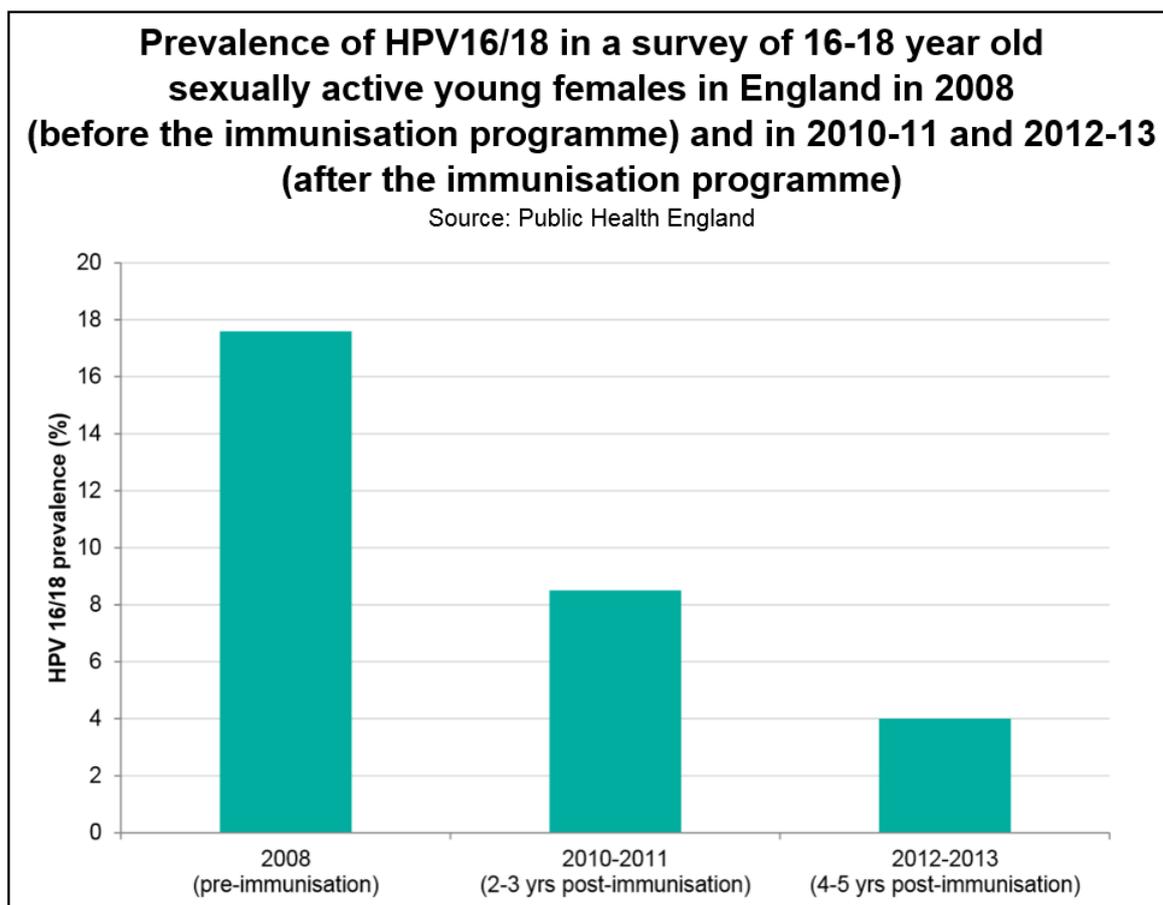


Figura 3 – Prevalencia de HPV em meninas após programa de imunização na Inglaterra.

([Vaccine Knowledge Project, 2014](#))

2.1.1 Vírus do Papiloma Humano - HPV

Os Papilomas vírus são vírus pertencentes à família Papillomaviridae com um DNA de fita dupla e circular. É considerado um vírus pequeno mas de alta complexidade. Quando está no estado sintomático, o paciente desenvolve lesões, placas esbranquiçadas e vermelhidão. Essas lesões são denominadas como condilomas, comumente conhecidas como verrugas, e se desenvolvem em áreas de membranas mucosas como boca, vagina e ânus, podendo se apresentar também na região orofaríngea, sendo responsável por 72% dos casos de câncer de boca (neoplasias que se desenvolvem na parte de trás da boca, amígdalas e garganta), segundo o CDC (Centers for disease control and prevention - EUA). O vírus tem, atualmente, mais de 150 tipos catalogados, e mais de 400 estimados, segundo a literatura. Classificado como uma infecção sexualmente transmissível (IST), o HPV é mais comumente adquirido através de relações sexuais sem proteção, também podendo ser transmitido de gestante para o feto, contato dermal ou através do compartilhamento de objetos contaminados. ([Centers for Disease Control and Prevention, 2020b](#))

O vírus tem três estados de vida em um organismo. Ele pode se manter latente no interior das células, não apresentando nenhum quadro clínico para o paciente, nesses casos o indivíduo não é vetor da doença, ou seja, não pode contaminar outras pessoas. Os outros dois estágios são infecção com e sem sintomas, onde o vírus HPV pode ser transmitido de pessoa para pessoa mesmo que não haja nenhum tipo de irritação aparente. Infecções por HPV podem demorar anos e até décadas para apresentar sintomas, fazendo com o que o indivíduo possa se tornar um vetor sem conhecimento.

Segundo estatísticas do CDC, praticamente toda a população sexualmente ativa irá entrar em contato com HPV em algum momento da vida e 4 de cada 5 mulheres irão adquirir HPV até alcançar os 50 anos. Felizmente, 95% dos casos irão ter resolução espontânea por ações do sistema imunológico, e boa parte até mesmo sem apresentar qualquer tipo de sintoma. O restante dos casos que não conseguem ser resolvidos podem se desenvolver em câncer de colo de útero causados pela evolução das lesões epiteliais ao longo do tempo. Os grandes fatores relacionados ao contágio por HPV são gênero, juventude e atividade sexual, segundo pesquisas do CDC. ([Centers for Disease Control and Prevention, 2020a](#))

As lesões causadas pelo HPV são classificadas de diversas formas de acordo com vários parâmetros, o ministério da saúde do Brasil utiliza uma classificação bem popular mundialmente, que se utiliza para diagnóstico histológico e divide as neoplasias intra-epitelial cervical pré-maligna (NIC) em três tipos: NIC1, displasia leve, NIC2, displasia moderada e NIC3, displasia grave ou Carcinoma in situ. Embora a maioria das NICs, de baixa intensidade, regredam, o desafio dos profissionais, ao identificá-las, é saber quais irão progredir para o carcinoma.

Como forma de prevenção contra HPV é recomendado o uso da camisinha, já que se trata de uma IST. Além disso, o mais importante avanço médico em relação ao vírus foi o desenvolvimento da vacina contra HPV, introduzida no mercado pela primeira vez em 2006. No Brasil, é distribuída pelo SUS (sistema único de saúde) para meninas de 9 a 15 anos, meninos de 11 a 15 anos e para mulheres e homens de 9 a 26 anos de idade que são portadores do vírus HIV a vacina quadrivalente. A vacina é chamada de quadrivalente pois trata os quatro subtipos de HPV que estão mais comumente ligados ao aparecimento de câncer de colo uterino (6, 11, 16, e 18, com prevalência dos subtipos 16 e 18, onde o 16 está relacionado a mais de 50% dos casos de câncer). Há também as vacinas bivalente, que trata os tipos 16 e 18 do vírus, e a nonavalente, que ataca os subtipos 6, 11, 16, 18, 31, 33, 45, 52 e 58, mas essas duas são administradas apenas por convênio particular, no Brasil. ([AULT, 2006](#))

2.2 Fatores Relevantes para o Risco

Nessa seção são apresentadas as variáveis que a literatura indica ter relação com o surgimento do câncer de colo de útero. Essas informações foram usadas para construção da aplicação para smartphone, explicada posteriormente no Capítulo 5. Cada uma delas tem influência direta ou indireta no desenvolvimento do câncer e estão relacionadas ao histórico do paciente, à epidemiologia da doença, aos fatores genéticos e/ou patologias e variáveis classificadas como atitudes de prevenção e/ou cuidado relacionadas a HPV e ao Câncer. Essa divisão foi importante pois a decisão sobre a forma como seriam apresentadas as perguntas na aplicação precisavam seguir um padrão para construir uma interface intuitiva. Essas variáveis foram estudadas e selecionadas pela Dr. Iolanda Matias, médica gineco e oncologista especializada em aparelho reprodutor feminino, com anos de experiência na atuação contra o câncer de colo uterino. A orientação dela foi muito importante para o entendimento dos fatores que são impactantes na patologia, e na relevância de cada um para definição do fator de risco.

2.2.1 Idade

A idade é um fator muito importante porque diz respeito ao comportamento sexual e a chance de desenvolver o câncer, pois a chance de uma pessoa contrair HPV e desenvolver a doença em idade mais avançada é mais baixa do que em um indivíduo jovem. Isso se deve ao fato de que a porcentagem de pessoas de meia idade que tem uma vida sexual ativa com vários parceiros é baixa, a chance de adquirir HPV é igualmente proporcional. Como uma lesão pode levar anos até desenvolver um câncer, uma pessoa que se previne e não acusou anomalias nos exames até 60 anos, por exemplo, tem uma porcentagem muito baixa de contrair a IST.

2.2.2 Idade da Primeira relação sexual

A segunda variável dessa classe é a Idade da primeira relação sexual. Trazendo uma gama de informações sobre a vida sexual do paciente, através desse dado é possível inferir a exposição ao vírus, uma vez que, se o paciente não teve primeira experiência sexual, a chance de contrair HPV é zero, exceto para bebês que contraíram HPV da mãe no parto. Uma pessoa que teve a primeira relação sexual em idade infantil, possivelmente foi vítima de um abuso, e como na maioria dos casos, houve lesão e sem presença de preservativo, mais uma vez, aumentando o risco do contágio por HPV. A idade da primeira relação sexual é importante porque quanto mais nova a pessoa, menos desenvolvido está o seu colo do útero e menos maturação nas células da região.

2.2.3 Gestações, Partos e Abortos

Essas três variáveis andam em conjunto: número de gestações, número de partos e número de abortos. Elas dizem respeito a quantidade de lesões sofridas na região do colo uterino. Quanto maior o número de gestações, maior estresse na região uterina. O número de partos é diferente do número de gestações porque a disparidade indica um alto índice de abortos, quanto mais partos, maior a quantidade de lesões e traumas na região do colo, e um alto número de abortos pode ser indicativo de diversos fatores, desde alterações cromossômicas no feto até problemas imunológicos e/ou uterinos na mãe, nesse caso, estamos interessados no último.

2.2.4 Alimentação

A presença de vitaminas antioxidantes A, C e E é benéfica para a prevenção do câncer de colo uterino, principalmente nos estágios iniciais da formação das lesões. Os compostos antioxidantes têm grande significância do combate às disfunções no organismo porque agem nas três principais linhas de defesa orgânicas do corpo: prevenção (lutando contra a formação de substâncias agressoras), interceptação (parando a ação dos radicais livres) e reparo (que age quando as duas primeiras não têm êxito). As vitaminas C e E são importantes para a manutenção do sistema imunológico, além de dificultar o surgimento dos carcinógenos, ambiente propício ao aparecimento de câncer. Enquanto o caroteno, presente da vitamina A, é capaz de inibir a formação de radicais livres (moléculas bastante instáveis no organismo, capazes de desenvolver doenças degenerativas e influenciar em formações cancerígenas). Como esses compostos estão presentes em frutas e verduras como laranja, cenoura, abacaxi, vegetais escuros etc. Se fez necessário o mapeamento dessa informação. (SAMPAIO; ALMEIDA, 2009)

2.2.5 Localização

Localidade é uma variável que se baseia na estatística cedida no site da organização mundial da saúde, que determina o índice de incidência de câncer cervical por 100.000 habitantes em cada país. É importante mapear esse dado porque a localização do paciente, no globo, demonstra o tipo de cultura, educação básica e sexual, qualidade de sistema de saúde e alimentação que ela tem acesso. Além de outros fatores secundários como clima e saneamento básico. Países como Canadá, que tem uma política de saúde pública muito rígida, relacionada à prevenção ao câncer de colo, se encontra entre os menores índices de desenvolvimento de casos. Nesse país, os profissionais de saúde tratam a lesão assim que detectadas, sem dar chance para o desenvolvimento do carcinoma. O Brasil se posiciona em mediano nessa relação de países, com uma política que atende mas não é muito eficiente em prevenir. Na Figura

4 pode ser visto que em países do centro e sul africano, onde o sistema de saúde não tem eficiência em conscientizar e prevenir, o alto índice é predominante. Além disso, outros fatores são responsáveis por esses índices, especialmente nos países que vivem em guerra civil, como abusos e fragilidade imunológica.

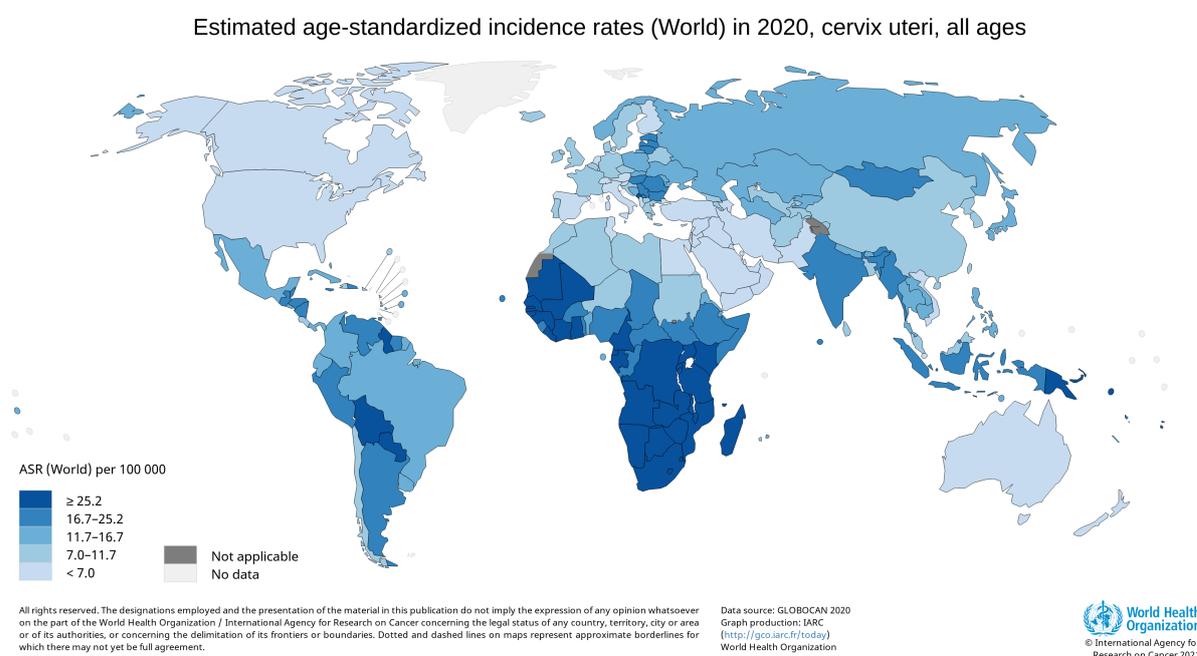


Figura 4 – Índice de incidência de câncer de colo de útero para cada 100.000 habitantes.

(International Agency for Research on Cancer, 2018)

2.2.6 Presença de ISTs

A presença de outras infecções sexualmente transmissíveis, além do HPV, pode aumentar a chance do paciente desenvolver câncer de colo de útero. Em diversos estudos, as pessoas são classificadas como portadoras ou não de outras ISTs, no grupo positivo a vulnerabilidade ao câncer de colo uterino é sempre mais alta. Para nosso estudo, consideramos 4 doenças: Aids, que ataca o sistema imunológico do paciente o deixando vulnerável a outras enfermidades; Hepatite B, transmitida por via sexual, a doença ataca o fígado e pode ser responsável por diversas males como cirrose e insuficiência hepática; sífilis, doença causada por bactéria que pode atacar a vagina e útero, causando feridas e podendo prejudicar na gestação; gonorreia, infecção bacteriana que atinge principalmente o colo do útero.

2.2.7 Tabagismo

O tabagismo é um hábito predominantemente masculino mas que começou a se difundir entre todas as pessoas no começo do século XX. É sabido que além de ser

um forte influenciador no surgimento do câncer de pulmão, o consumo de cigarros, principalmente com nicotina, é objeto de estudo como fator de risco de vários outros tipos de doenças, incluindo câncer de colo. Segundo o INCA, mulheres que fumam antes da gravidez tem duas vezes mais probabilidade no atraso da concepção devido a concentração de nicotina no fluido folicular do ovário. Por essa razão, mapear o uso de cigarros é muito importante para o trabalho. ([American Cancer Society, 2020](#))

2.2.8 Histórico Familiar

Nesta subseção, são relacionados os fatores de herança genética. Como é sabido, o surgimento de alguns tipos de câncer se dá por influência genética e predisposição a patologia. Um famoso caso é o da atriz Angelina Jolie, que no ano de 2013 se submeteu a uma mastectomia bilateral profilática, removendo as mamas completamente como tratamento de prevenção ao câncer de mama. Após um mapeamento genético foi observado que a atriz possuía um gene associado ao desenvolvimento de câncer defeituoso, com chance de 87% de desenvolvimento de carcinoma mamário. O histórico familiar de Angelina justificava a alteração, ela perdeu a mãe, avó e tia para os cânceres de ovário e mama, sendo que na mãe houve incidência dos dois tipos. Dois anos após a remoção mamária, a atriz recebeu o diagnóstico precoce de um possível câncer de ovário, novamente de submetendo a uma operação preventiva, dessa vez de ooforectomia (remoção dos dois ou um ovário). Após terapia hormonal, necessária por causa das duas cirurgia, Angelina desenvolveu menopausa precoce.

A partir disso, é notado que o estudo do histórico familiar do paciente é importante, esses casos são chamados de cânceres hereditários (correspondem a uma taxa de 5 a 10% de todos os casos). A presença de casos de câncer em parentes de primeiro grau (pai e mãe) foi considerada um fator que corrobora com a possibilidade de haver mutações genéticas. Além disso, a presença de outros tipos de neoplasias específicas é indicativo de preocupação, por isso elas foram especificadas neste teste. Para quantificar isso, a usuária deve responder se houve histórico de casos de câncer de mama, câncer de útero, câncer de tireóide ou câncer de intestino nos pais. ([Instituto Oncoguia, 2015](#))

2.2.9 Diabetes, Hipertensão e Uso de Corticoides

A diabetes do tipo 2 e obesidade podem ser associadas ao desenvolvimento de alguns tipos de câncer, dentre eles o câncer de colo de útero, devido a um processo chamado síndrome metabólica, que acaba aumentando as chances de danos celulares decorrentes da produção de fatores inflamatórios, radicais livres e espécies reativas de oxigênio. Tudo isso associado ao estresse metabólico causado pela hipertensão e a diabetes aumentam as chances de desenvolver mutação neoplásica. Os corticoides

por sua vez estão associados a imunossupressão que o uso contínuo desses medicamentos causa, o que acaba baixando as barreiras do organismo no reconhecimento de defeitos celulares e facilitando infecções.

2.2.10 Número de Parceiros

As últimas perguntas são mais relacionadas a fatores ligados à prevenção do câncer, foram destacadas variáveis que quantificam a exposição do paciente ao HPV, diante disto, é importante saber qual foi o número de parceiros no último ano. Esses parceiros são apenas considerados do sexo masculino, pois poucos são os casos de transmissão de HPV em casais de lésbicas que evoluíram para câncer, a estatística é irrelevante. Além disso, está associado ao comportamento masculino características como pouca preocupação com prevenção, pouco cuidado com ISTs e alto número de parceiras, por isso a vacina para o HPV é dada para adolescentes de ambos os sexos, uma vez que os homens, epidemiologicamente falando, são o vetor de transmissão de HPV.

2.2.11 Uso de Preservativo

O uso de camisinha é fortemente ligado à prevenção, uma vez que é, hoje, o meio mais eficaz de se proteger de qualquer infecção sexualmente transmissível, não só HPV mas HIV, sífilis, gonorreia e Hepatite, já citadas e incluídas na estatística. É importante frisar que essa informação é muito curiosa, uma vez que segundo o ministério da saúde (Brasil), 6 em cada 10 jovens, de 15 a 19 anos, tem relações sexuais sem preservativo, justamente a idade onde, para as meninas, a região uterina está com alta taxa de atividade celular.

2.2.12 Citologia Oncótica

Papanicolau ou citologia oncótica é um exame que consiste na observação e coleta do material do colo uterino, na região de junção escamocolunar, local onde as lesões podem acontecer. Esse procedimento serve para rastrear lesões que podem ser precursoras do câncer. O resultado do exame pode ser negativo, indeterminado, lesão de baixo grau (LSIL, que indica uma displasia branda, ou seja, uma lesão com poucas chances de desenvolver câncer cervical, causada por qualquer tipo de HPV) ou lesão de alto grau (HSIL, indica que há alterações de tamanho e formato nas células anormais, indicando que é possível haver lesão NIC 3 ou até mesmo um câncer já estabelecido, por isso, caso esse seja o resultado, o paciente precisa se submeter a colposcopia e biópsia). Além disso, o exame serve para detectar alterações bacterianas na região do colo uterino.

2.2.13 Colposcopia

A colposcopia, citada no último tópico como exame importante em caso de HSIL é um procedimento feito de forma visual, com um aparelho chamado colposcópico, que funciona como um microscópio para o médico, aumentando a capacidade de visualização dentro da região do colo uterino, na vagina e na vulva. Identificada a lesão e indicada para colposcopia, o resultado pode ser atípico (ou anormal, em casos de lesões sugestiva ou cancro) ou normal (satisfatório, onde toda a região indicada é avaliada e nenhuma lesão precursora é identificada). A colposcopia e o papanicolau são exames feitos gratuitamente pelo SUS e muito utilizados na prevenção.

2.2.14 Teste de HPV

O HPV teste é um dado que deve ser lido com bastante cautela. No Brasil, o SUS não faz o teste de forma gratuita, então, boa parte dos pacientes não teriam o resultado desse exame para responder a questão. Quando ele é feito, em geral por convênio médico privado, o paciente recebe o resultado se tem ou não HPV e qual a classificação do seu tipo, oncogênico ou não-oncogênico, essa classificação é feita identificando o subtipo do vírus, 18 e 16 (combatidos pela vacina Bi e Quadrivalente) são os tipos mais associados ao desenvolvimento de câncer, quando o subtipo não está nesse conjunto, o exame tem resultado não-oncogênico. Em geral, esse exame apenas é feito quando o papanicolau dá alguma alteração. Mais uma vez, isso pode ser associado a idade da pessoa, retomando a ideia do comportamento sexual, se ela já tem uma idade acima dos 60 anos e o exame de HPV entrega que não há presença do vírus no organismo, a chance de ela contrair é muito baixa, excluindo-a de um possível grupo de alto risco.

2.2.15 Vacina Contra HPV

O último dado e talvez o mais importante relacionado a prevenção é a vacina contra o HPV, distribuída pelo Sistema Único de Saúde para jovens de ambos os sexos, é o modo mais eficaz de combater a infecção sexualmente transmissível, de forma a se garantir a eficiência da medicação, a vacina deve ser renovada a cada cinco anos.

3 Modelo Usados para Estimar Risco em Doenças

3.1 Trabalhos Relacionados

Yılmaz e Ayan (YILMAZ; AYAN, 2013), apresentam um estudo, baseado em lógica fuzzy sobre três tipos de câncer: Pulmão, Mama e câncer colorretal, escolhidos por causa da alta taxa de incidência. O objetivo deles é calcular a chance do paciente desenvolver os cânceres, fundamentado nos fatores de influência, entender a associação do estresse com o surgimento das doenças e desenvolver uma aplicação para ser usada por médicos e pacientes, capaz de responder o risco de surgimento. Para fundamentar a pesquisa, foram usados dados de pacientes cedidos pelos serviços de oncologia, neles constava as variáveis captadas pelo estudo e a divisão entre pacientes diagnosticados com o tipo de câncer examinado e pessoas que estão em observação mas não tiveram o câncer confirmado. Para o câncer de mama por exemplo, as variáveis usadas foram: sexo, idade, idade da menarca, idade da menopausa, consumo e álcool, hábitos nutricionais, relação genética e idade da primeira gravidez. Com essas informações, os cientistas construíram 3 modelos adaptados para cada realidade, utilizando regras de inferência fuzzy e funções de pertinência para calcular o grau de pertencimento dos pacientes em quatro categorias: saudável, ligeiramente em risco, risco, ou risco excessivo.

Laghetto (LAGHETTO, 2016), teve como objetivo estudar formas de modelar matematicamente incertezas por meio da teoria de conjuntos fuzzy e a partir disso, construir um modelo para calcular a chance de uma pessoa desenvolver câncer de pulmão. Para esse cálculo foram considerados fatores de risco relacionados ao surgimento da doença, foram eles: tabagismo, poluição, histórico de doenças pulmonares, histórico familiar e contato com agentes químicos. A escolha da doença foi feita devido ao diagnóstico tardio do câncer de pulmão, como não apresenta sintomas em estágios iniciais, a descoberta é feita apenas em etapas mais avançadas. A proposta da autora foi de definir os fatores de risco como variáveis e a classificação deles como conjuntos fuzzy de um modelo, por exemplo, para a variável poluição há dois conjuntos relacionados: pouca e muita, o modelo vai calcular o grau de pertencimento do fator em relação aos conjuntos. O modelo foi construído no software Matlab, usando o método de inferência de Mamdani e a função centroide para defuzzificar as respostas. O modelo trouxe resultados satisfatórios para a pesquisadora, estando de acordo com a literatura mas não pôde ser validado com dados reais de pacientes investigados ou confirmados

com câncer de pulmão.

Bassani (BASSANI, 2016), desenvolveu um modelo baseado em lógica fuzzy para estudar o desenvolvimento de doenças cardiovasculares e propor uma forma alternativa de prever o risco da incidência. Para construção do sistema foram estudadas as variáveis idade, pressão sistólica, tabagismo, diabetes e índice de massa corporal de cada paciente. A autora destaca um método já conhecido de observação da saúde cardiovascular do paciente utilizando o Escore de Risco de Framingham, que consiste em uma contagem de pontos obtidos a partir dos dados idade, gênero, colesterol total, lipoproteína de alta densidade (HDL) colesterol, pressão arterial sistólica e tabagismo. Apesar desse recurso ser reconhecido mundialmente, muitos pesquisadores da área afirmam ser superestimado, segundo a autora. As variáveis foram selecionadas utilizando análise de correlação linear. O modelo foi validado com dados de pacientes reais. O paciente é classificado com pertinência em cinco conjuntos, são eles: baixo, médio baixo, médio, médio alto e alto. O método de inferência utilizado foi Mamdani e o software Matlab. Os resultados foram satisfatórios com algumas divergências de um teste executado com o Escore de Risco de Framingham mas justificadas pela diferença geográfica de construção dos dois modelos além da carência de dados femininos. Uma das vantagens do modelo é a capacidade de observar a evolução do risco individual de cada paciente, fixando as variáveis, alterando a idade e utilizando equações diferenciais.

Nogueira, Ximenes e Brito (SOUSA; XIMENES; BRITO, 2020), Apresentam um modelo baseado em lógica fuzzy para auxiliar no diagnóstico da Covid-19. Para mapear o comportamento do vírus no organismo humano foram observados os principais sintomas associados: tosse seca, febre, cansaço, perda de olfato e paladar, dor de garganta, dor de cabeça, dor no peito e falta de ar. A partir disso, foi definido um único conjunto formado pelas variáveis linguísticas: leve, médio e grave. O estudo consiste em observar o grau de pertencimento de cada sintoma para as funções citadas para inferir a chance do paciente ter coronavírus. O método escolhido para classificação foi o Método de Sugeno. O modelo não foi validado com casos reais e o autor sugere a construção de uma alternativa com método de Mamdani de inferência, a fim de comparar os resultados obtidos e ver qual se aproxima mais da realidade.

Modelos matemático-computacionais têm sido usados em aplicações específicas para o câncer cervical. Cabrera-Bacerril, Vargas-De-Léon, Hernández, Miramontes e Peralta (CABRERA-BECERRIL et al., 2017), apresentam dois modelos para tentar compreender como se dá o comportamento cromossômico nas células precursoras do câncer de colo de útero através das mutações delétricas e não-deletérias. O primeiro é um modelo baseado em agentes, muito utilizado em estudos de sistemas complexos, onde há muitas interações. A ideia principal da análise com essa técnica é replicar,

por meio de estímulos, algumas das interações entre os componentes do sistema em observação, nesse caso, seguindo três pontos principais: os agentes, as regras que fazem as interações entre os agentes e o ambiente de interação. As células presentes nos tecidos foram representadas como os agentes, e elas recebem sinais externos do ambiente em que estão inseridas e dos agentes ao redor, dessa forma, cada célula interage com as outras ao redor (chamada de vizinhança). O segundo modelo é o de autômatos celulares, que permite que sejam feitas idealizações de sistemas físicos, por causa do caráter discreto e contínuo do modelo. Como as características celulares podem ser classificadas como discretas, dependendo dos conjuntos de células da vizinhança regras podem ser atribuídas para transições celulares. Por fim, os autômatos foram utilizados para modelar os comportamentos celulares (as células em si) enquanto o modelo ABM foi utilizado para simular as dinâmicas celulares, para determinar a atividade das alterações deletérias e não deletérias nos tecidos do colo uterino. As conclusões referentes ao estudo das células deletérias e não deletérias mostrou que nos casos onde o câncer de colo prevaleceu, a presença de células não deletérias foi maior, em séries temporais. Analisando as amostras dos experimentos, foi observado que as células não deletérias se comportavam como um mecanismo de seleção natural, ao passo que as deletérias sofrem morte celular. Desse modo, em lesões precursoras do câncer, a presença de mutações deletérias foi mais forte, enquanto as não deletérias se mostraram maior quando o câncer já se estabeleceu no tecido, favorecendo o tumor com características presente nas mutações celulares.

Saldaña, Korobeinikov e Barradas ([SALDAÑA; KOROBEINIKOV; BARRADAS, 2019](#)) utilizam um modelo compartimental baseado em equações diferenciais para investigar as melhores estratégias de vacinação e triagem relacionadas ao câncer de colo de útero, e assim minimizar a taxa de mortalidade da doença e otimizar os custos de intervenção. O modelo divide as pessoas por gênero e trabalha com três variáveis ao longo do tempo: vacinação em adolescentes, vacinação em adultos e triagem, as três podendo acontecer simultaneamente. A pesquisa é dividida em dois casos de problemas funcionais: O primeiro corresponde à proteção do grupo vulnerável e o problema de controle consiste em minimizar o nível cumulativo de mulheres infectadas em um intervalo de tempo fixo. O segundo funcional visa eliminar a infecção e o problema de controle consiste em minimizar a prevalência total no final do tempo estipulado. O trabalho questiona as vantagens e custos de estender a vacina contra HPV para ambos os sexos. O objetivo é usar uma teoria que eles chamam de optimal control theory para combinar a vacina e a triagem de forma que reduza o espalhamento do vírus e consequentemente o custo de intervenções. Para isso foi preciso examinar a dinâmica do espalhamento do HPV na sociedade e formular duas funções objetivo para simular as situações. O resultado da pesquisa foi que a vacina deve ser administrada tanto antes quanto depois do início da vida sexual, tanto em homens quanto em mulheres.

As duas simulações mostraram que a vacinação deve priorizar as mulheres, como já era esperado, mas ficou provado que a vacinação de homens é quase tão benéfica quanto da parcela feminina. Se o objetivo for de proteger os grupos vulneráveis, as simulações numéricas mostraram que devem ser aplicadas altas taxas de vacinação, inicialmente, que podem ser reduzidas gradualmente, levando a uma taxa baixa de indivíduos infectados. Quando o objetivo é a erradicação da infecção, os experimentos mostraram que os controles ideais começam em taxas baixas, permitindo que a prevalência aumente, no entanto, as taxas ótimas de vacinação e triagem aumentam ao longo do tempo, alcançando a eliminação da infecção no final do intervalo de tempo.

3.2 Análise de Risco para Câncer de Colo de Útero

Muitos modelos para estimar risco de desenvolver câncer cervical são baseados em métodos estatísticos sobre dados reais. Em (AL-MADANI et al., 2019), foi proposto um estudo para melhorar a rastreabilidade de câncer de colo uterino usando fatores de risco além do exame papanicolau. Foram usados dados de 300 mulheres sintomáticas com suspeita de serem portadoras do câncer cervical, coletados a partir de um formulário aplicado entre elas para obter as informações: idade, altura, peso, doenças crônicas, uso de anticoncepcionais e histórico de câncer na família. Além disso, dados sobre os resultados dos exames dessas pessoas foram recuperados e associados para realizar a pesquisa. Para examinar a associação dos fatores com a presença de tecido cervical maligno foram usadas técnicas de análises qualitativas. Por fim, os quatro fatores que mais se associaram ao alto risco de desenvolver câncer foram: histórico familiar, sangramento vaginal, hipertensão e resultado anormal do papanicolau.

Em (LANDY et al., 2018), os autores examinam de forma crítica três métodos de análise de registros eletrônicos de saúde para estimar risco de desenvolvimento de câncer de colo de útero, são eles: método de Kaplan-Meier, estimador de Turnbull e modelo de Weibull, esses dois últimos foram adaptados para a realidade estudada. O objetivo é mostrar como a escolha do método estatístico impacta na estimativa final de risco. Para executar os testes foram usados dados de 41.067 mulheres com idades de 30 a 64 anos que fizeram ao menos um teste de rastreio cervical. O estudo concluiu que para dados que são tratados em intervalos, o método de Kaplan-Meier não se comportou bem porque produziu resultados tendenciosos, como muitos registros eletrônicos trabalham com intervalos grandes, o método não é apropriado. A estimativa feita por Turnbull foi imparcial, porém instável. O algoritmo de Weibull assim como o Turnbull superestimaram o risco de doenças prevalentes em amostras grandes, mas ele teve um bom desempenho, exceto quando os tempos dos eventos não seguiram uma distribuição Weibull.

Bicheng et al. ([BICHENG et al., 2014](#)), propuseram um modelo para estimar o risco de câncer de colo de útero usando análise de regressão múltipla e uma rede neural artificial para identificar os fatores de risco associados ao câncer. Os pesquisadores coletaram amostras de sangue e informações pessoais de 68 pacientes diagnosticadas com carcinoma nas células cervicais em um hospital na China. Os dados das pacientes foram inseridos em um algoritmo de rede neural artificial e usadas funções logarítmicas e lineares para calcular a variável de saída. A pesquisa concluiu que a presença do vírus HPV e dos alelos HLA-DRB1 * 13-2 e HLA-DRB1 * 3 aumenta a chance de desenvolvimento do câncer de colo uterino, assim como a relação entre as variáveis idade do primeiro parto e escolaridade também estão associadas a chance. Dessa forma, o estudo pode ajudar a selecionar os pacientes que precisam passar por tratamentos de intervenção para a doença. Os testes e comparações com dados concluíram que o modelo teve uma taxa assertiva de 95% no grupo de mulheres com câncer.

Rothberg et al. ([ROTHBERG et al., 2018](#)), desenvolveram um estudo estatístico baseado em uma base formada por 99319 mulheres a partir dos registros eletrônicos de saúde, coletando dados demográficos e de histórico médico, foram eles: idade, raça, tabagismo, cobertura de plano de saúde, estado civil e renda mediana, para prever o surgimento de neoplasias intraepiteliais cervicais 2 (NIC 2) ou superiores, pois essas lesões são consideradas perigosas no surgimento do câncer de colo de útero. No final do estudo eles compararam os resultados obtidos no modelo com os casos reais. O estudo utilizou mulheres de mais de 30 anos que tinham feito teste de HPV e papanicolau. Após análise ficou concluído que todos os fatores, exceto raça, estavam relacionados ao surgimento do câncer.

4 Teoria De Conjuntos Fuzzy

Nessa seção serão utilizados (LAGHETTO, 2016), (SABINO, 2016), (BASSA-NEZI; BARROS, 2010) e (SILVA, 2011) para explicar e contextualizar a teoria dos conjuntos fuzzy, história, funcionamento e aplicações, além de instrumentos que foram usados para construir o modelo de cálculo de risco descrito no Capítulo 5.

4.1 História da Teoria Fuzzy

Em 1965, na Universidade de Berkley, na Califórnia, um matemático, engenheiro eletrônico e cientista da computação chamado Lotfali Askar-Zadeh introduziu o conceito de lógica difusa ou lógica nebulosa. Para ele, a lógica se diferenciava da matemática clássica por quebrar a ideia de verdade absoluta, ou seja, ao invés de um método booleano onde uma proposição só pode ser verdadeira ou falsa, a lógica batizada como fuzzy trabalha com o conceito de intervalo entre completamente verdadeiro e completamente falso, $[0,1]$.

O estudo de Zadeh começou em meados da década de 1960, quando ele observou que os sistemas clássicos de representação não conseguiam propor um modelo ideal para problemas de natureza industrial ou biológica que resultassem em respostas ambíguas e incertas. Após a publicação do seu artigo teorizando os conjuntos fuzzy, o conceito começou a ser estudado mais a fundo por outros matemáticos, um deles foi o Professor Ebrahim Mamdani da Queen Mary College, Londres, idealizador do modelo de inferência que leva seu nome (explicado na seção 4.7 deste Capítulo). Em 1974 Mamdani conseguiu controlar uma máquina a vapor utilizando os conceitos de lógica difusa propostos por Zadeh. A lógica Fuzzy então começou a ser implementada em campos de controle de operações em indústrias, refinarias, processos químicos, plantas nucleares e sistemas de operações de trens.

4.2 Conceitos Básicos

A lógica fuzzy se destaca por representar o mundo em que vivemos de um modo não discreto e incerto, contendo variáveis que não são absolutamente verdadeiras nem falsas. No nosso cotidiano usamos conceitos subjetivos para expressar sentimentos em relação às coisas, por exemplo: Pedro é alto; siga alguns metros e dobre a direita. A definição de alto ou baixo na determinação da altura de Pedro é um conceito puramente individual e incerto, o mesmo vale para a expressão "alguns metros". Fuzzy significa difuso ou nebuloso, o contexto fuzzy pode ser representado por situações que

se pode responder sim ou não com certeza e é feito uso de adjetivos ou substantivos que expressam essas respostas de forma verbal.

Na teoria clássica de conjuntos, o conceito de pertinência é bem definido em **pertence** ou **não pertence**. Observe a função a seguir tomando o conjunto A e verificando o elemento x:

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

Na teoria fuzzy há um grau de pertinência de **cada elemento** para os conjuntos tratados. Por exemplo:

- conjunto das pessoas ricas.
- conjunto das pessoas de altura mediana.
- conjunto das pessoas que tem colesterol alto.

Para cada conjunto em questão não há um valor (fronteira) exato que diga se um elemento pertence ou não a eles, mas podem ser criados intervalos e observar o **grau de pertencimento** de cada elemento para cada conjunto abordado. Abaixo encontra-se uma representação gráfica do comportamento da lógica clássica (também chamada de crisp) e lógica fuzzy:

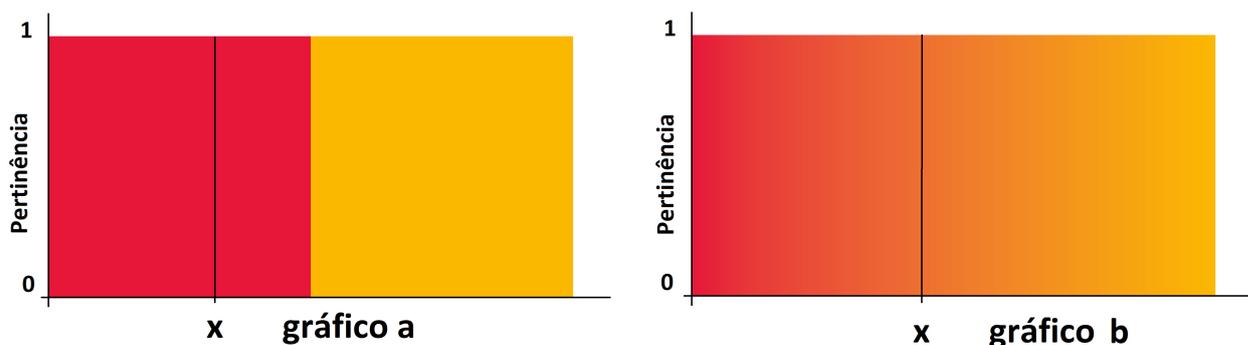


Figura 5 – diferença entre lógica crisp(a) e lógica fuzzy(b)

Na Figura 5, pode ser observado que no gráfico que representa um sistema clássico (a) o elemento x tem um fronteira bem definida delimitando as duas regiões desenhadas e faz um salto brusco entre elas a partir de um valor exato. Já no gráfico b, que representa um sistema fuzzy, é notado que as regiões se mesclam e não é possível determinar em que momento há a passagem do elemento x do vermelho para o amarelo. Nessa imagem, as duas cores representam conjuntos e as diferentes tonalidades entra elas é o que mais tarde será explicado como função de pertinência.

4.3 Conjuntos Fuzzy

Os conjuntos fuzzy são formados a partir do domínio (conhecido como conjunto Universo ou U). Para cada conjunto fuzzy há uma **função de pertinência** que determina quanto um elemento faz parte daquele conjunto, para essa classificação é usado um intervalo numérico de 0 a 1. Um subconjunto fuzzy A de U é representado pela função de pertinência:

$$\mu_A : U \rightarrow [0, 1].$$

O grau de pertencimento de cada elemento x do conjunto U no subconjunto fuzzy A é representado por:

$$\mu_A(x)$$

Exemplo: Seja o conjunto U um conjunto universo que representa uma lista de alturas representado pelos valores $\{ 1.30, 1.50, 1.70, 1.80, 1.90 \}$. Sejam A , B e C conjuntos fuzzy que representam as classificações da altura de uma pessoa, nessa ordem: baixo, médio e alto. Podem ser atribuídos graus de pertinência dos elementos do conjunto universo para os conjuntos fuzzy, descritos na Tabela 1:

Altura	baixo	médio	alto
1.30	1	0.2	0
1.50	0.6	0.8	0.3
1.70	0.2	0.9	0.6
1.80	0	0.5	0.9
1.90	0	0	1

Tabela 1 – Exemplo de grau de pertinência para o conjunto Universo.

Na Tabela 1 é possível observar que cada elemento do conjunto Universo tem um grau de pertinência para os conjuntos fuzzy baixo, médio e alto. Por exemplo, o elemento 1.50 pertence ao conjunto baixo com grau 0.6, ao conjunto médio com grau 0.8 e ao conjunto alto com grau de pertencimento 0.3.

Alguns conceitos são importantes para serem notados na teoria dos conjuntos Fuzzy. O primeiro deles é o **suporte de um conjunto fuzzy A** no conjunto Universo (U), que indica o conjunto de todos os elementos de U que tem grau de pertinência maior que zero, e é dado da seguinte maneira:

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

Por exemplo, Na Tabela 1, o suporte do conjunto fuzzy baixo é o conjunto clássico:

$$\text{sup}(\text{baixo}) = \{1, 0.6, 0.2\}$$

O próximo conceito é a **cardinalidade de um conjunto fuzzy A** sobre um conjunto universo U, que é dada pela soma de todos os graus de pertencimento dos elementos de A. É representada como:

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

Por exemplo, na Tabela 1, a cardinalidade do conjunto fuzzy alto é dada por:

$$|alto| = 0 + 0.3 + 0.6 + 0.9 + 1 = 2.8$$

4.4 Operações Entre Conjuntos Fuzzy

Assim como os conjuntos clássicos, os conjuntos fuzzy também estão sujeitos à operações, são elas:

- O conjunto fuzzy A é **subconjunto** do conjunto fuzzy B se o grau de pertinência de todos os elementos do conjunto A é menor que o grau de pertinência de todos os elementos do conjunto B. É indicado por:

$$A \subseteq B.$$

- O conjunto A é **igual** ao conjunto B se o grau de pertinência de todos os elementos de A for igual ao grau de todos os elementos de B. É indicado por:

$$A = B.$$

- O conjunto A **não é igual** ao conjunto B se pelo menos **um** elemento do conjunto A tem grau de pertinência que não existe no conjunto B. É indicado por:

$$A \neq B$$

- O conjunto fuzzy A é **subconjunto próprio** do conjunto fuzzy B quando A é subconjunto de B e $A \neq B$.

- A **união** dos conjuntos Fuzzy A e B é o conjunto Fuzzy $A \cup B$ tal que

$$\forall x \in U \mu_{A \cup B}(x) = \max(\mu_A(x), \mu_B(x))$$

Exemplo, na Tabela 1 a união entre os conjuntos fuzzy médio e alto é dada pelo conjunto $\{0.2, 0.8, 0.9, 0.9, 1\}$.

- A **intersecção** dos conjuntos Fuzzy A e B é o conjunto Fuzzy $A \cap B$ tal que

$$\forall x \in U \mu_{A \cap B}(x) = \min(\mu_A(x), \mu_B(x))$$

Na Tabela 1, sejam os conjuntos baixo e médio, a intersecção entre eles é conjunto $\{0.2, 0.6, 0.2, 0, 0\}$.

- O **complemento** do conjunto fuzzy A em relação ao conjunto Universo U é descrito por A' e sua função de pertinência é:

$$\mu'_A(x) = 1 - \mu_A(x), \forall x \in \mathbb{U}$$

Por exemplo, na Tabela 1, o elemento 1.50 tem grau de pertinência 0.3 no Conjunto fuzzy alto, para o conjunto alto' seu grau de pertinência é 0.7.

4.5 Normas e Conormas Triangulares

Normas triangulares são operações usadas para generalizar o conceito de união (t-conorma) e intersecção (t-normas). Elas foram propostas inicialmente por Zadeh, sendo descritas como operações de soma (∇) e produto algébrico (Δ), representados pela nomenclatura Max e Min respectivamente. As equações abaixo representam essas duas operações:

$$\mu_{A \cup B}(x) = \mu_A(x) + \mu_B(x) - \mu_A(x) * \mu_B(x) \text{ (soma algébrica)}$$

$$\mu_{A \cap B}(x) = \mu_A(x) * \mu_B(x) \text{ (produto algébrico)}$$

Mais a frente será explicado o conceito de método de Mamdani que utiliza a ideia de t-norma e t-conorma para calcular a inferência das variáveis no sistema fuzzy.

4.6 Sistema Baseado em Regras Fuzzy

Nesta seção será apresentado o sistema baseado em regras Fuzzy, também conhecido como SBRF, um tipo de sistema utilizado para modelar os problemas do nosso cotidiano baseado na lógica fuzzy.

4.6.1 Regras de Inferência

Regras de inferência são proposições lógicas que se combinam para formar uma sentença lógica do tipo "se x é A e y é B , então z é C", elas são construídas relacionando os conjuntos fuzzy de entrada com os de saída. Essas regras são desenvolvidas utilizando os conjuntos fuzzy selecionados no problema a ser abordado e podem contar diversos elementos que concordam para chegar a algum resultado. Existem diversos modos de se construir regras de inferência, por isso é importante observar a estrutura do sistema fuzzy escolhida para construí-las. Elas são fundamentais na modelagem de um sistema fuzzy pois fazem parte da etapa de inferência, como mostrado abaixo:

"Se Altura é baixa e Peso é alto, então IMC é alto"

Pode ser notado que no exemplo acima Altura e Peso são os conjuntos fuzzy de entrada, também conhecidos como **antecedentes** ou premissas, enquanto o con-

junto fuzzy IMC faz parte da saída do problema, também conhecida como parte **consequente** ou de conclusão. Na etapa de inferência é onde ocorre a combinação desses dois elementos.

Existem diversas formas de trabalhar com as variáveis, isso é o que diferencia os métodos de inferência, um dos mais conhecidos é o método de inferência de Mamdani que será mais explicado na seção 4.7 desse mesmo Capítulo. Basicamente, o modelo de Mandani mapeia os antecedentes e consequentes em conjuntos fuzzy para calcular um resultado. Outro método bastante conhecido também é o de Takagi-Sugeno.

É importante destacar aqui que a complexidade da construção das regras de inferência é diretamente proporcional a quantidade de variáveis linguísticas e conjuntos fuzzy definidos, principalmente no modelo de Mamdani que será explicado mais a frente, pois trabalha com produto cartesiano. No desenvolvimento do Modelo Fuzzy explicado no Capítulo 5, foi necessário limitar a quantidade de variáveis porque comprometia o tempo de execução do sistema.

4.6.2 Variáveis Linguísticas

Variáveis linguísticas são variáveis que descrevem de forma verbal uma informação que geralmente é quantitativa, ou seja, seus valores são palavras ou frases. Essas variáveis recebem classificações de pertinência de acordo com os conjuntos fuzzy que são determinados a partir das funções de pertinência (explicadas na seção 4.6.3) e são chamados de **termos linguísticos**. Por exemplo, na Tabela 1, a altura de uma pessoa pode ser classificada de três maneiras: baixa, média e alta, essas classificações são os conjuntos fuzzy e a altura é a variável linguística que vai ser estudada de acordo com o grau de pertinência em cada conjunto. É importante notar que uma variável não se associa apenas com um conjunto, ela é classificada de acordo com o nível de pertencimento para cada conjunto fuzzy, então, uma variável linguística pode ter relacionamento com dois ou mais conjuntos, porém, com graus de pertinência diferentes. Como é visto no exemplo da Tabela 1 o elemento 1.80 se associa com os conjuntos fuzzy médio e alto em graus diferentes. A Figura 6 ilustra o que são variáveis linguísticas e termos linguísticos e como usá-los na representação de funções de pertinência.

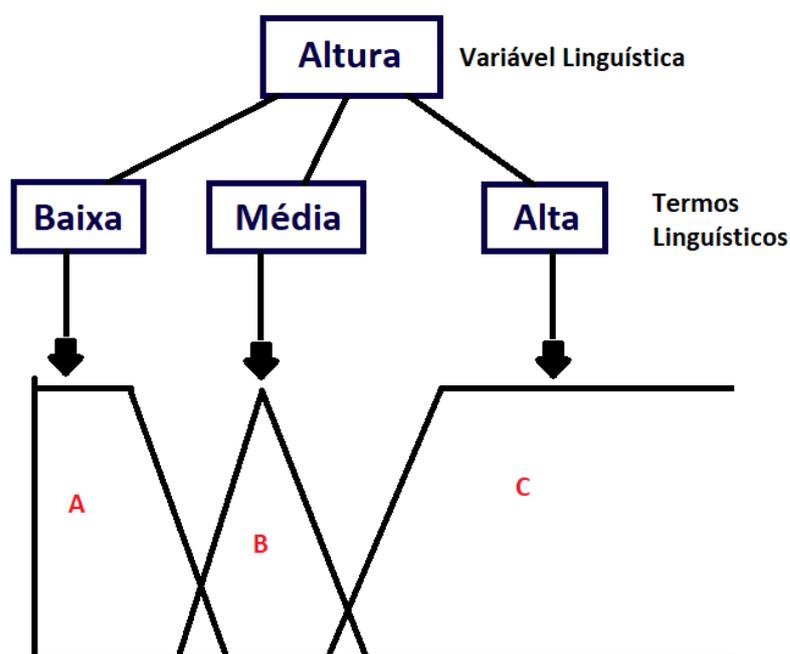


Figura 6 – Funcionamento de uma variável linguística e termos linguísticos.

4.6.3 Função de Pertinência

A função de pertinência é uma lógica numérica que atribui uma divisão fuzzy para valores discretos de uma variável. Existem diversos tipos de funções de pertinência e elas se adequam à realidades diferentes, por isso é importante a atuação de um especialista na área estudada para observar o comportamento do gráfico das funções. Tomando como exemplo a variável altura proposta na Tabela 1, essa variável linguística está associada aos termos linguísticos baixo, médio e alto, que correspondem aos conjuntos fuzzy A, B e C, onde cada um tem uma função de pertinência para a variável em questão. A representação gráfica desse cenário pode ser vista na Figura 7. A variável linguística altura é comparada com os 3 conjuntos fuzzy de acordo com sua pertinência. Para a classificação "baixo", é observado que o nível de pertencimento é 1 até a altura de 1.30m e após isso ele tende a cair à medida que a altura aumenta. Uma pessoa de 1.75m tem grau de pertencimento igual a zero para o conjunto baixo e começa a crescer no conjunto alto, dessa forma deve ser feita a leitura das funções. É importante destacar que foram utilizadas dois tipos de funções, trapezoidal e triangular, mas existem outras, como por exemplo a gaussiana e a função de gbell. Na prática, a decisão da função de pertinência deve respeitar o contexto e ser observada de perto pelo especialista da área para verificar se o comportamento das curvas representa uma situação real.

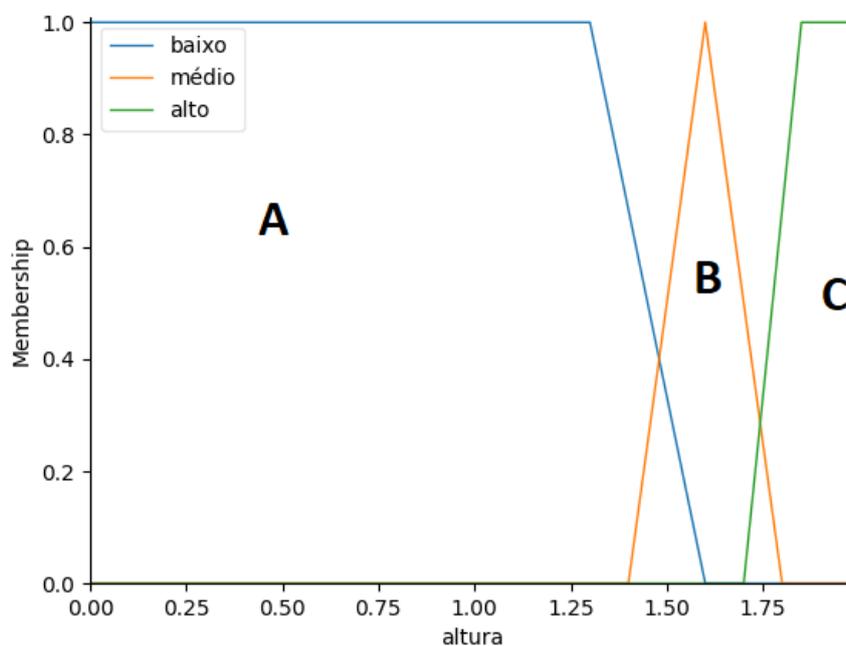


Figura 7 – Funções de pertinência trapezoidal(a e c) e triangular(b)

Na Figura 8 podem ser vistos alguns exemplos de funções de pertinência mais utilizadas:

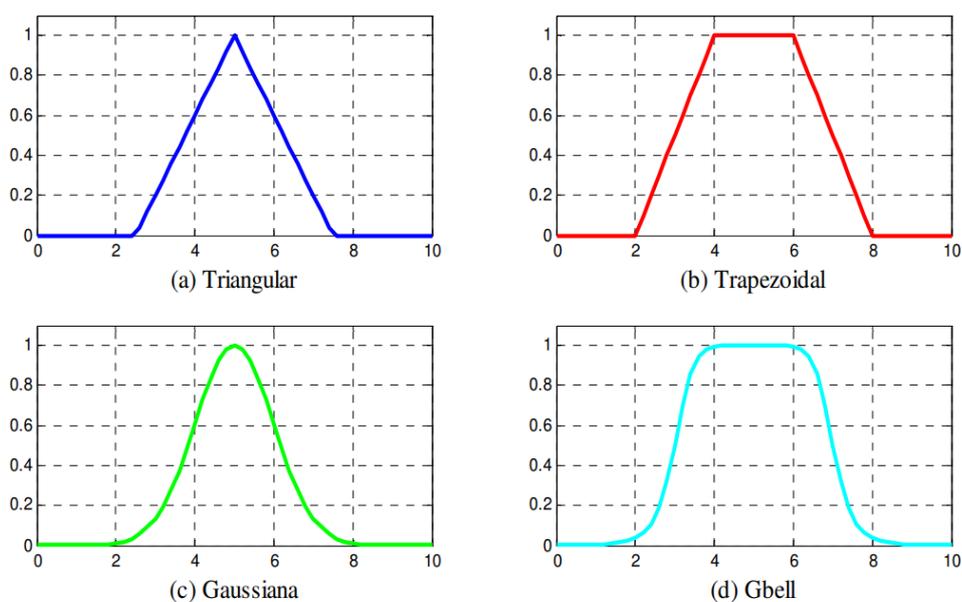


Figura 8 – Exemplo de funções de pertinência.
(SABINO, 2016)

Para construir essas funções é necessário compreender como elas funcionam

primeiramente, e após isso, criar intervalos que representam as áreas do gráfico. Por exemplo, na Figura 8 foi selecionado o intervalo [1.40, 1.80] para representar o conjunto fuzzy B(triangular) com pico na altura 1.6. Esses três valores funcionam para o tipo de figura em questão, mas para cada função deve ser observado como o intervalo é explicitado.

4.6.4 Sistemas Baseados em Regras Fuzzy (SBRF)

Nesta seção será explicado como ocorre o processo de modelagem de um sistema baseado em regras fuzzy utilizando o conteúdo abordado nos tópicos anteriores deste Capítulo e apresentados dois novos conceitos que fazem parte do início e fim de um modelo fuzzy. Abaixo, na Figura 9, pode ser visto o diagrama de fluxo geral de como funciona a construção do sistema:

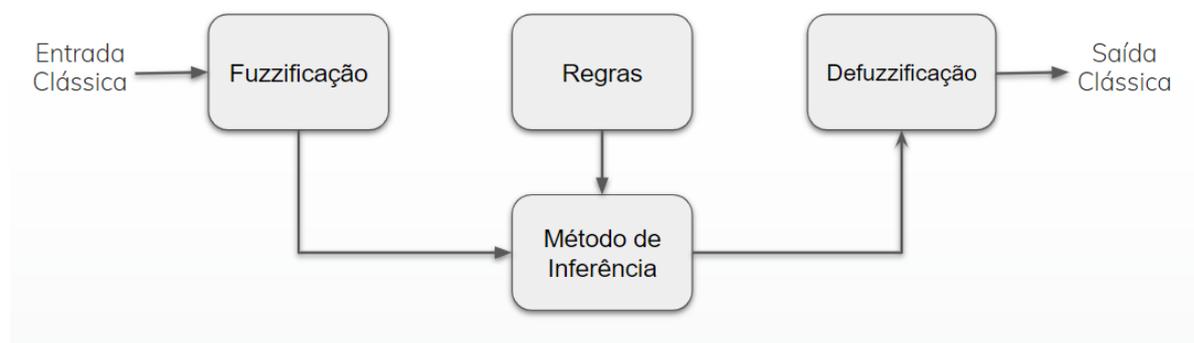


Figura 9 – Diagrama de um Modelo Fuzzy.

Como descrito na Figura 9, o processo de modelagem de um sistema fuzzy é composto basicamente de três etapas: Fuzzificação, métodos de inferência e defuzzificação. Na primeira etapa são consideradas entradas clássicas que são transformadas em conjuntos fuzzy; a segunda etapa é constituída pela base de regras que relacionam as variáveis de entrada e saída para construir as regras de inferência; por fim é necessário fazer a tradução dessa informação obtida a partir da variável consequente.

- Processo de Entrada (Fuzzificação): No componente inicial de um modelo fuzzy, são consideradas entradas clássicas que passam por um processo de chamado de fuzzificação, nesse momento é onde as entradas são mapeadas em conjuntos fuzzy para criação de variáveis linguísticas, termos linguísticos e funções de pertinência de acordo com seus respectivos domínios. No exemplo da Tabela 1 tem-se a variável linguística altura e três conjuntos fuzzy classificados da seguinte maneira: baixo (0 a 1.6m), médio (1.40m a 1.80m) e alto (1.70m a 2.00m). É importante que as funções de pertencimento se interceptem para manter a regra da incerteza do modelo e particularidade das respostas, por isso é essencial

a presença de um especialista que vai saber determinar onde esse cruzamento de funções vai ser nebuloso.

- Regras: Nessa etapa serão montadas as regras de inferência obedecendo a lógica do modelo fuzzy escolhida, como Mamdani ou Takagi-Sugeno, por exemplo, citados da seção 4.6.1. Para isso, é necessário observar as variáveis e termos linguísticos gerados na fase inicial para organizá-los em formato lógico.
- Processo Saída (Defuzzificação): Na modelagem de um sistema fuzzy a defuzzificação é o processo onde um resultado fuzzy é traduzido em uma saída clássica. Para isso, é efetuada a interpretação dos conjuntos fuzzy consequentes usando métodos de defuzzificação. Nesse momento é estudado de que forma as variáveis serão abordadas para obter as respostas desejadas. O método mais conhecido de defuzzificação é a centro de área (também chamado de centroide ou centro de gravidade) que retorna a área formada pela média de todas as figuras construídas após os cálculos de pertinência. O cálculo do resultado fuzzy leva em conta cada grau de pertencimento atribuído às variáveis linguísticas, não é apenas uma comparação de pertencimento de conjuntos.

4.7 Método de Inferência de Mamdani

Também conhecido como modelo de Mamdani, o método desenvolvido por Ebrahim Mamdani ([LAGHETTO, 2016](#)) é uma das técnicas mais conhecidas de obtenção de resultados fuzzy em um modelo. Na sua idealização, as regras de inferência são construídas a partir do produto cartesiano dos conjuntos fuzzy antecedentes e consequentes da regra. Neles são aplicados as regras de composição de inferência max-min que agrega as regras através do operador lógico OU (representado pela t-conorma ∇ , máximo) e as variáveis linguísticas de cada regras com o operador lógico E (representado pela t-norma Δ , mínimo). A resposta final é o resultado da agregação das saídas inferidas pelas regras de inferência. Abaixo estão listadas algumas regras de inferência de Mamdani:

- se (x é A) e (y é B) então z é C.
- se (peso é baixo) e (altura é mediana) então IMC é normal.
- se (diabetes é alta) e (obesidade é alta) então risco de síndrome metabólica é alto.

Na Figura 10, é ilustrado como funciona o processo de resolução de uma regra de inferência usando a técnica de Mamdani. Nas regras R1 e R2 é observado que as

duas entradas x e y formam as saídas C_1 e C_2 a partir dos seus graus de pertinência utilizando a técnica de mínimo, ou seja, prevalece o grau de pertinência menor.

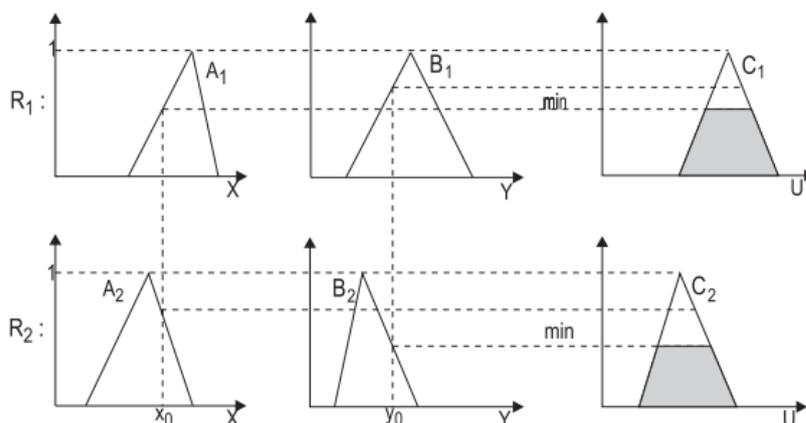


Figura 10 – Saídas de regras baseadas no método de Mamdani usando t-norma. (SILVA, 2011)

Já na Figura 11, é demonstrado como as saídas C_1 e C_2 , resultantes das regras de inferência do modelo são usadas com a t-conorma para produzir o grau de pertinência final. É observado que o conjunto fuzzy final é construído a partir da união dos resultados anteriores.

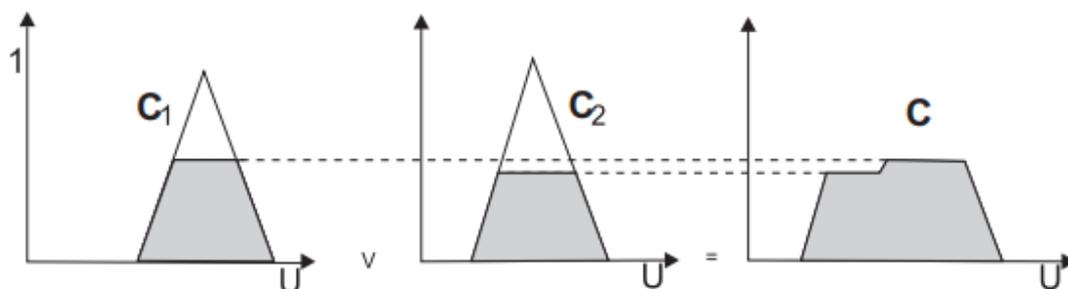


Figura 11 – Saída final fuzzy no método de Mamdani usando t-conorma. (SILVA, 2011)

5 Um Modelo Baseado em Lógica Fuzzy para Estimar o Risco de Desenvolver Câncer de Colo de Útero

Este capítulo trata do desenvolvimento do trabalho em si. Inicialmente, será apresentada a descrição do aplicativo móvel para coleta de dados e em seguida o modelo fuzzy proposto para estimar o risco de desenvolvimento da doença. Por fim, são descritos alguns resultados preliminares obtidos ao aplicar o modelo à dados hipotéticos.

5.1 O Aplicativo

O aplicativo foi desenvolvido com a mentoria de Iolanda Matias, médica ginecologista pesquisadora do LIKA-UFPE, que atuou como especialista no assunto. O objetivo é coletar dados sobre os pacientes enquanto eles respondem a um questionário contendo os fatores de risco e proteção associados ao câncer de colo de útero, descritos no Capítulo 2. São eles: idade, idade da primeira relação sexual, número de gestações, número de abortos, número de partos, qualidade da alimentação, localização, presença de infecções sexualmente transmissíveis, tabagismo, histórico de câncer nos pais, diabetes, hipertensão, uso de corticoides, número de parceiros, uso de preservativo, resultado do papaniculau, resultado da colposcopia, resultado do HPV teste e vacinação contra HPV. Com a coleta desses dados o propósito futuro é integrar a aplicação com o Modelo fuzzy para calcular o risco de desenvolvimento de câncer cervical usando as perguntas como fonte de dados.

O aplicativo funciona da seguinte maneira: o usuário é apresentado a uma tela inicial onde é explicado o objetivo do sistema, em seguida, ele responde uma série de perguntas que foram distribuídas de acordo com as categorias histórico do paciente, epidemiologia da doença, fatores genéticos e/ou patologias e variáveis classificadas como atitudes de prevenção e/ou cuidado relacionadas a HPV e ao câncer. Essas informações são salvas em variáveis e ao final do processo enviadas para um banco de dados onde ficam estruturadas em formato de árvore, na qual o exame corresponde à cabeça da árvore, os filhos são as variáveis e os netos são os resultados referentes a cada variável, seguindo esquema representado na Figura 12.

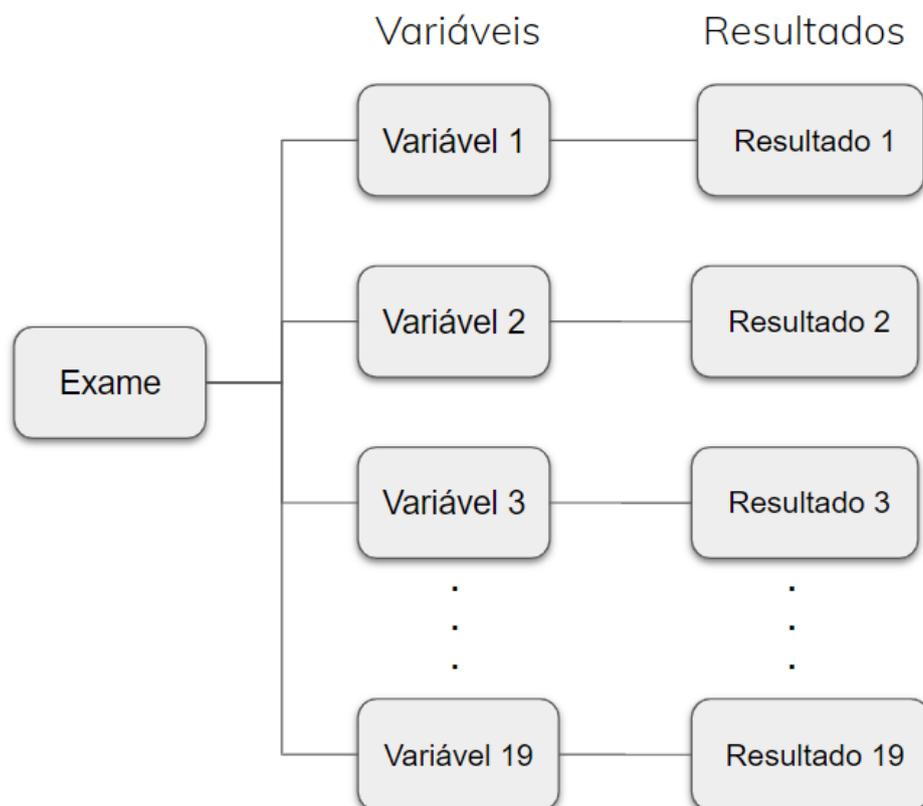


Figura 12 – Esquema do banco de dados do aplicativo.

Em um cenário futuro, com possível integração, ao final do questionário, o paciente é informado em que classificação se encaixa de acordo com a chance de desenvolver câncer cervical, são elas: muito baixo, baixo, moderado, alto e muito alto. Além disso, são observadas quais foram as informações dadas e comparadas com respostas ideais. Caso alguma resposta esteja fora do recomendado como boas práticas de prevenção e cuidados, o usuário recebe indicações de medidas que podem ser tomadas para preservar sua saúde. Por exemplo, para um paciente que responde que nunca tomou ou está com a vacina contra HPV vencida, é indicado que procure um médico para ter acesso à medicação.

O objetivo é que o software possa ser usado por médicos para auxiliar em diagnósticos de câncer de colo. Além de ser útil a pacientes que desejem ter consciência do seu risco de exposição e procurar um profissional munido de informação e recomendações para seu estado de saúde. Na Figura 13 podem ser vistas algumas telas da aplicação.

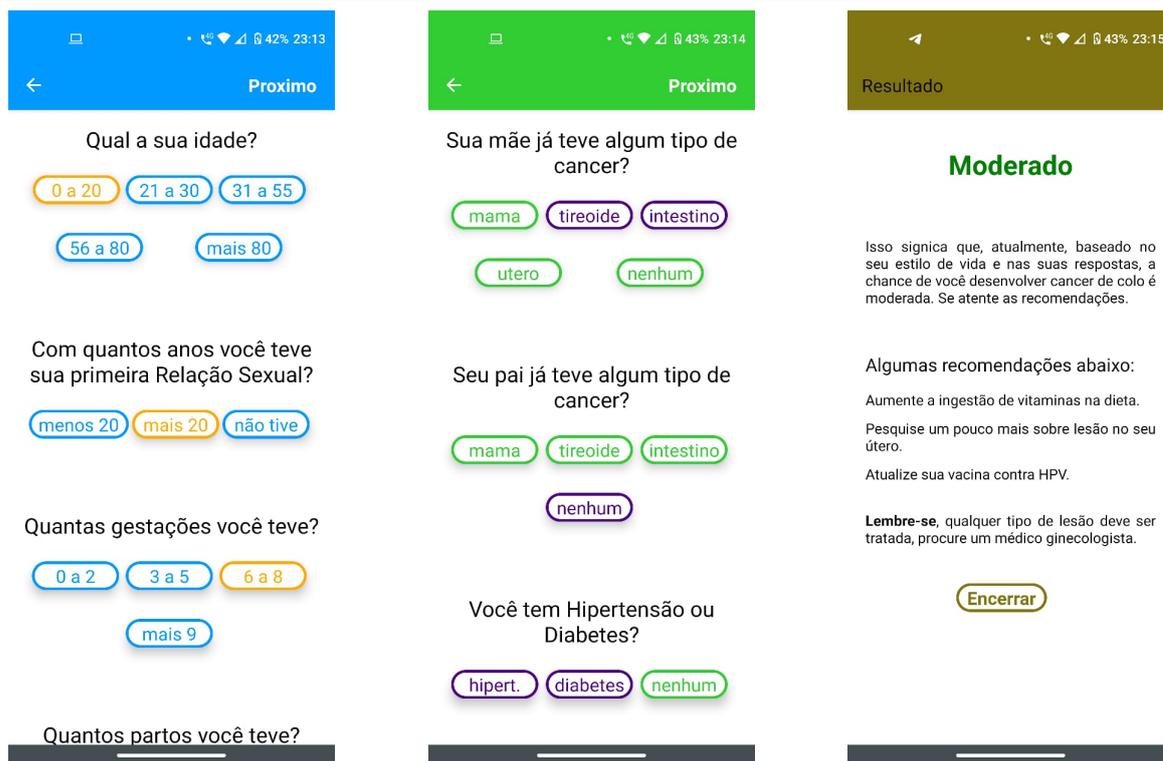


Figura 13 – Telas do aplicativo: As duas primeiras são telas de questionário, a última representa a tela de resultado como ficaria após o cálculo de risco.

5.1.1 Arquitetura do App

O aplicativo foi desenvolvido em React Native (React Native, 2021), utilizando a linguagem de programação Javascript para o backend e React Native view para construção das telas. As informações coletadas ficam armazenadas no banco de dados Firebase, construído de forma não relacional. Na Figura 14 pode ser visto como ficou o diagrama de classes do sistema.

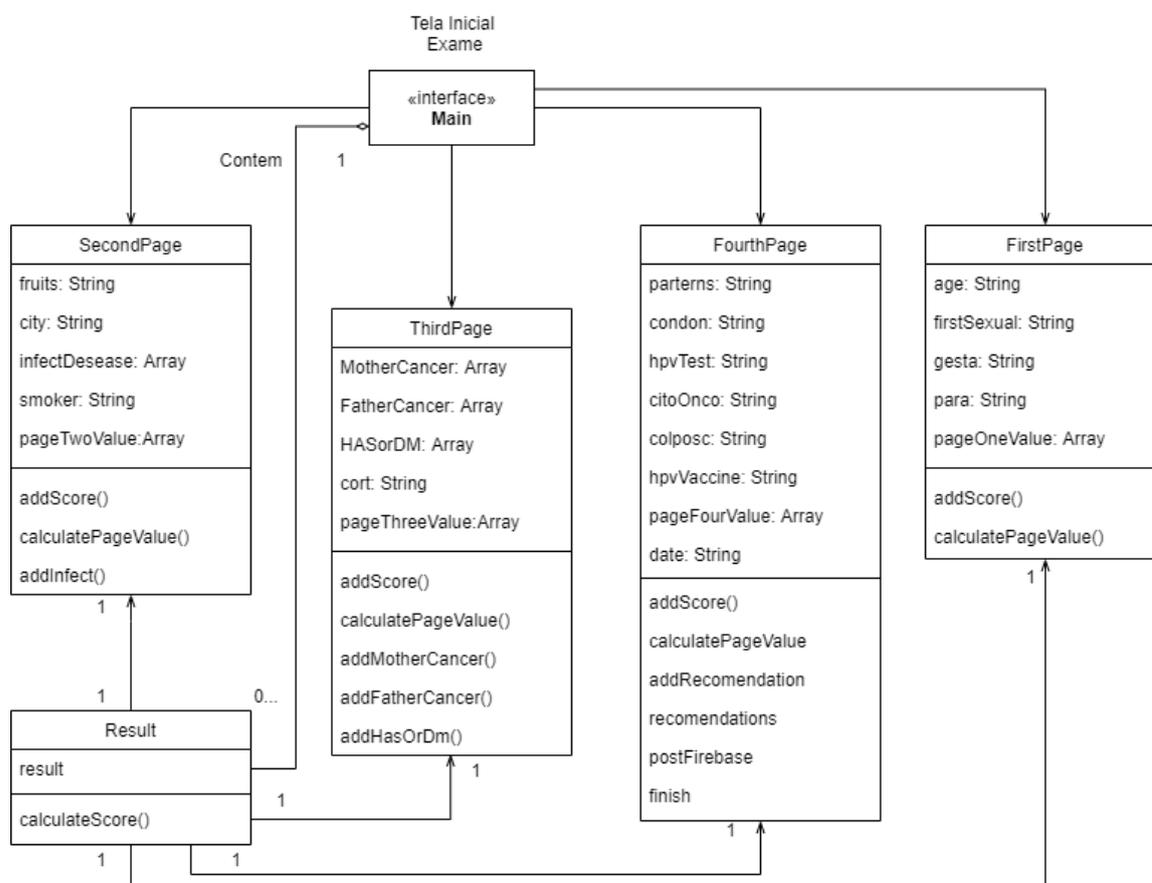


Figura 14 – Diagrama de Classes do App.

A aplicação foi desenvolvida de forma que cada tela corresponde a uma classe e nela são explicitadas as variáveis e métodos utilizados. No diagrama, em cada classe, as variáveis correspondem ao primeiro quadrante e os métodos ao segundo. A primeira tela é o Main, tela principal do aplicativo onde se inicia o exame e está relacionada com cada classe, que equivalem às telas de questionários. A interface final é composta pela classe Result, que é criada após receber os dados de cada entidade de cada tela anterior.

5.2 Modelo Fuzzy

Nessa seção será apresentada a construção de um modelo baseado em lógica fuzzy para estimar o risco de uma pessoa desenvolver câncer de colo uterino. Para isso, foi utilizada a teoria dos conjuntos fuzzy apresentada no Capítulo 4.

5.2.1 O Modelo

O modelo foi desenvolvido em python usando a biblioteca SKfuzzy ¹ que organiza o ambiente de forma a trabalhar diretamente com lógica baseada em sistemas fuzzy. Essa biblioteca é ideal para a criação de aplicações fuzzy porque traz métodos e funções nativas para criar as variáveis linguísticas, definir os termos linguísticos, as funções de pertinência e método de defuzzificação.

Para modelar o risco do câncer foi necessário selecionar um subconjunto das 19 variáveis citadas no Capítulo 2, devido ao tempo de execução do sistema, que aumenta à medida que a quantidade de variáveis linguísticas cresce, como descrito no Capítulo 4. Portanto, foi preciso escolher um subconjunto de variáveis para modelagem. Foram elas: idade, idade da primeira relação sexual, histórico familiar de câncer, presenças de outras infecções sexualmente transmissíveis, tabagismo, quantidade de parceiros e vacinação contra HPV. O método de inferência escolhido foi o de Mamdani, onde foram criadas 648 regras de inferência abordando todas as variáveis linguísticas e os conjuntos fuzzy. Para extrair uma resposta numérica clássica a partir da solução de Mamdani foi selecionada a função centroide. Na Figura 15 pode ser visto como ficou o resumo do modelo contendo as sete variáveis linguísticas antecedentes e a consequente.

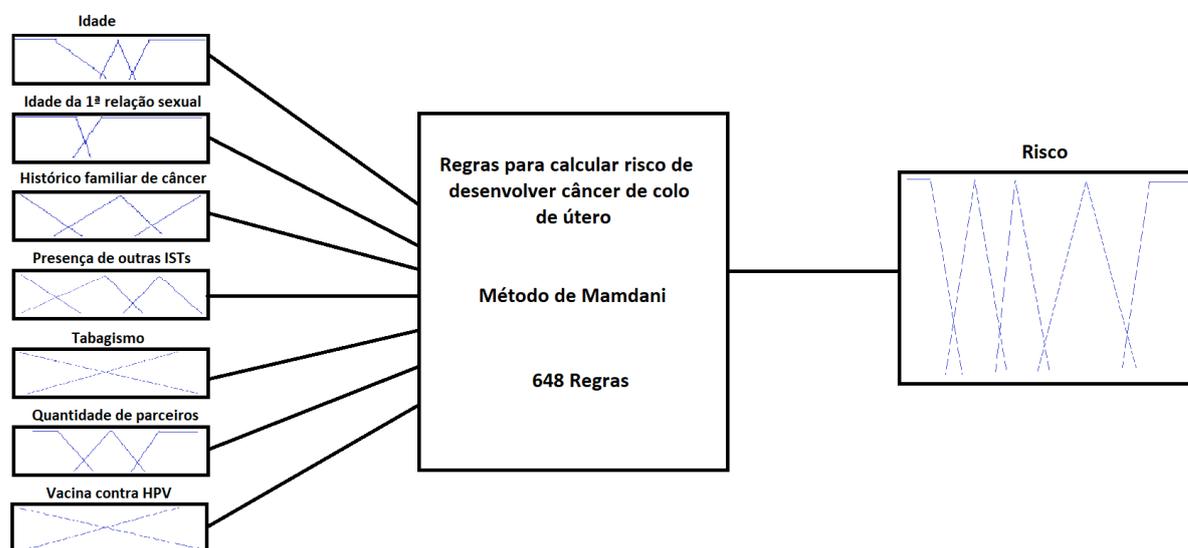


Figura 15 – Resumo do modelo fuzzy para cálculo de risco de câncer de colo de útero.

5.2.2 Variáveis Antecedentes e Consequente

As variáveis linguísticas abordadas foram definidas a partir do conjunto de todos os fatores que interferem no surgimento do câncer de colo, sugeridos por Iolanda Ma-

¹ <https://pythonhosted.org/scikit-fuzzy/>

tias. Para isso foi necessário investigar quais delas têm mais influência na doença. Essa investigação foi feita estudando trabalhos relacionados ao surgimento de câncer de colo e quais as causas associadas. As variáveis antecedentes foram desenvolvidas e gráficos de pertinência foram criados a partir da mesma biblioteca SKFuzzy no compilador Pycharm:

- Idade: A variável idade foi relacionada com três subconjuntos fuzzy, jovem (0 a 35 anos), moderado (30 a 55 anos) e avançado (50 a 80 anos). Para tanto, foi gerado um gráfico de pertencimento da variável usando uma função triangular e duas trapezoidais. A Figura 16 representa este gráfico, observe que para representar os conjuntos fuzzy jovem e avançado foram usadas funções trapezoidais. Note que até 12 anos uma pessoa é considerada jovem e após isso ela começa a perder pertinência para o conjunto. O oposto acontece com o conjunto avançado, onde uma pessoa a partir dos 57 anos começa a ter idade mais avançada:

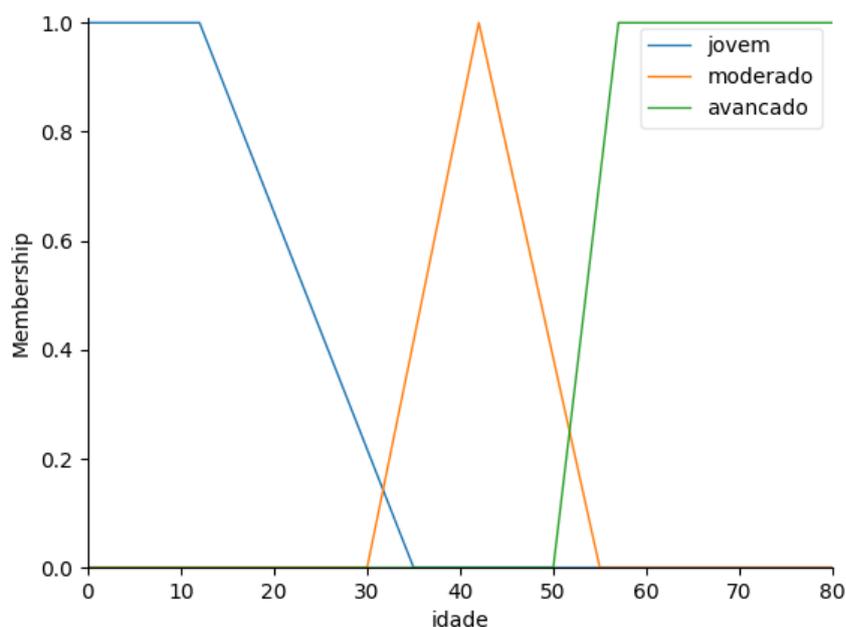


Figura 16 – Representação gráfica da variável Idade.

- Idade da primeira relação sexual: Aqui foi observado se o paciente teve a primeira relação sexual muito jovem. Para isso foram criados dois conjuntos fuzzy: criança/adolesc (0 a 20 anos) e jovem/adult (15 a 50 anos). O gráfico desta variável é apresentado na Figura 17. Os dois conjuntos foram representados por funções trapezoidais de pertinência: Até 15 anos o grau de pertinência ao conjunto criança/adolesc é igual a 1. A partir daí começa a decair até zerar ao atingir

20 anos. O mesmo vale para a extremidade direita do gráfico, onde o grau de pertencimento 1 de jovem/adulto começa aos 20 anos e permanece assim durante o decorrer da função.

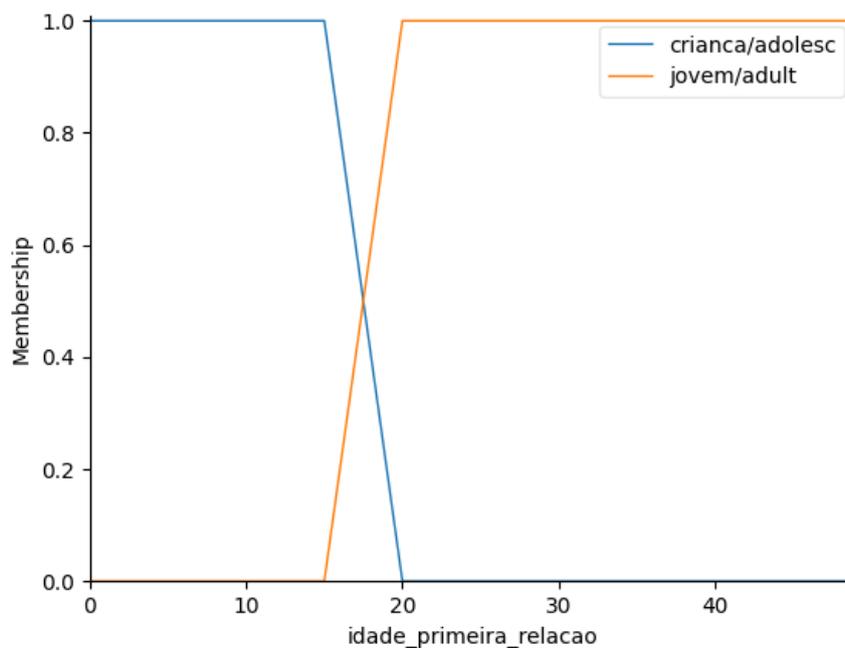


Figura 17 – Representação gráfica da variável Idade da primeira relação sexual.

- Histórico familiar de câncer: Para simplificar o modelo, vamos restringir esta variável à quantidade de cânceres presentes nos pais do paciente. Os tipos são os mesmos observados no Capítulo 2: câncer de Mama, Útero, Tireóide ou Intestino. Os conjuntos fuzzy criados foram desenvolvidos pensando na classificação do indivíduo em relação ao histórico, podendo ser fraco (0 a 2 cânceres), médio (1 a 5 cânceres) ou forte (4 a 7 cânceres). A Figura 18 apresenta os intervalos de pertencimento para os três conjuntos. As funções selecionadas foram todas triangulares pois representam o decaimento em linha reta da taxa de pertencimento. Na seleção dessa variável é importante notar que o número máximo de cânceres no histórico dos pais é sete, sendo mama, tireóide, intestino e útero (esse último apenas pode estar presente na mãe), por isso, a função de pertencimento para conjunto forte tem grau 1 ao atingir 7 cânceres.

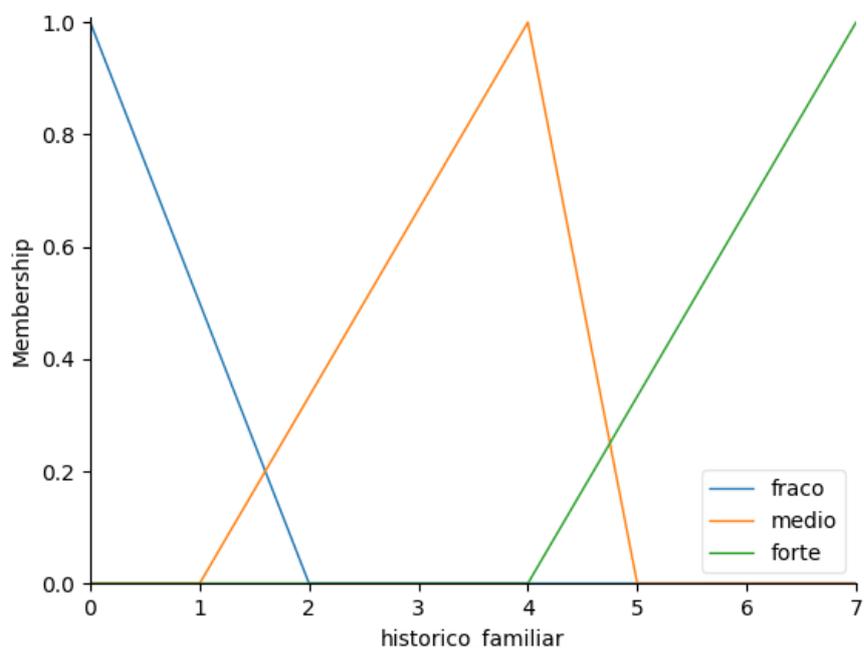


Figura 18 – Representação gráfica da variável Históricio Familiar.

- Presença de outras ISTs: Essa variável segue a mesma ideia da anterior, o paciente é avaliado quanto a quantidade de doenças sexualmente transmissíveis que porta e classificado em três conjuntos: fraco (0 a 1 IST), médio (0 a 3 ISTs) ou forte (2 a 4 ISTs). É importante destacar que foram observadas as quatro infecções citadas no capítulo 2. Na Figura 19 pode ser observado que os três conjuntos de pertencimento foram definidos usando funções triangulares, da mesma forma que a variável anterior, utilizando a mesma justificativa da quantidade máxima de quatro infecções sexualmente transmissíveis.

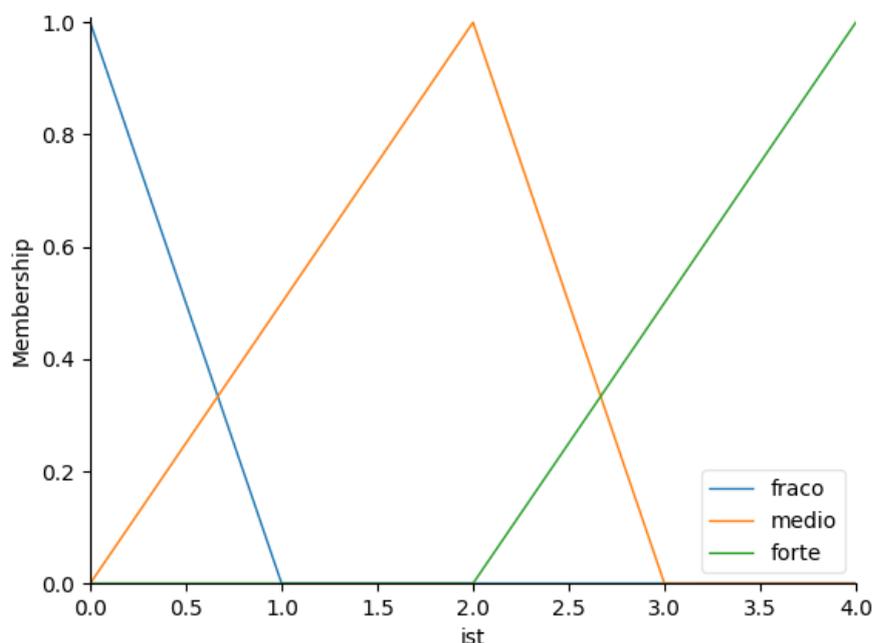


Figura 19 – Representação gráfica da variável Presença de ISTs.

- Tabagismo: Essa variável foi representada pela palavra "fumante" no modelo. Diferente das outras que possuem pertinência em conjuntos fuzzy ela é uma variável clássica tendo valor {sim, não}, para atribuir valor a ela foi usada a regra booleana do 0 (sim) ou 1 (não). A representação gráfica é encontrada na Figura 20, onde é possível observar que as duas funções se interceptam tendo valor de pertinência unicamente 0 ou 1 pois não existem respostas intermediárias entre verdadeiro e falso.

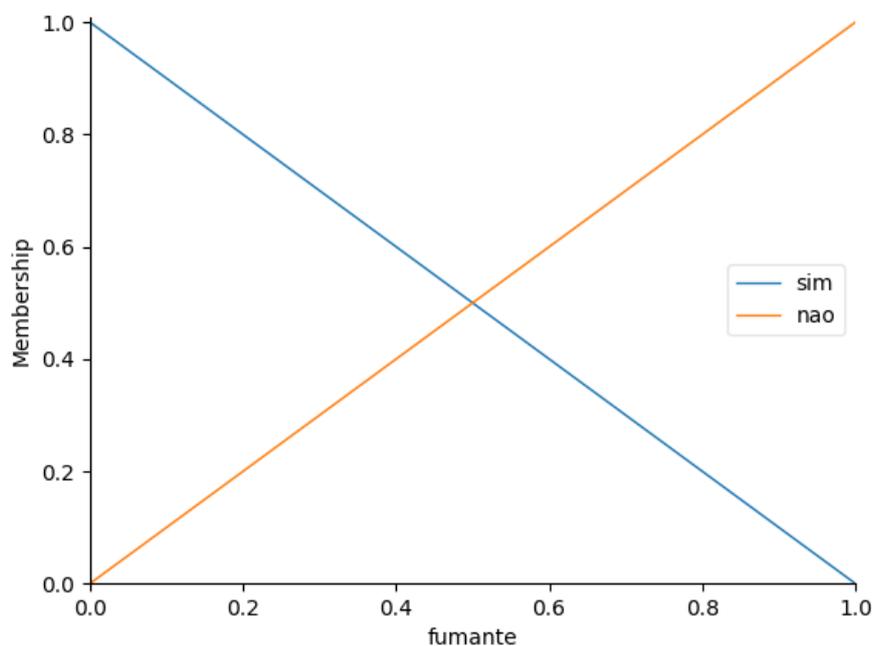


Figura 20 – Representação gráfica da variável Tabagismo.

- Quantidade de Parceiros: É observado aqui a quantidade de parceiros do sexo masculino que o paciente teve no último ano, classificando o resultado em três conjuntos fuzzy: pouco (0 a 7 parceiros), moderado (5 a 18 parceiros) e muito (15 a 25 parceiros). Na figura 21 é apresentado o gráfico da variável linguística em questão. Foram usadas duas funções trapezoidais e uma triangular nos conjuntos, isso porque o início do gráfico é mantido com grau de pertinência 1 até a quantidade de 2 parceiros e após isso há queda no grau, assim como na função do extremo direito, a partir de 20 parceiros por ano é considerado grau de de pertencimento 1. Já na função triangular, foi observado um intervalo grande onde pode ser considerado um número moderado de parceiros até começar o decaimento.

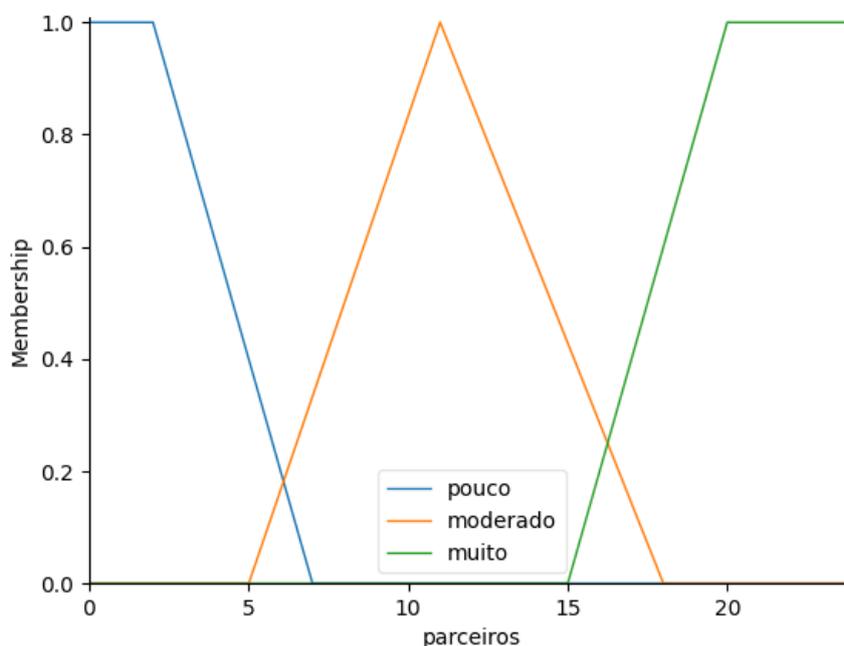


Figura 21 – Representação gráfica da variável Quantidade de Parceiros.

- Vacina contra HPV: Por último foi criada a variável clássica para saber se o paciente tomou ou não a vacina contra HPV. Diferente da aplicação nós não precisamos saber o tempo da vacinação e classificamos em dois conjuntos crisp sim, nao, seguindo o mesmo padrão da variável tabagismo, 0 (sim) e 1 (não). A representação da variável pode ser vista na Figura 22.

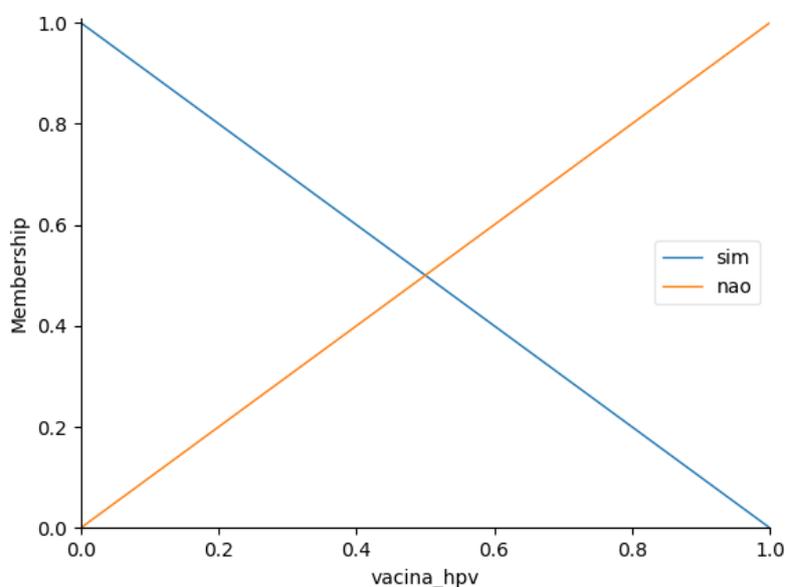


Figura 22 – Representação gráfica da variável Vacina Contra HPV.

Para obter o valor final do cálculo de risco foi desenvolvida uma variável resposta consequente, chamada de Chance. Essa variável linguística se relaciona com cinco conjuntos de pertinência: muito baixa (0 a 12), baixa (9 a 20), moderada (17 a 29), alta (26 a 52) e muito alta (49 a 55). Nela foram usados dois tipos de funções de pertinência, duas trapezoidais, representando os dois extremos e três triangulares. A Figura 23 representa o gráfico dessa variável linguística.

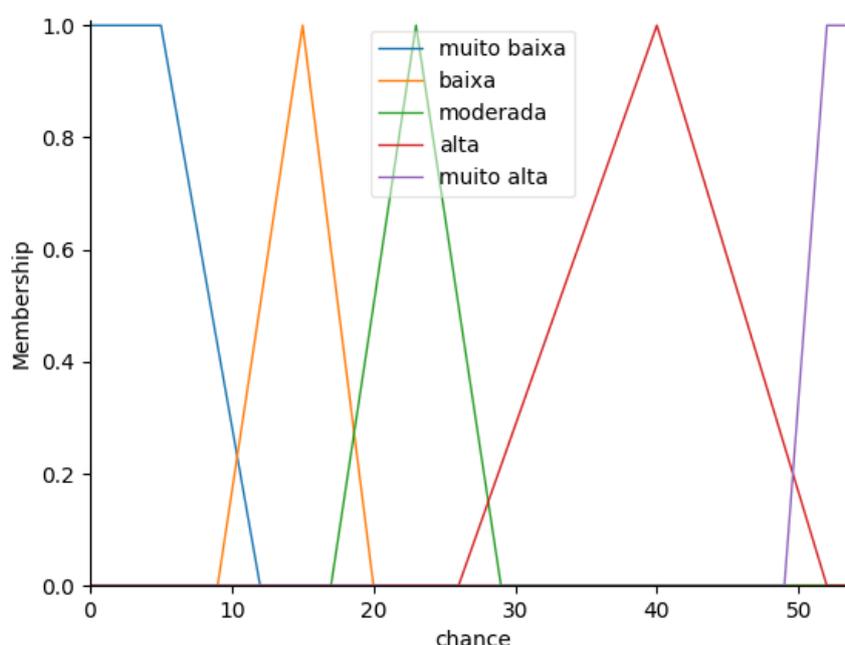


Figura 23 – Representação gráfica da variável Chance

A Tabela 2 resume as variáveis linguísticas e seus respectivos termos linguísticos. Nela é possível ver os parâmetros antecedentes e o consequente que foram traduzidos a partir da legenda: I: idade; IRS: idade da primeira relação sexual, H: histórico familiar; ISTs: presença de infecções sexualmente transmissíveis, T: tabagismo; QP: quantidade de parceiros sexuais no último ano; V: vacina contra HPV.

I	IRS	H	ISTs	T	QP	V	Chance
jovem	criança/adolesc	fraco	fraco	sim	pouco	sim	muito baixa
moderado	jovem/adult	medio	medio	não	moderado	não	baixa
avanzado		forte	forte		muito		moderada
							alta
							muito alta

Tabela 2 – Resumo das variáveis linguísticas e termos linguísticos.

5.3 Regras de Inferência

Seguindo o modelo de Mamdani, as regras de inferência foram montadas a partir do produto cartesiano das variáveis linguísticas e seus conjuntos. Com base nisso, foram geradas 648 regras únicas utilizando um script em python de combinação aleatória que gerou um arquivo contendo todos os arranjos possíveis. O cálculo foi feito multiplicando entre si as quantidades de termos linguísticos de cada variável: $3 \times 2 \times 3 \times 3 \times 2 \times 3 \times 2 = 648$. Alguns exemplos de regras montadas:

- Se (idade é avançado) e (idade da primeira relação é jovem/adult) e (presença de ISTs é forte) e (tabagismo é não) e (histórico familiar é forte) e (parceiros é moderado) e (vacina contra hpv é não), então chance é alta.
- Se (idade é jovem) e (idade da primeira relação é jovem/adult) e (presença de ISTs é fraco) e (tabagismo é sim) e (histórico familiar é medio) e (parceiros é muito) e (vacina contra hpv é sim), então chance é moderada.
- Se (idade é jovem) e (idade da primeira relação é criança/adolesc) e (presença de ISTs é medio) e (tabagismo é não) e (histórico familiar é fraco) e (parceiros é pouco) e (vacina contra hpv é não), então chance é baixa.
- Se (idade é moderado) e (idade da primeira relação é criança/adolesc) e (presença de ISTs é forte) e (tabagismo é sim) e (histórico familiar é medio) e (parceiros é pouco) e (vacina contra hpv é não), então chance é alta.

5.4 Resultados Preliminares

Para testar o modelo foram criados cinco casos de pacientes hipotéticos A, B, C, D e E e atribuídos valores às variáveis linguísticas com objetivo de avaliar o método de inferência. Para o paciente A foram gerados os gráficos de todas as variáveis antecedentes e consequente para demonstrar o comportamento das funções e grau de pertinência na prática. Para os pacientes B, C, D e E foram colocados os gráficos da variável consequente chance, junto com os dados referentes as variáveis antecedentes.

Para o Paciente A foram considerados os seguintes valores:

- Idade: 28 anos
- Idade da primeira relação sexual: 12 anos
- Histórico familiar: 2 tipos de câncer
- Presença de ISTs: 1 infecção

- Tabagismo: 1 (indicando que o paciente NÃO é fumante)
- Quantidade de parceiros: 2
- Vacina contra HPV: 0 (indicando que o paciente TEM a vacina)

Na Figura 24 pode ser visto o gráfico da variável linguística consequente Chance para o paciente A, que contém o resultado do cálculo final indicando que o paciente se enquadra em risco moderado, com valor numérico de aproximadamente 23 na escala de classificação. A região pintada de verde representa a figura formada pela união de todas as figuras formadas a partir das variáveis antecedentes do exemplo, apresentadas na Figura 25.

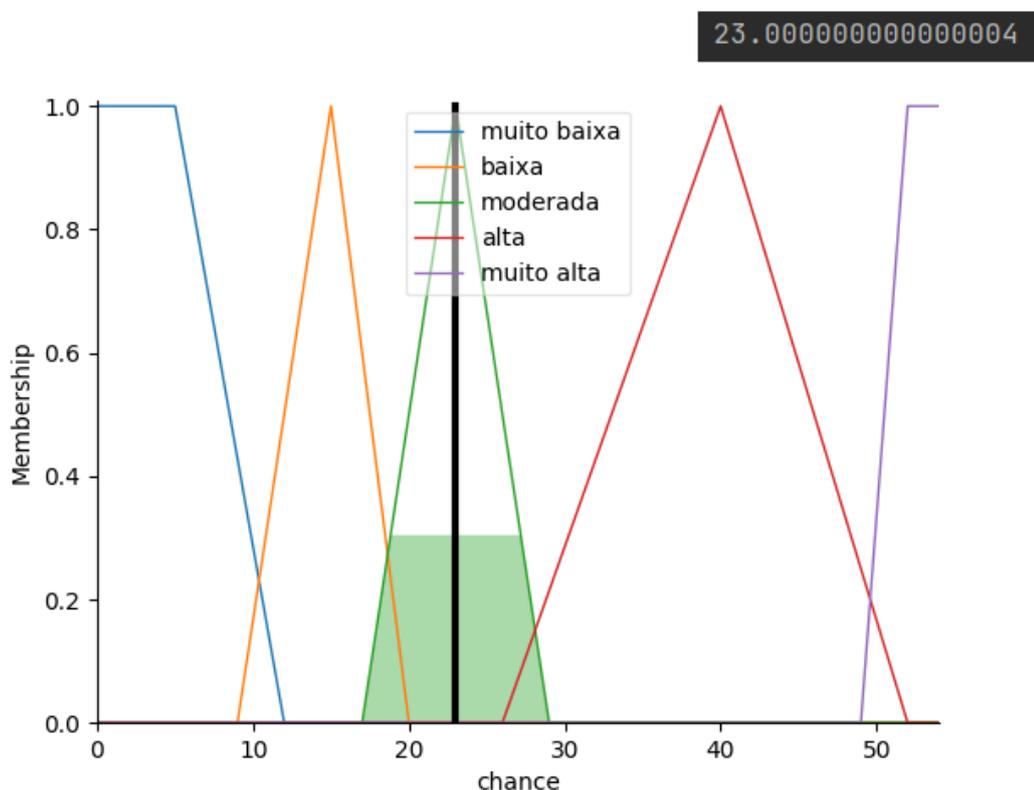


Figura 24 – Resultado da variável chance para o paciente A.

A Figura 25 representa os resultados individuais das variáveis antecedentes para o paciente A. As regiões pintadas representam o grau de pertinência da função de pertinência para cada variável calculado a partir da t-norma.

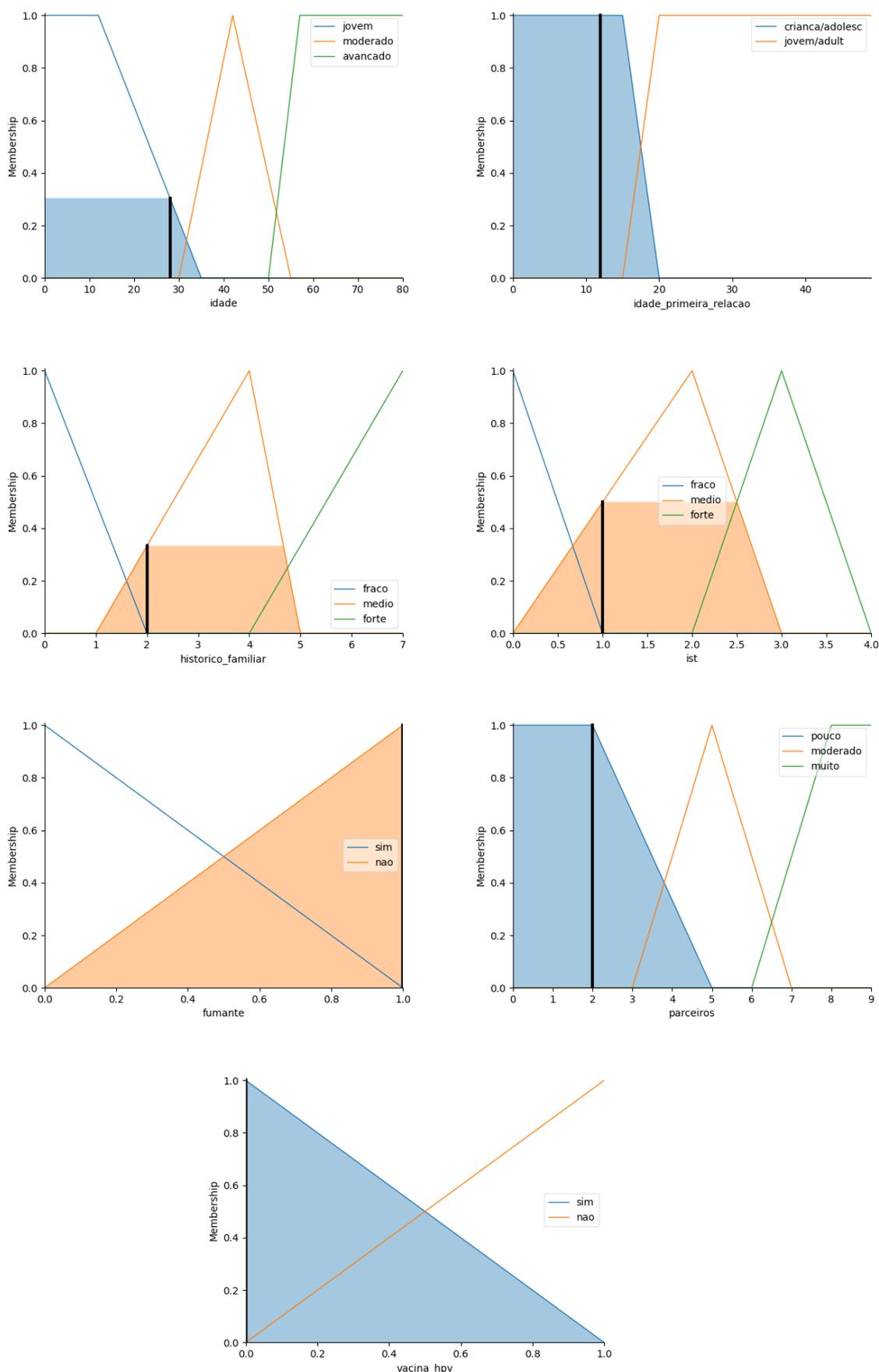


Figura 25 – Resultados parciais das variáveis idade, idade da primeira relação sexual, histórico de câncer, presença de outras ISTs, tabagismo, quantidade de parceiros e vacina contra HPV do paciente A.

Para o Paciente B foram considerados os seguintes valores:

- Idade: 43 anos
- Idade da primeira relação sexual: 19 anos
- Histórico familiar: nenhum tipo de câncer
- Presença de ISTs: nenhuma infecção
- Tabagismo: 1 (indicando que o paciente NÃO é fumante)
- Quantidade de parceiros: 1
- Vacina contra HPV: 0 (indicando que o paciente TEM a vacina)

A Figura 26 representa o resultado final de cálculo de chance para o paciente B. É possível observar que sua classificação é no grupo muito baixa, com risco de aproximadamente 4.74.

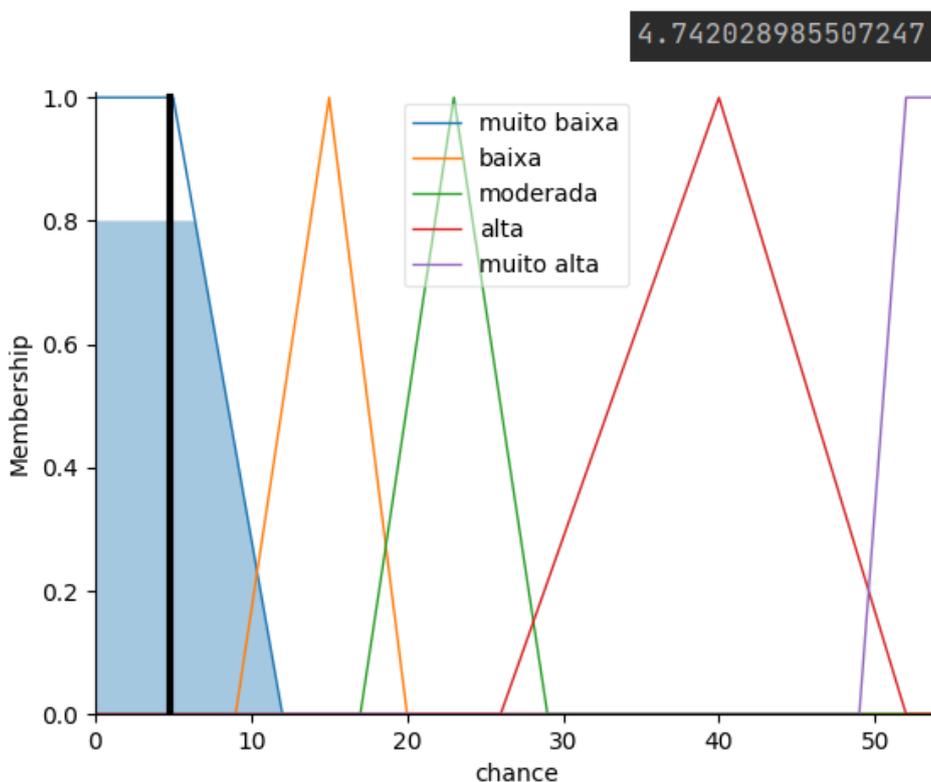


Figura 26 – Resultado da variável chance para o paciente B.

Para o Paciente C foram considerados os seguintes valores:

- Idade: 19 anos

- Idade da primeira relação sexual: 15 anos
- Histórico familiar: 5 tipos de câncer
- Presença de ISTs: 4 infecções
- Tabagismo: 0 (indicando que o paciente É fumante)
- Quantidade de parceiros: 10
- Vacina contra HPV: 1 (indicando que o paciente NÃO tem a vacina)

Na Figura 27 é possível ver o gráfico da variável final chance do paciente C calculada usando a t-conorma.

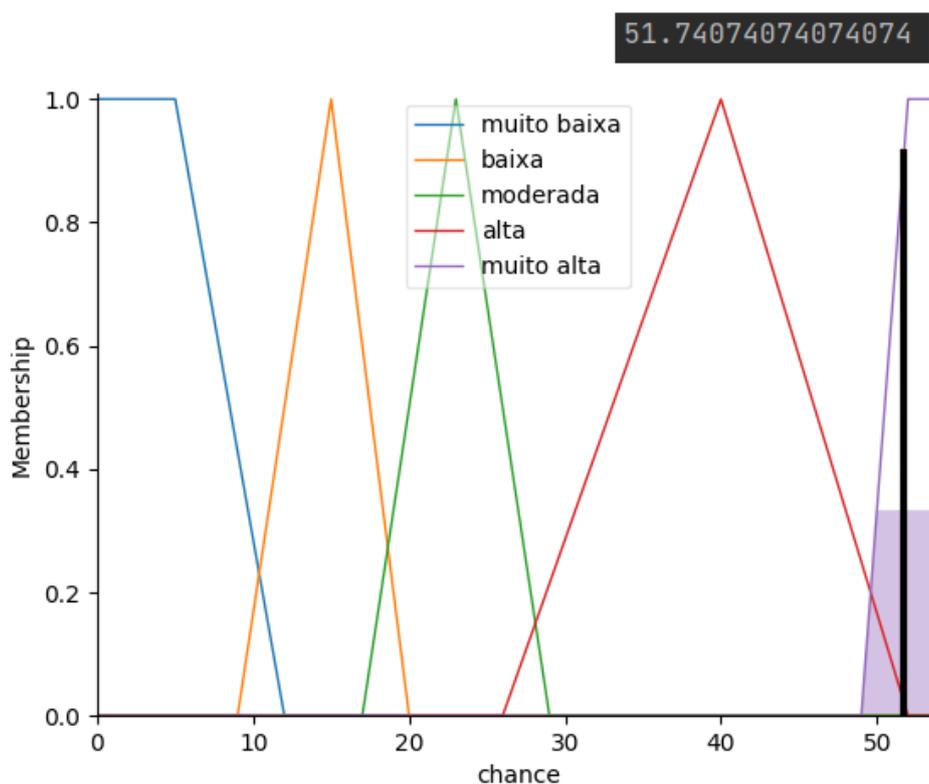


Figura 27 – Resultado da variável chance para o paciente C.

Para o Paciente D foram considerados os seguintes valores:

- Idade: 33 anos
- Idade da primeira relação sexual: 22 anos
- Histórico familiar: 4 tipos de câncer
- Presença de ISTs: 2 infecções

- Tabagismo: 0 (indicando que o paciente É fumante)
- Quantidade de parceiros: 12
- Vacina contra HPV: 1 (indicando que o paciente NÃO tem a vacina)

A Figura 28 representa o resultado final para o paciente D.

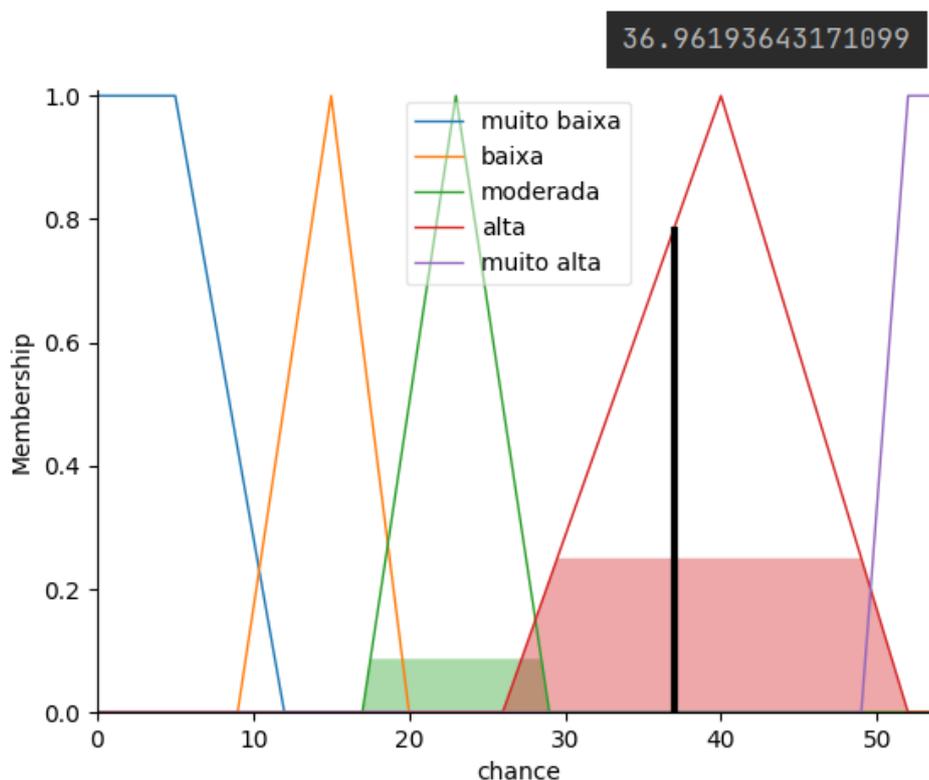


Figura 28 – Resultado da variável chance para o paciente D.

Para o Paciente E foram considerados os seguintes valores:

- Idade: 65 anos
- Idade da primeira relação sexual: 14 anos
- Histórico familiar: 6 tipos de câncer
- Presença de ISTs: 3 infecções
- Tabagismo: 1 (indicando que o paciente NÃO é fumante)
- Quantidade de parceiros: 17
- Vacina contra HPV: 1 (indicando que o paciente NÃO tem a vacina)

Na Figura 29 é possível observar que o paciente E se enquadra no risco alto, com chance de aproximadamente 43 na escala.

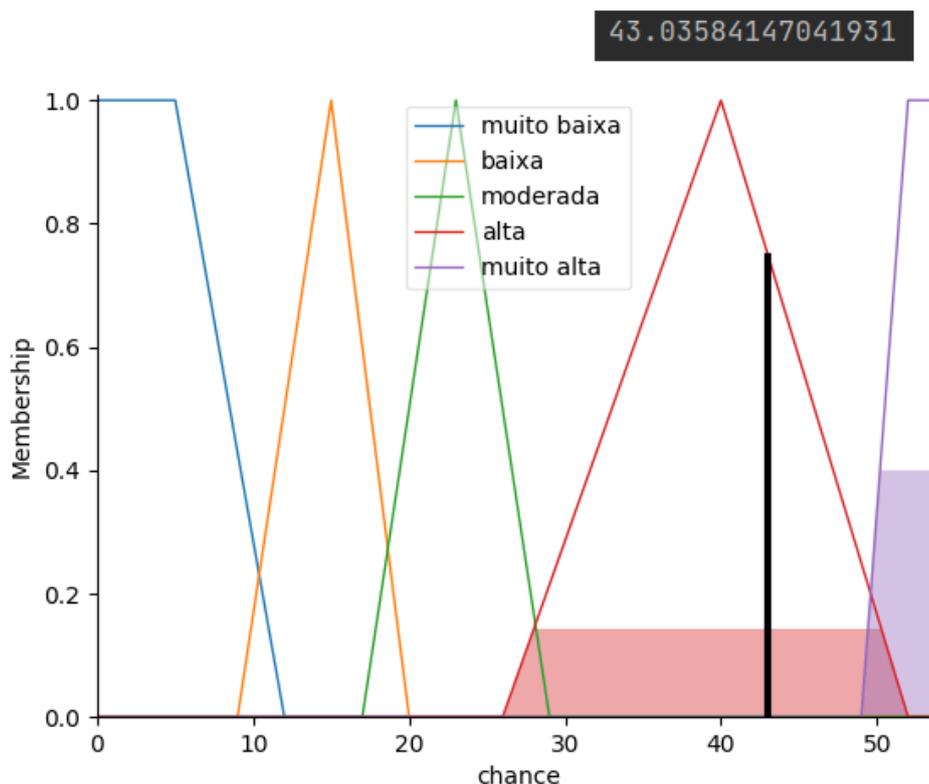


Figura 29 – Resultado da variável chance para o paciente E.

Pode ser visto em algumas imagens o cruzamento das funções de pertinência, por exemplo da Figura 33, a variável quantidade de parceiros assume um valor intermediário entre os conjuntos fuzzy moderado e muito. Essa representação explicita a teoria fuzzy de pertencimento em dois conjuntos, diferente da teoria clássica exclusiva.

A Tabela 3 representa o resumo dos resultados obtidos a partir das simulações dos pacientes A, B, C, D e E seguindo a legenda: I: idade; IRS: idade da primeira relação sexual, H: histórico familiar; ISTs: presença de infecções sexualmente transmissíveis, T: tabagismo; QP: quantidade de parceiros sexuais no último ano; V: vacina contra HPV.

Paciente	I	IRS	H	ISTs	T	QP	V	Chance	Categoria
A	28	12	2	1	não	2	sim	≈ 23	moderada
B	43	19	0	0	não	1	sim	≈ 4.75	muito baixa
C	19	15	5	4	sim	10	não	≈ 51.74	muito alta
D	33	22	4	2	sim	12	não	≈ 36.96	alta
E	65	14	6	3	não	17	não	≈ 43.04	alta

Tabela 3 – Resumo dos resultados dos casos de teste.

6 Considerações Finais

Com o uso da lógica fuzzy foi possível construir um modelo para estimar o risco do surgimento de câncer de colo de útero. Este é um modelo inicial que ainda precisa ser aprimorado e validado com o auxílio de especialistas da área. Para simplificar o modelo, foi selecionado um subconjunto com 07 variáveis, das 19 que foram inicialmente indicadas pela Dr. Iolanda, especialista na área. O modelo simplificado contém 648 regras e exige ao menos 07 minutos de tempo de processamento em um computador Windows i7 com 8GB de Ram e 1 terabyte de hard drive. Um modelo com todas as 19 variáveis gera um total de 6.298.560 regras de inferência, o que levaria um tempo 100000 vezes maior para rodar. Do ponto de vista computacional é possível evoluir o modelo matemático acrescentando mais variáveis e trabalhando na otimização do código. Uma possível solução é utilizar variações do algoritmo que trata as regras fuzzy, como por exemplo processamento em duas fases ou representação de vetores de bytes em tdlis ([Mangalampalli; Pudi, 2009](#)). O modelo desenvolvido é um protótipo inicial, no futuro, é possível dividir essas variáveis em classes e criar vários subsistemas fuzzy para estimar risco de desenvolver câncer de colo. Além disso, também seria ideal integrar o modelo fuzzy com o aplicativo para calcular em tempo real a classificação.

O modelo não pôde ser testado com uma base de dados reais de paciente, o que não invalida a pesquisa mas não garante assertividade nos resultados falando de uma perspectiva de diagnóstico. O aplicativo não foi utilizado para coleta devido às limitações de tempo para aprovação do projeto no comitê de ética por causa do uso de dados pessoais de pacientes.

Referências

AL-MADANI, W. et al. Modelling risk assessment for cervical cancer in symptomatic saudi women. *Saudi medical journal*, Saudi medical journal, v. 2019, 2019. Citado na página 29.

American Cancer Society. *Risk Factors for Cervical Cancer*. 2020. <<https://www.cancer.org/cancer/cervical-cancer/causes-risks-prevention/risk-factors.html>>. Acessado em: 2021-02-13. Citado na página 23.

AULT, K. A. Epidemiology and natural history of human papillomavirus infections in the female genital tract. *Infectious Diseases in Obstetrics and Gynecology*, Hindawi Publishing Corporation, v. 2006, 2006. Citado na página 19.

BASSANEZI, R.; BARROS, L. *Tópicos de Lógica Fuzzy e Biomatemática*. [S.l.: s.n.], 2010. ISBN 85-87185-05-5. Citado na página 31.

BASSANI, L. T. *Sistema de base de regras fuzzy como método alternativo para avaliar o risco de doenças cardiovasculares*. Dissertação (Mestrado) — IMECC, 2016. Citado na página 27.

BICHENG, H. et al. A risk evaluation model of cervical cancer based on etiology and human leukocyte antigen allele susceptibility. *International Journal of Infectious Diseases*, v. 28, p. 8–12, 2014. ISSN 1201-9712. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1201971214015434>>. Citado na página 30.

CABRERA-BECERRIL, A. et al. Modeling the dynamics of progression of chromosomal alterations in cervical cancer: a computational model. *bioRxiv*, Cold Spring Harbor Laboratory, p. 121814, 2017. Citado 2 vezes nas páginas 13 e 27.

Centers for Disease Control and Prevention. *High-Risk and Low-Risk HPV Types*. 2020. <https://www.cdc.gov/cancer/hpv/basic_info/index.htm>. Acessado em: 12/06/2020. Citado na página 19.

Centers for Disease Control and Prevention. *HPV Diseases and Cancers*. 2020. <<https://www.cdc.gov/hpv/parents/cancer.html>>. Acessado em: 09/06/2020. Citado na página 18.

Healthline. *Human Papillomavirus (HPV) and Cervical Cancer*. 2018. <<https://www.healthline.com/health/cervical-cancer/symptoms#prevention>>. Acessado em: 15/03/2021. Citado na página 16.

Instituto Nacional de Câncer. *Conceito e Magnitude*. 2021. <<https://www.inca.gov.br/controlado-cancer-do-colo-do-utero/conceito-e-magnitude>>. Acessado em: 01/06/2020. Citado 2 vezes nas páginas 12 e 15.

Instituto Oncoguia. *Risco Genético e Câncer*. 2015. <<http://www.oncoguia.org.br/conteudo/risco-genetico-e-cancer/99/6/>>. Acessado em: 18/06/2020. Citado na página 23.

International Agency for Research on Cancer. *Estimated age-standardized incidence rates (World) in 2020, cervix uteri, females, all ages*. 2018. <https://gco.iarc.fr/today/online-analysis-map?v=2018&mode=population&mode_population=continents&population=900&populations=900&key=asr&sex=2&cancer=23&type=0&statistic=5&prevalence=0&population_group=0&ages_group%255B%255D=0&ages_group%255B%255D=17&nb_items=10&group_cancer=0&include_nmssc=1&include_nmssc_other=1&projection=natural-earth&color_palette=default&map_scale=quantile&map_nb_colors=5&continent=0&rotate=%25255B19%25252C21.75%25252C0%25255D>. Acessado em: /05/2020. Citado na página 22.

LAGHETTO, B. K. Um modelo matemático para estimar o risco de desenvolver câncer de pulmão por meio de sistemas fuzzy. *Repositório Institucional UFSCar*, 2016. Citado 4 vezes nas páginas 13, 26, 31 e 40.

LANDY, R. et al. Challenges in risk estimation using routinely collected clinical data: The example of estimating cervical cancer risks from electronic health-records. *Preventive Medicine*, v. 111, p. 429–435, 2018. ISSN 0091-7435. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0091743517304826>>. Citado na página 29.

Mangalampalli, A.; Pudi, V. Fuzzy association rule mining algorithm for fast and efficient performance on very large datasets. In: *2009 IEEE International Conference on Fuzzy Systems*. [S.l.: s.n.], 2009. p. 1163–1168. Citado na página 61.

Medical Xpress. *Twenty five years of learning to combat cervical cancer*. 2018. <<https://medicalxpress.com/news/2019-11-twenty-years-combat-cervical-cancer.html>>. Acessado em: 01/07/2020. Citado na página 16.

React Native. *React Native*. 2021. <<https://reactnative.dev/>>. Acessado em: 08/06/2020. Citado na página 44.

ROTHBERG, M. et al. A risk prediction model to allow personalized screening for cervical cancer. *pumed*, v. 2018, 2018. Citado 2 vezes nas páginas 13 e 30.

SABINO, L. *Redes Neurais Artificiais, Lógica Nebulosa e Sistemas Neuro-Fuzzy na Previsão de Carga Elétrica em Curto Prazo*. Dissertação (Mestrado) — PUC-RIO - PONTIFÍCIA UNIVERSIDADE CATÓLICA DO RIO DE JANEIRO, 2016. Citado 2 vezes nas páginas 31 e 38.

SALDAÑA, F.; KOROBEINIKOV, A.; BARRADAS, I. Optimal control against the human papillomavirus: Protection versus eradication of the infection. In: HINDAWI. *Abstract and Applied Analysis*. [S.l.], 2019. v. 2019. Citado 2 vezes nas páginas 13 e 28.

SAMPAIO, L.; ALMEIDA, C. Vitaminas antioxidantes na prevenção do câncer do colo uterino. *Rev. bras. cancerol*, Biblioteca Geraldo Matos de Sá . Hospital do Câncer, v. 2009, 2009. Citado na página 21.

SENE, B.; MARIN, R. *Fuzzy Logic in Medicine*. [S.l.]: Physica-Verlag Heidelberg, 2002. ISBN 978-3-7908-1804-8. Citado 2 vezes nas páginas 13 e 14.

SILVA, F. F. B. *Desvendando a lógica Fuzzy*. Dissertação (Mestrado) — Universidade Federal de Uberlândia, 2011. Citado 2 vezes nas páginas 31 e 41.

SOUSA, R. de; XIMENES, J. de S.; BRITO, R. de. Sistema fuzzy como mecanismo auxiliador na predição da doença viral covid-19. In: *Anais da VIII Escola Regional de Computação do Ceará, Maranhão e Piauí*. Porto Alegre, RS, Brasil: SBC, 2020. p. 30–37. ISSN 0000-0000. Disponível em: <<https://sol.sbc.org.br/index.php/ercemapi/article/view/11465>>. Citado 2 vezes nas páginas 14 e 27.

United Nations Development Programme. *Human Development Report 2020*. 2020. <<http://hdr.undp.org/sites/default/files/hdr2020.pdf>>. Acessado em: 15/03/2021. Citado na página 15.

Vaccine Knowledge Project. *HPV (Human Papillomavirus)*. 2014. <<https://vk.ovg.ox.ac.uk/vk/hpv>>. Acessado em: 01/07/2020. Citado na página 18.

World Health Organization. *UN Joint Global Programme on Cervical Cancer Prevention and Control*. 2016. <<https://www.who.int/ncds/un-task-force/un-joint-action-cervical-cancer-leaflet.pdf?ua=1>>. Acessado em: 09/06/2020. Citado na página 17.

World Health Organization. *Tools for implementing WHO PEN (Package of essential noncommunicable disease interventions)*. 2021. <[https://www.who.int/ncds/management/pen_tools/en/#:~:text=The%20WHO%20Package%20of%20Essential,even%20in%20resource%2Dpoor%20settings](https://www.who.int/ncds/management/pen_tools/en/#:~:text=The%20WHO%20Package%20of%20Essential,even%20in%20resource%2Dpoor%20settings.)>. Acessado em: 04/06/2020. Citado na página 16.

YILMAZ, A.; AYAN, K. Cancer risk analysis by fuzzy logic approach and performance status of the model. *Turkish Journal of Electrical Engineering and Computer Sciences*, v. 21, p. 897–912, 01 2013. Citado 2 vezes nas páginas 14 e 26.