

# Previsões de máximo de casos confirmados e óbitos de COVID-19 no Brasil

## Predictions of the maximum number of confirmed cases and deaths of COVID-19 in Brazil

### RESUMO

Carlos Maximiliano Dutra\* 

**Introdução:** A COVID-19 teve início em dezembro de 2019 na China e em fevereiro de 2020 no Brasil. Até 29 de agosto de 2020, o Brasil possuía 3.717.156 casos confirmados (60,6% do total da América e 15,0% do total mundial) e 117.665 óbitos (59,4% do total da América e 14,0% do total mundial), ocupando o segundo lugar no mundo. **Objetivo:** Estimar limites máximos para os quantitativos de casos confirmados e óbitos por COVID-19 no Brasil. **Método:** A estimativa dos máximos de casos confirmados e óbitos foi realizada através da aplicação de regressão não linear com ajuste da função Gompertz aos dados de COVID-19 registrados até final de agosto de 2020. **Resultados:** Como resultados foram obtidos máximo de 7.189.300 casos confirmados, com 95,0% do máximo atingidos em fevereiro 2021 e 164.810 óbitos, com 95,0% do máximo atingidos em dezembro 2020. **Conclusões:** Essas previsões expressam a tendência da evolução temporal dos dados notificados de COVID-19 pelo *Our World in Data*. Eles não consideram a subnotificação no número de contaminados e de óbitos, tendo em conta a limitada capacidade de testagem da população e a existência de contaminados assintomáticos não contabilizados. Estes fatores somados ao relaxamento descontrolado das medidas preventivas individuais e do distanciamento social poderão ter um efeito significativos nas previsões do modelo matemático usado neste trabalho e no entendimento da transmissão do vírus no Brasil.

**PALAVRAS-CHAVE:** Previsão; Coronavírus; COVID-19; Gompertz

### ABSTRACT

**Introduction:** COVID-19 started in December 2019 in China and in February 2020 in Brazil. As of August 29, 2020, Brazil had 3,717,156 confirmed cases (60.6% of the total in America and 15.0% of the total World) and 117,665 deaths (59.4% of the total in America and 14.0% of the World total), the second place in the world. **Objective:** To estimate maximum limits for the quantitative number of confirmed cases and deaths of COVID-19 in Brazil. **Method:** The maximum estimates of confirmed cases and deaths were estimated by applying nonlinear regression with adjustment of the Gompertz function to the COVID-19 data recorded until the end of August 2020. **Results:** As results were obtained a maximum of 7,189,300 confirmed cases, with 95.0% of the maximum reached in February 2021 and 164,810 deaths, with 95.0% of the maximum reached in December 2020. **Conclusions:** These forecasts express the trend of temporal evolution of COVID-19 reported data by *Our World in Data*. They do not consider underreporting in the number of contaminated people and deaths, taking into account the limited testing capacity of the population and the existence of unaccounted asymptomatic contaminated people. These factors added to the uncontrolled relaxation of individual preventive measures and social distancing may have a significant effect on the predictions of the mathematical model used in this work and on the understanding of the transmission of the virus in Brazil.

Universidade Federal do Pampa  
(Unipampa), Uruguaiana, RS, Brasil

\* E-mail: [carlosdutra@unipampa.edu.br](mailto:carlosdutra@unipampa.edu.br)

Recebido: 31 ago 2020  
Aprovado: 09 nov 2020

**KEYWORDS:** Prediction; Coronavirus; Covid-19; Gompertz



## INTRODUÇÃO

Os primeiros casos de infecção pelo vírus SARS-CoV2 ocorreram na China em dezembro de 2019, conforme a Organização Mundial da Saúde (OMS)<sup>1</sup>. Em 11 de março de 2020, a OMS declarou a COVID-19 como pandemia, quando havia 118.319 casos confirmados e 4.292 óbitos em 113 países<sup>2</sup>. Até essa data a América do Sul contava com 109 casos confirmados e um óbito, enquanto o Brasil possuía apenas 34 casos confirmados<sup>2</sup>, com o primeiro óbito registrado apenas em 12 de março de 2020. Cerca de 5,5 meses depois, em 29 de agosto de 2020 (10h30, horário de Londres), tendo como fonte o banco de dados de COVID-19 da *Our World in Data* (disponível no site: <https://ourworldindata.org/coronavirus-source-data>), o mundo registrava 24.760.129 casos confirmados e 837.579 óbitos; enquanto que a América do Sul alcançou aproximadamente 6,1 milhões de casos confirmados e 198,1 mil óbitos. Ainda nessa data de 29 de agosto de 2020, o Brasil possuía 3.717.156 casos confirmados (60,6% do total da América e 15,0% do total mundial) e 117.665 óbitos (59,4% do total da América e 14,0% do total mundial), com o Brasil ocupando o segundo lugar no mundo como país com mais casos confirmados de COVID-19.

Esta rápida evolução no número de casos confirmados e no número de óbitos vem sendo acompanhada de uma forma sem precedentes pelos cidadãos do mundo com atualizações diárias através de sites como: COVID-19 *Dashboard* - John Hopkins University (<https://coronavirus.jhu.edu/map.html>); Bing COVID (<https://www.bing.com/covid>); e *Worldometer Coronavirus* (<https://www.worldometers.info/coronavirus/>), bem como em repositórios de dados como *Our World in Data*. Com a disponibilidade de dados da evolução do número de casos e de óbitos com o tempo, vários trabalhos têm sido produzidos para realizar previsões sobre o impacto da COVID-19 em diferentes países. Perez et al.<sup>3</sup> utilizaram o método de regressão não linear da função de crescimento Gompertz junto aos dados acumulativos de casos confirmados e óbitos até 11 de maio de 2020 para a Itália, Espanha, China e Coreia do Sul e obtiveram previsões com valores compatíveis com os efetivamente verificados nos bancos de dados citados anteriormente. Medina-Mendleta et al.<sup>4</sup> realizaram ajuste de regressão não linear com os modelos Logístico e Gompertz junto aos dados de casos confirmados/óbitos de COVID-19 da Espanha e da Itália, determinando corretamente as datas de picos na distribuição de casos confirmados/óbitos, previram o limite máximo para os valores acumulativos de casos confirmados/óbitos e ainda realizaram previsões para Cuba. Essas predições corresponderam a primeira “onda” de infecção da COVID-19. Em novembro de 2020, países da Europa como a Espanha (<https://www.worldometers.info/coronavirus/country/spain/>) e a Itália (<https://www.worldometers.info/coronavirus/country/italy/>) estão desenvolvendo uma segunda “onda” com números significativamente maiores de contágio do que a primeira “onda”. Esta nova fase de crescimento poderá ser novamente modelada através de regressão não linear para novas estimativas nesses países.

No presente trabalho utilizou-se o método de regressão não linear da função Gompertz para realizar a estimativa do número total de casos confirmados e de óbitos para o Brasil.

## MÉTODO

Conforme Domingues<sup>5</sup>, a função Gompertz foi desenvolvida por Benjamin Gompertz em 1938 para estimar o limite para o crescimento de uma população de tumores sólidos. Ao longo do tempo verifica-se que este crescimento acumulativo apresenta as seguintes fases: *Lag* - fase inicial de crescimento lento; Exponencial - com crescimento acelerado e passa a ter crescimento desacelerado no chamado ponto de inflexão; *Plateau* - crescimento convergindo para um valor limite máximo da população.

Na equação (1) tem-se o modelo Gompertz definido conforme os parâmetros que utilizados para caracterizar o crescimento acumulativo de óbitos devido à COVID-19.

$$N(t) = Ae^{-e^{B(t-C)}} \quad (1)$$

Onde  $N(t)$  é a função do número de indivíduos acumulados da população de interesse em função do tempo “ $t$ ”;  $A$  é o valor máximo ou valor limite assintótico para o número de óbitos acumulados;  $B$  é o crescimento relativo no ponto de inflexão (aproximadamente número de casos novos dividido pelo número de casos acumulados no instante do ponto de inflexão); e  $C$  corresponde ao instante de tempo  $t$  em que ocorre o ponto de inflexão.

Dutra et al.<sup>6</sup> aplicaram o método de regressão não linear com modelo Gompertz aos dados acumulativos de casos confirmados de oito países que estavam em fase de estabilidade da pandemia e verificaram que, quando o método é aplicado após o país ter atingido o pico na distribuição de casos novos, ocorre uma convergência que permite reproduzir os dados já registrados na fase de estabilidade da pandemia nesses países, demonstrando que as previsões por meio do método de regressão não linear da função Gompertz nessas condições tendem a se concretizar.

Para determinação dos limites máximos de casos confirmados e de óbitos para o Brasil, foi realizada primeiramente a identificação junto à distribuição ao longo do tempo dos dados de casos novos diários e óbitos novos diários, se os mesmos já abandonaram a fase de crescimento exponencial acelerado e se alcançaram um pico de distribuição. Em caso afirmativo, foi realizada a regressão não linear por mínimos quadrados do modelo Gompertz aos dados de casos confirmados e óbitos acumulados por dia. Por meio desse método estatístico foram determinados os valores dos parâmetros da função Gompertz, principalmente o parâmetro “ $A$ ” que resulta o limite máximo para o número de casos confirmados ou óbitos acumulados para o Brasil.

A partir da Equação (1) pode-se determinar o tempo transcorrido da pandemia para alcançar determinado número de casos confirmados ou óbitos acumulados isolando a variável tempo “ $t$ ”, conforme apresentado na Equação (2)

$$t = \frac{1}{B} \cdot \ln \left( \ln \frac{A}{N(t)} \right) + C \quad (2)$$

Onde os parâmetros  $A$ ,  $B$  e  $C$  são obtidos pela regressão não linear e descrevem a função  $N(t)$ . Para realizar as previsões



temporais em relação a pandemia se considera a Equação (2), quando a função Gompertz  $N(t)$  atingirá determinada porcentagem do valor máximo  $A$ : 80,0%; 90,0% e 95,0%.

## RESULTADOS E DISCUSSÃO

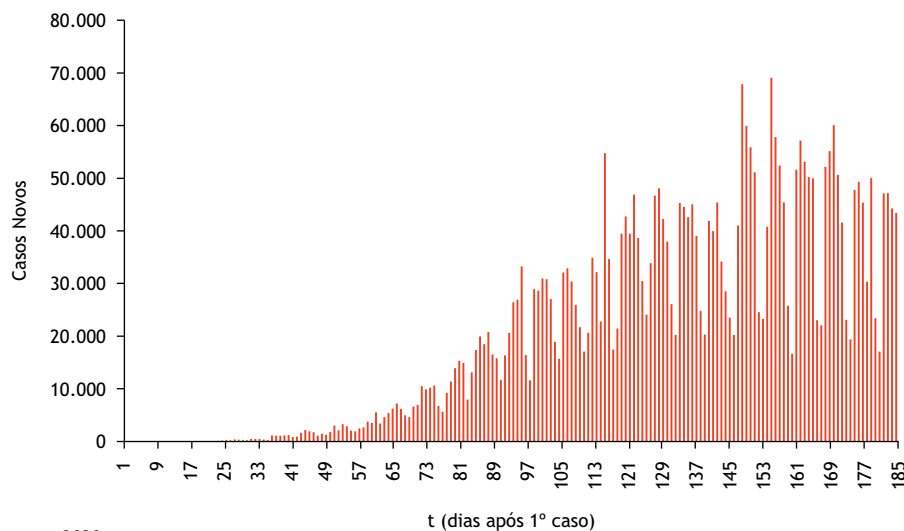
Para realização das estimativas definidas pelo método foram utilizados os dados de casos confirmados e óbitos da COVID-19 no Brasil, atualizados até 29 de agosto de 2020 (10h30 horário de Londres), disponíveis no repositório de dados de COVID-19 *Our World in Data*.

Na Figura 1 é apresentada a evolução temporal dos casos confirmados diários ao longo dos dias após o primeiro caso registrado no Brasil. Observa-se que o crescimento acelerado cessou e já ocorreu a formação de um pico. Essa configuração indica que a

variação de casos confirmados acumulados com o tempo pode ser modelada pela função Gompertz.

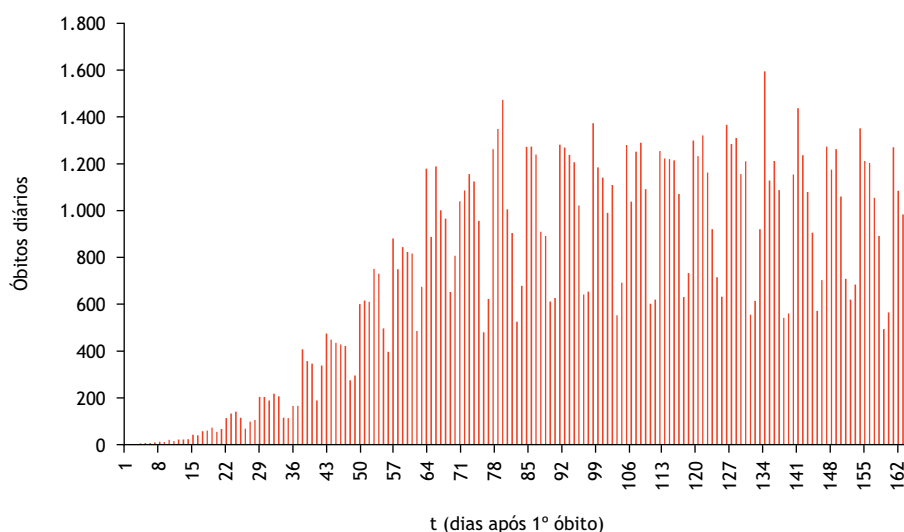
Na Figura 2 é apresentada a evolução temporal de óbitos diários ao longo dos dias após o primeiro óbito registrado no Brasil. Verifica-se que a distribuição temporal de óbitos também abandonou o crescimento exponencial e acelerado com um crescimento constante e estável com a formação de um pico alargado.

A partir da análise das Figuras 1 e 2 é possível utilizar a regressão não linear da função Gompertz junto às distribuições de casos confirmados e óbitos acumulados com o tempo para realizar a estimativa do máximo de casos confirmados e óbitos para o Brasil. Utilizando o *software* livre estatístico PAST<sup>7</sup> (disponível no *site*: <https://folk.uio.no/ohammer/past/>), foi realizada a regressão não linear determinando os parâmetros da função Gompertz conforme apresentado no Quadro 1.



Fonte: Elaborada pelo autor, 2020.

Figura 1. Distribuição de casos novos de COVID-19 ao longo dos dias após o primeiro caso registrado no Brasil.



Fonte: Elaborada pelo autor, 2020.

Figura 2. Distribuição de óbitos diários devido à COVID-19 em função do tempo a partir do 1º óbito registrado no Brasil.

**Quadro 1.** Sumário dos resultados dos parâmetros ajustados da função Gompertz aos dados de casos confirmados e óbitos acumulados de COVID-19 no Brasil.

Acumulados	A	B	C
Casos confirmados	7,1893e06	-0,0160	157,86
Intervalos de confiança 95%	(7,024e06;7,348e06)	(-0,0163; -0,0157)	(156,41; 159,24)
Óbitos	164810	-0.0185	106,72
Intervalos de confiança 95%	(161200;169100)	(-0,0189; -0,0182)	(105,50; 107,25)

Fonte: Elaborado pelo autor, 2020.

A partir dos parâmetros determinados no Quadro 1, tem-se estimativa de máximo de casos confirmados acumulados de 7.189.300 e máximo de óbitos acumulados de 164.810 para o Brasil. Esses resultados refletem a análise de regressão não linear da evolução temporal dos casos confirmados e óbitos de COVID-19 devidamente registrados pelos sistemas de saúde e compilados pelo Ministério da Saúde. Deve-se destacar que estes números de casos confirmados não representam os números reais de infectados pelo vírus SARS-CoV2. A subnotificação de casos tem sido variável ao longo da evolução da pandemia da COVID-19 no Brasil pelo uso limitado de testes ou orientação política de realização de testes rápidos e RT-PCR adotada nos estados brasileiros. Até 20 de abril de 2020 foi estimada uma taxa de notificação de casos confirmados de COVID-19 de 9,2%, indicando uma subnotificação de que o número real de infectados seria 8-11 vezes maior que o número de casos notificados<sup>8</sup>.

Na Figura 3 é apresentado o resultado gráfico do ajuste de regressão não linear, onde a linha vermelha representa a função Gompertz ajustada e os pontos azuis, os dados de casos acumulados.

Na Figura 4 é apresentado o resultado gráfico do ajuste de regressão não linear, onde a linha vermelha representa a função Gompertz ajustada e os pontos azuis, os dados de óbitos acumulados.

A partir do máximo de casos confirmados acumulados de 7.189.300 e máximo de óbitos acumulados de 164.810, é possível

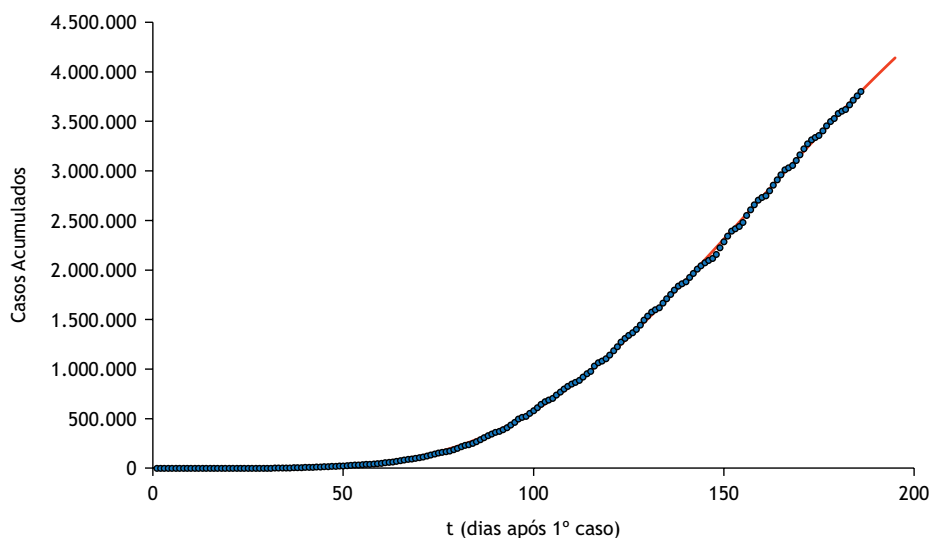
estimar a taxa de letalidade (definida como o número de óbitos por determinada doença e o número total de doentes multiplicada por 100) que, neste caso, será de 2,29% para o COVID-19 no Brasil.

Utilizando-se a Equação 2 e os parâmetros determinados no Quadro 1 é possível realizar as seguintes previsões para a pandemia de COVID-19 alcançar 80,0%, 90,0% e 95,0% dos máximos preditos pela regressão não linear da função Gompertz para os casos confirmados e óbitos no Brasil conforme o Quadro 2.

A partir do Quadro 2 observa-se que 95,0% do máximo de casos confirmados acumulados ocorrerá em fevereiro de 2021 e 95,0% do máximo de óbitos acumulados ocorrerá em dezembro de 2020.

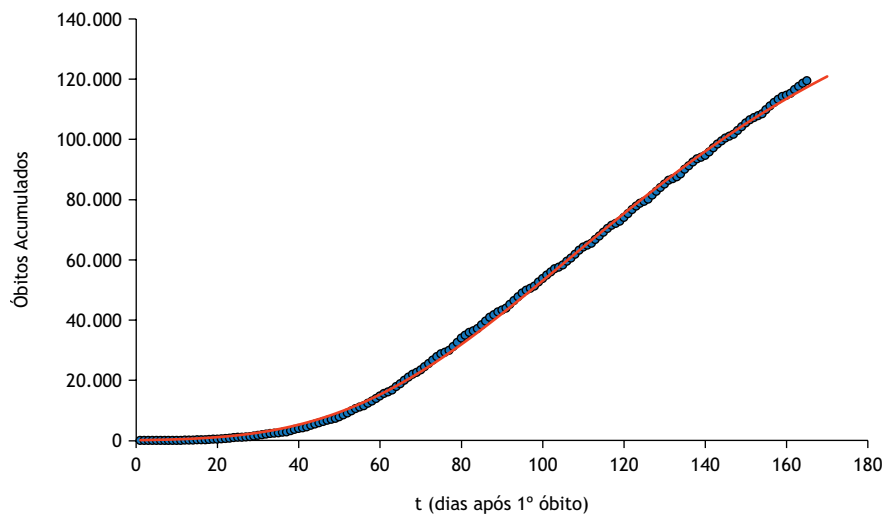
## CONCLUSÕES

O presente trabalho realizou uma estimativa de máximo de casos confirmados e de óbitos devido à COVID-19 para o Brasil utilizando o método de regressão não linear da função Gompertz junto à evolução temporal dos dados registrados de COVID-19 até a data de 29 de agosto de 2020, conforme dados compilados pelo Ministério da Saúde. Essa mesma metodologia foi utilizada com êxito em outros estudos internacionais de evolução da COVID-19. Foi possível estimar máximo de 7.189.300 casos confirmados e 164.810 óbitos com uma letalidade de 2,29%, com 95,0% do máximo



Fonte: Elaborada pelo autor, 2020.

**Figura 3.** Resultado gráfico da regressão não linear Gompertz aos dados de casos acumulados de COVID-19 no Brasil.



Fonte: Elaborada pelo autor, 2020.

Figura 4. Resultado gráfico da regressão não linear Gompertz aos dados de óbitos acumulados de COVID-19 no Brasil.

Quadro 2. Sumário das previsões aos dados de casos confirmados e óbitos acumulados COVID-19 no Brasil.

Casos confirmados	80,0% do máximo	90,0% do máximo	95,0% do máximo
	04/11/2020	20/12/2020	03/02/2021
Óbitos	80,0% do máximo	90,0% do máximo	95,0% do máximo
	22/09/2020	01/11/2020	10/12/2020

Fonte: Elaborado pelo autor, 2020.

de casos confirmados atingidos em fevereiro 2021 e 95,0% do máximo óbitos atingidos em dezembro 2020. As previsões deste trabalho devem ser vistas como projeções da evolução do número de contaminados e óbitos contabilizados pelo Ministério da Saúde. São valores aproximados em relação aos números reais que teremos no final da pandemia devido ao número limitado de testes executados no país e do desconhecimento do número de contaminados assintomáticos, ao longo de todo o processo da pandemia.

Determinar os números e o exato momento do término da Pandemia no Brasil é muito difícil, tendo em vista o desenvolvimento da transmissão do vírus variar de acordo com a eficiência das medidas de distanciamento social, a capacidade de testagem e de atendimento do sistema de saúde. Essa incerteza quanto à evolução da pandemia de COVID-19 torna ainda mais importante o aprimoramento da testagem e o acompanhamento do seu desenvolvimento de forma a dimensionar da melhor forma possível o seu impacto na sociedade.

## REFERÊNCIAS

1. World Health Organization - WHO. Coronavirus disease 2019 (COVID-19) situation report 1. Geneva: World Health Organization; 2020[acesso 17 ago 2020]. Disponível em: <https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200121-sitrep-1-2019-ncov.pdf>
2. World Health Organization - WHO. Coronavirus disease 2019 (COVID-19) situation report 51. Geneva: World Health Organization; 2020[acesso 17 ago 2020]. Disponível em: <https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200311-sitrep-51-covid-19.pdf>
3. Perez FJD, Chinarro D, Otin RP, Mouhaffel AG. Growth forecast of the COVID-19 with the gompertz function, case study: Italy, Spain, Hubei (China) and South Korea. *Int J Adv Eng Res Sci.* 2020;7(7):67-77.
4. Medina-Mendieta JF, Cortés M, Iglesias M. COVID-19 forecasts for Cuba using logistic regression and gompertz curves. *Medic Rev.* 2020;22(3):32-9. <https://doi.org/10.37757/MR2020.V22.N3.8>
5. Domingues JS. Análise do modelo de Gompertz no crescimento de tumores sólidos e inserção de um fator de tratamento. *Biomat.* 2011;21:103-12.
6. Dutra CM, Farias FM, Melo CAR. New approach of non-linear fitting to estimate the temporal trajectory of the COVID-19 cases. *Braz J Health Rev.* 2020;3(3):6341-56. <https://doi.org/10.34119/bjhrv3n3-186>
7. Hammer O, Harper DAT, Ryan PD. PAST: paleontological statistic software package for education and data analysis. *Palaent Eletr.* 2001;4(1):1-9.



8. Prado MF, Antunes BBP, Bastos LSS, Peres IT, Silva AAB, Dantas LF et al. Analysis of COVID-19 under-reporting

in Brazil. Rev Bras Ter Intensiva. 2020;32(2):224-8.  
<https://doi.org/10.5935/0103-507X.20200030>

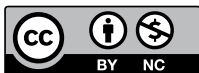
---

#### Contribuição do Autor

Dutra CM - Concepção, planejamento (desenho do estudo), aquisição, análise, interpretação dos dados e redação do trabalho. O autor aprovou a versão final do trabalho.

#### Conflito de Interesse

O autor informa não haver qualquer potencial conflito de interesse com pares e instituições, políticos ou financeiros deste estudo.



Licença CC BY-NC atribuição não comercial. Com essa licença é permitido acessar, baixar (download), copiar, imprimir, compartilhar, reutilizar e distribuir os artigos, desde que para uso não comercial e com a citação da fonte, conferindo os devidos créditos de autoria e menção à Visa em Debate. Nesses casos, nenhuma permissão é necessária por parte dos autores ou dos editores.