

# Risco de desenvolver câncer de pulmão por meio de um sistema *fuzzy*

Beatriz K. Laghetto,<sup>1</sup> Magda S. Peixoto,<sup>2</sup>  
DFQM, Universidade Federal de São Carlos, 18052-780, Sorocaba/SP.

**Resumo.** Neste artigo propomos um modelo matemático para estimar o risco de um indivíduo desenvolver câncer de pulmão por meio de um sistema baseado em regras *fuzzy*. Consideramos, para isso, os principais fatores de risco (tabagismo, poluição, histórico de doenças pulmonares, genética e contato com agentes químicos) como variáveis de entrada do sistema. O câncer de pulmão é uma doença que não apresenta sintomas em suas fases iniciais, tornando o diagnóstico mais difícil de ser feito e, por isso, a maioria descobre quando o câncer já está avançado. Esse tipo de câncer é altamente letal e frequente na população, apresentando aumento de 2% ao ano na sua incidência mundial.

**Palavras-chave:** conjuntos *fuzzy*; câncer de pulmão; fatores de risco, Método de inferência de Mamdani.

## 1. Introdução

O câncer é uma doença que resulta, principalmente, de alterações genéticas, fatores ambientais e do estilo de vida. Dentre os muitos tipos de câncer, propomos um estudo do câncer de pulmão que tem como principal causa o tabagismo.

Segundo Zamboni (2002), o câncer de pulmão passou de uma doença rara no passado para a doença neoplásica mais comum e mais mortal em todo o mundo. Ainda segundo seus dados, este tipo de câncer é a principal causa de morte por câncer entre os homens da América do Norte e Europa e sua mortalidade vem aumentando, significativamente, entre a população da Ásia, América Latina e da África.

---

<sup>1</sup>biakralla@hotmail.com

<sup>2</sup>magda@ufscar.br

As modificações no comportamento da doença começaram a ser notadas a partir da década de 20, quando o número de casos começou a crescer progressivamente, transformando-se numa verdadeira epidemia mundial no início do século XXI. Entre 1960 e 1980, sua mortalidade aumentou 76% entre as mulheres e cerca de 135% entre os homens (Zamboni, 2002).

Os primeiros estudos controlados sobre a doença começaram a ser feitos em 1950, demonstrando que o tabagismo tinha relação direta com o câncer de pulmão e que junto com a explosão do hábito de fumar a partir dos anos 20, veio também a explosão dos casos de câncer de pulmão (Zamboni, 2002).

O fumo tornou-se então um grave problema de saúde pública, podendo se tornar, até 2030, a maior causa isolada de mortalidade, matando até 10 milhões de pessoas por ano. A indústria do fumo é voltada para aqueles menos privilegiados economicamente e socialmente, tendo como principal público alvo os adolescentes (Menezes et al., 2002).

O objetivo deste trabalho é elaborar um modelo matemático para estimar o risco de um indivíduo desenvolver câncer de pulmão por meio de um sistema baseado em regras *fuzzy*. Consideramos, para isso, os fatores de risco, tabagismo, poluição, histórico de doenças pulmonares, genética e contato com agentes químicos como variáveis de entrada do sistema *fuzzy*.

## 2. Sistemas baseados em regras *fuzzy*

As ações humanas controlam os mais diversos sistemas do mundo real por meio de informações imprecisas. Cada indivíduo recebe informações que são interpretadas de acordo seus parâmetros e, então, decide que atitudes tomar. O controle e execução de tarefas devem seguir uma sequência de “ordens” linguísticas, traduzidas por um conjunto de regras, capazes de serem codificadas por um controlador. E baseado nisso, são esses componentes que um sistema baseado em regras *fuzzy* possui: um processador de entrada (ou *fuzzificador*), um conjunto de regras linguísticas, um método de inferência *fuzzy* e um processador de saída (ou *defuzzificador*), gerando um número real como saída (ver Figura 1) (Barros e Bassanezi, 2010), (Peixoto, 2005).

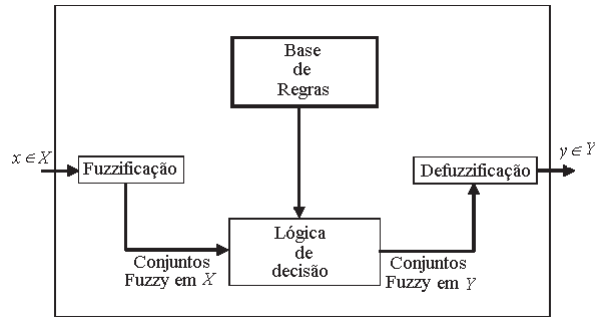


Figura 1: Esquema geral de um controlador *fuzzy*.

No sistema, cada uma das variáveis de entrada é modelada matematicamente por um conjunto *fuzzy*. Chamamos esse processo de *fuzzificação*.

Um subconjunto clássico é definido de tal forma: Seja  $U$  um conjunto universo e  $A$  um subconjunto de  $U$ . A função característica de  $A$  é dada por

$$X_A(x) = \begin{cases} 1, & \text{se } x \in A \\ 0, & \text{se } x \notin A \end{cases} \quad (2.1)$$

Essa função tem objetivo de indicar quando um elemento  $x \in U$  pertence ou não a  $A$ , dependendo da sua imagem em  $\{0,1\}$ . Assim, a função característica descreve completamente o conjunto  $A$ , uma vez que indica quais elementos do conjunto  $U$  são elementos de  $A$  (Barros e Bassanezi, 2010). A Teoria dos Conjuntos *Fuzzy* faz essa classificação de forma gradual, e tem como definição: Seja  $U$  um conjunto (clássico); um subconjunto *fuzzy*  $F$  de  $U$  é caracterizado por uma função de pertinência dada por

$$\varphi_F : U \rightarrow [0, 1] \quad (2.2)$$

que indica o grau de pertinência de  $x \in U$  a  $F$  (Zadeh, 1965).

A base de regras é composta por uma coleção de proposições condicionais *fuzzy* na forma de regras SE-ENTÃO. Um especialista do fenômeno analisado fornece as informações para se formular o conjunto de regras *fuzzy* que ativa as associações de entradas/saídas linguísticas.

O método de inferência que utilizamos é o de Mamdani, que agrega as regras por meio do operador lógico OU, modelado pelo operador máximo e, em cada regra, os operadores lógicos E e ENTÃO são modelados pelo operador mínimo (Klir e Yuan, 1995).

Por fim, o método de *defuzzificação* utilizado nesse trabalho é o centróide. O centróide dá a média das áreas de todas as figuras que representam os graus de pertinência de um subconjunto *fuzzy* (Barros e Bassanezi, 2010).

### 3. Formulação do modelo *fuzzy*

Para a elaboração deste modelo, utilizou-se os dados de Uehara et al. (1998), Zamboni (2002), Omenn et al. (1996) e CETESB (2016), ou seja:

- Tabagismo - Nunca fumou e não é fumante passivo = NÃO RISCO; Até 1 ano de fumo ou fumante passivo = POUCO RISCO e acima de um ano de fumo = RISCO;
- Poluição - Até nível 4 = NÃO RISCO, do nível 5 em diante = RISCO;
- Histórico de doenças pulmonares - Sim = RISCO, Não = NÃO RISCO;
- Histórico familiar - Não = Não Risco, Parentesco de 2<sup>o</sup> grau = MEIO RISCO e Parentesco do 1<sup>o</sup> grau = RISCO
- Contato com Agentes Químicos- Não = NÃO RISCO, sim = RISCO;

A proposta desse modelo é definir os fatores acima como conjuntos *fuzzy*, ou seja, de forma mais gradual. Estas informações qualitativas, encontradas nas referências citadas no início dessa seção, permitem propor regras que relacionem os fatores acima com o risco de um indivíduo desenvolver o câncer de pulmão.

Dessa forma, o sistema *fuzzy* aqui proposto possui cinco variáveis de entrada (Tabagismo, Poluição, Histórico de Doenças Pulmonares, Histórico familiar, Contato com Agentes Químicos) e uma variável de saída (risco de um indivíduo desenvolver um câncer de pulmão):

- A variável *fuzzy* tabagismo é definida pelos subconjuntos *fuzzy* {não fumante, fumante} (ver Figura 2).
- A variável *fuzzy* Poluição é definida pelos subconjuntos *fuzzy* {pouca, muita} de acordo com classificação feita pela CETESB de acordo com a Tabela 1 (CETESB, 2016) (ver Figura 3).

Tabela 1: Qualidade do Ar e Efeitos à Saúde. Fonte: CETESB (2016)

Qualidade	Índice	Significado
N1 - Boa	0-40	-
N2 - Moderada	41-80	Pessoas de grupos sensíveis (crianças, idosos e pessoas com doenças respiratórias e cardíacas) podem apresentar sintomas como tosse seca e cansaço. A população, em geral, não é afetada.
N3 - Ruim	81-120	Toda a população pode apresentar sintomas como tosse seca, cansaço, ardor nos olhos, nariz e garganta. Pessoas de grupos sensíveis (crianças, idosos e pessoas com doenças respiratórias e cardíacas) podem apresentar efeitos mais sérios na saúde
N4 - Muito Ruim	121-200	Toda a população pode apresentar agravamento dos sintomas como tosse seca, cansaço, ardor nos olhos, nariz e garganta e ainda falta de ar e respiração ofegante. Efeitos ainda mais graves à saúde de grupos sensíveis (crianças, idosos e pessoas com doenças respiratórias e cardíacas).
N5 - Péssima	200	Toda a população pode apresentar sérios riscos de manifestações de doenças respiratórias e cardiovasculares. Aumento de mortes prematuras em pessoas de grupos sensíveis.

- A variável *fuzzy* Histórico de Doenças Pulmonares é definida pelos subconjuntos *crisp* {sim, não} (ver Figura 4).
- A variável *fuzzy* Histórico familiar é definida pelos subconjuntos *fuzzy* {não, sim} (ver Figura 5).
- A variável *fuzzy* Contato com agentes químicos é definida pelos subconjuntos *crisp* {sim, não} (ver Figura 6).
- A variável *fuzzy* Risco é definida pelos subconjuntos *fuzzy* {baixo, médio, alto} (ver Figura 7)

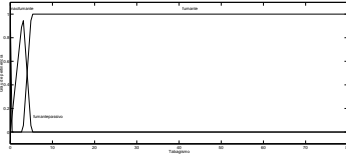


Figura 2: Variável Tabagismo.

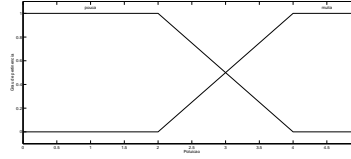


Figura 3: Variável Poluição.

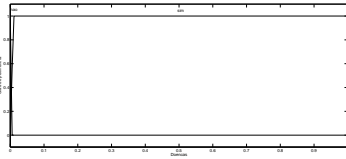


Figura 4: Variável Histórico de Doenças Pulmonares.

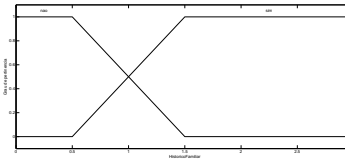


Figura 5: Variável Histórico familiar.

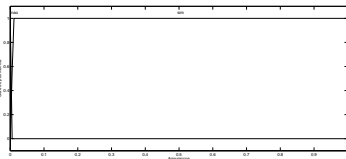


Figura 6: Variável Contato com Agentes Químicos.

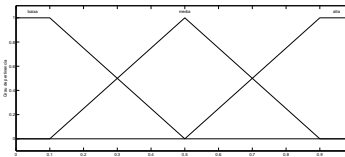


Figura 7: Variável de saída Risco.

Foi elaborada uma base de regras, num total de 32 regras, do tipo: 1. “Se (tabagismo é não fumante) e (poluição é pouca) e (Histórico de Doenças Pulmonares é não) e (Histórico familiar é não) e (Contatos com Agentes Químicos é não), então (risco é baixo)”;

2. “Se (tabagismo é fumante) e (poluição é muita) e (Histórico de Doenças Pulmonares é sim) e (Histórico Familiar é sim) e (Contatos com Agentes Químicos é sim), então (risco é alto)”.

Adotou-se o Método de Inferência de Mamdani e o centróide como método de *defuzzificação*.

O sistema *fuzzy* pode ser resumido no esquema dado pela Figura 8.

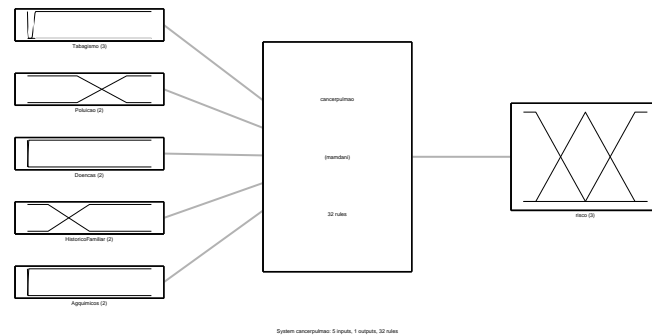


Figura 8: Sistema fuzzy.

## 4. Resultados

O modelo matemático foi desenvolvido utilizando-se a ferramenta *Toolbox Fuzzy* do software *MatLab*.

Por exemplo, considere os casos de dois pacientes hipotéticos a as simulações obtidas por meio do programa:

**Exemplo 1:** Suponha um indivíduo que

- Fuma há mais de 30 anos;
- Mora em uma cidade grande e poluída;
- Possui histórico de doenças pulmonares;
- Possui parentes de 1<sup>o</sup> grau com a doença;
- Já esteve em contato com Agentes Químicos;

Como o programa é auto-explicativo, deve digitar de acordo com os comandos:

*Para Variável “tabagismo” considere:*

*Se você não for fumante e não conviver com um fumante digite zero.*

*Se você não for fumante, porém conviver com um fumante digite uma nota de 1 a 5.*

*Se você for (ou foi) fumante digite há quantos anos você fuma (ou fumou).*

*Digite o valor atribuído à variável “tabagismo”:* **40**

*De 0 a 5, qual o nível de poluição você está exposto diariamente:* **5**

*Você teve, ou têm alguma doença pulmonar grave (como, tuberculose ou Doença Pulmonar Obstrutiva Crônica)? Se sim, digite 1, se não, digite 0:* **1**

*Você possui histórico familiar de câncer de pulmão? Se não, digite 0. Se sim e o familiar for de 2<sup>o</sup> grau, digite 1. Ou, se sim e o familiar for de 1<sup>o</sup> grau, digite 2:* **2**

*Você já esteve exposto a algum agente tóxico como: Arsênico, amianto, asbesto, berílio, cromo, radônio, níquel, cádmio ou cloreto em seu trabalho? Se sim, digite 1, se não, digite 0:* **1**

*Risco de desenvolver câncer de pulmão:* **0.8817**

Nas condições citadas acima, numa escala de 0 a 1, o indivíduo possui aproximadamente 0,9 de risco de desenvolver um câncer de pulmão, isto é, estimativa gerada pelo sistema *fuzzy* proposto.

**Exemplo 2:** Considere agora um indivíduo que

- Nunca Fumou e teve pouco contato com fumantes;
- Mora em uma cidade do interior em contato com a natureza;
- Não possui histórico de doenças pulmonares;
- Possui parentes de 2<sup>o</sup> grau com a doença;
- Nunca esteve em contato com Agentes Químicos;

A pessoa, então, deve responder as perguntas:

*Para Variável “tabagismo” considere:*

*Se você não for fumante e não conviver com um fumante digite zero.*

*Se você não for fumante, porém conviver com um fumante digite uma nota de 1 a 5.*

*Se você for (ou foi) fumante digite há quantos anos você fuma (ou fumou).*

*Digite o valor atribuído à variável “tabagismo”:* **2**

*De 0 a 5, qual grau de poluição você está exposto diariamente:* **3**

*Você teve, ou têm alguma doença pulmonar grave (como, tuberculose ou Doença Pulmonar Obstrutiva Crônica)? Se sim, digite 1, se não, digite 0:* **0**

*Você possui histórico familiar de câncer de pulmão? Se não, digite 0. Se sim e o familiar for de 2<sup>o</sup> grau, digite 1. Ou, se sim e o familiar for de 1<sup>o</sup> grau, digite 2:* **1**



*Você já esteve exposto a algum agente tóxico como: Arsênico, amianto, asbesto, berílio, cromo, radônio, níquel, cádmio ou cloreto em seu trabalho? Se sim, digite 1, se não, digite 0: 0*

*Risco de desenvolver câncer de pulmão: 0.4022*

Nas condições citadas acima, numa escala de 0 a 1, o indivíduo possui aproximadamente 0,4 de risco de desenvolver um câncer de pulmão, isto é, estimativa gerada pelo sistema *fuzzy* proposto.

## 5 Conclusões

Na modelagem de fenômenos biológicos é comum se deparar com incertezas, envolvendo variáveis que não são precisamente quantificadas. Muitas das incertezas que encontramos nos fenômenos são provenientes apenas da subjetividade da nossa linguagem. Esse é o caso típico de alguns procedimentos adotados em Biomedicina para diagnosticar e controlar alguma doença em paciente. Por exemplo, para se controlar determinada doença, observam-se os sinais ou sintomas apresentados pelo paciente. A gravidade destes sinais indicará o procedimento médico a ser adotado. O termo “gravidade” é subjetivo no sentido de apresentar graduações. Pois bem, incertezas devido à gradualidade são tipicamente tratadas por meio de métodos *fuzzy*. Esta teoria tem se mostrado mais adequada no tratamento de variáveis incertas e subjetivas do que a matemática clássica. Esse artigo sugere a utilização da Teoria dos Conjuntos *Fuzzy* em Medicina.

Utilizamos um sistema baseado em regras *fuzzy* para elaborar um modelo matemático para estimar o risco de um indivíduo desenvolver câncer de pulmão. Para isso, consideramos os principais fatores de risco, isto é, tabagismo, poluição, histórico de doenças pulmonares, genética e contato com agentes químicos, de acordo com dados da literatura, como variáveis de entrada do sistema *fuzzy*.

Nos modelos *fuzzy* usam-se base de regras ao invés de equações explícitas. Utilizamos o método de inferência de Mamdani e o centroide como defuzzificador. As simulações foram realizadas no *Toolbox Fuzzy* do ambiente *Matlab*. Para encontrar a solução, digitou-se dados fictícios de um indivíduo. As variáveis e a base de regras do sistema *fuzzy* foram elaboradas segundo dados da literatura (Uehara et al., 1998; Zamboni, 2002; Omenn et al., 1996; CETESB, 2016).

## Agradecimentos

A primeira autora agradece à CAPES pela bolsa de Mestrado.

## Referências

- Barros, L. C. e Bassanezi, R. C. (2010). *Tópicos de Lógica Fuzzy e Biomatemática*, volume 5 of *Coleção Textos Didáticos*. IMECC–UNICAMP.
- CETESB (2016). Makeindex: An index processor for latex. URL: <http://www.cetesb.sp.gov.br/> Acesso em: 15/02/2016.
- Klir, G. J. e Yuan, B. (1995). *Fuzzy Sets and Fuzzy Logic - Theory and Application*. Prentice Hall, N.Jersey.
- Menezes, A., Horta, B. L., Oliveira, A. L. B., Kaufmann, R. A. C., Duquia, R., Diniz, A., Motta, L. H., Centeno, M. S., Estanislau, G., e Gomes, L. (2002). Risco de câncer de pulmão, laringe e esôfago atribuível ao fumo. *Rev. Saúde Pública*, 36(2):129–134.
- Omenn, G. S., Goodman, G. E., Thornquist, M. D., Balmes, J., Cullen, M. R., Glass, A., Keogh, J. P., Jr., F. L. M., Valanis, B., Jr., J. H. W., Barnhart, S., Cherniack, M. G., Brodtkin, C. A., e Hammar, S. (1996). Risk factors for lung cancer and for intervention effects in caret, the beta-carotene and retinol efficacy trial. *J. National Cancer Institute*, 88(21):1550–1559.
- Peixoto, M. S. (2005). Sistemas dinâmicos e controladores *Fuzzy*: um estudo da dispersão da morte súbita dos citros em são paulo. Master's thesis, IMECC–UNICAMP, Campinas/SP.
- Uehara, C., Jamnik, S., e Santoro, I. L. (1998). Câncer de pulmão. *Medicina*, 31:266–276.
- Zadeh, L. A. (1965). Fuzzy sets. *Information and Control*, 8:338–353.
- Zamboni, M. (2002). Epidemiologia do câncer do pulmão. *J. Pneumol.*, 28(1):41–47.