



A COVID-19 E OS FATORES QUE INFLUENCIAM A PROBABILIDADE DE MORTALIDADE: UMA ANÁLISE DE REGRESSÃO LOGÍSTICA COM DADOS DO SISTEMA DE ÚNICO DE SAÚDE DO BRASIL

Hugo Arruda de Albuquerque, Miriam Karla Rocha

Resumo: No final de 2019 foi descoberto um novo vírus altamente contagioso que se espalhou no mundo inteiro, a COVID-19. Ele causa diversos sintomas, em que o mais grave é a síndrome respiratória aguda grave, podendo levar o paciente a óbito. No Brasil, ele surgiu no início do ano de 2020 e vem se espalhando por todo o País. Como a doença vem crescendo em número de contaminados e vítimas, surgiu a necessidade de estudar os dados e características da doença, como ela pode reagir nos indivíduos, estudando de maneira probabilística as suas consequências. Procurando assim, trazer e entender como a doença pode reagir em cada pessoa de acordo com o seu sexo, idade, peso e outras características. Os resultados obtidos nessa pesquisa mostram que a equação alcançada através da regressão logística é condizente com a realidade do Brasil e com outras pesquisas publicadas, em que estas mostram as estatísticas com a covid-19 para cura ou óbito e suas correlações com as características dos pacientes. Pode-se com o resultado obtido, ter uma probabilidade de como a doença irá evoluir de acordo com as informações do paciente e sintomas manifestados pela covid-19. Foi possível identificar quais fatores contribuí mais para o óbito, como a idade mais avançada e quais doenças preexistentes afetam mais ou menos que outras. Atualmente, o País ainda vive uma crise epidemiológica, e saber trabalhar com os dados, é muito importante para saber qual estratégia pode-se tomar como forma de prevenção, já que outros problemas que a doença causa não é só na saúde pública, mas também na economia de um país que gera desempregos e queda de renda para as famílias.

Palavras-chave: Covid-19, Probabilidade, Regressão Logística, Pacientes, Óbitos.

1. INTRODUÇÃO

Segundo o Ministério da Saúde [1], a Covid-19, *corona virus disease*, é uma doença infecciosa causada pelo coronavírus, Sars-CoV-2. A maioria das pessoas que adoece tem sintomas leves e se recuperam sem tratamento especial. Ela se transmite por meio de gotículas geradas quando uma pessoa infectada tosse, espirra ou exala. Essas gotículas ficam no ar por um determinado tempo e se depositam sobre as superfícies. Qualquer pessoa pode se infectar ao inalar essas gotículas no ar ou se tiver contato através da superfície contaminada e em seguida colocar as mãos na boca, nariz ou olhos.

De acordo com Gruder [2], o primeiro caso surgiu em 2019 na cidade de Wuhan na China, onde se acredita que o 1º contágio deve ter ocorrido de um animal para o ser humano, uma vez que os primeiros casos confirmados foram detectados em um grupo de pessoas que estiveram em um mercado popular na cidade de Wuhan, e nesta cidade, eram vendidas várias espécies de animais selvagens vivos, como cobras, morcegos e roedores, que poderiam estar doentes e ter transmitido o vírus para as pessoas.

Após os primeiros casos de contágio em Wuhan, outras pessoas foram apresentando sintomas parecidos, mas que nunca estiveram no mercado da cidade em questão, apoiando a hipótese de que o vírus tenha se adaptado e estava sendo transmitido entre humanos, através das gotículas suspensas no ar, por meio da tosse, espirro ou exalação de ar. Logo, o vírus foi se espalhando pelos países inteiros e aos poucos, foram detectadas as formas em que ele se comportava, diante de cada cenário, como também, sendo detectados os principais sintomas, desde leves, onde não precisa de cuidados especiais, a casos mais graves, sendo necessário internação e auxílio de respiradores, uma vez que o vírus compromete o sistema respiratório, podendo evoluir para uma pneumonia, até levar o paciente a óbito.

Em 20 de janeiro de 2020, a Organização Mundial de Saúde (OMS) [3], classificou o surto como Emergência de Saúde Pública de Âmbito Internacional e, em 11 de março de 2020, como pandemia. Segundo a OMS [4], até 30 de março de 2021, 127.528.304 casos foram confirmados em 192 países e territórios, com 2.790.109 mortes atribuídas à doença, tornando-se uma das pandemias mais mortais da história.

De acordo com o Ministério da Saúde, no Brasil, já foram detectados mais de 12 milhões de casos e 314 mil mortes até 30 de março de 2021, quantidade esta que vem aumentando ao longo do tempo e que precisa ser controlado através de ações públicas como cuidados sanitários, distanciamento de pessoas doentes, vacinação, entre outros.

A doença não só trouxe problemas de saúde como também um abalo socioeconômico, o que fez aumentar ainda mais a desigualdade social e aumento no desemprego. De acordo com Silveira [5], muitas empresas decretaram falência devido ao isolamento social, ao atraso no ensino educacional para as crianças, devido às escolas estarem fechadas, apontado por Bota [6], o meio ambiente que devido a diminuição no consumo e queima de combustíveis houve uma melhora no ar e a diminuição da poluição de acordo com a *Agence France-Presse* (AFP) [7], eventos culturais foram interrompidos devido a aglomeração que poderia acelerar ainda mais a proliferação do vírus, e a saúde mental devido a soma de todos esses problemas relatados as pessoas ficaram mais em casa para evitar o contágio mas com a incerteza acabaram criando problemas mentais [8,9].

O presente trabalho tem como objetivo analisar os dados referentes a Covid-19 no Brasil, buscar entender como se manifesta em diferentes pessoas, de sexo e idade diferentes, bem como analisar quais doenças pré-existentes pode agravar o quadro clínico como; diabetes, asma, doença cardiovascular, obesidade e outras que

podem agravar ainda mais os sintomas da Covid-19. Com esses dados, pode-se identificar qual a maior probabilidade desses casos evoluírem a óbito ou internação.

A pesquisa foi um trabalho original e de revisão de literatura, na qual, Marconi e Lakatos [10], vêm dizer que a pesquisa, seja de que natureza for, é necessária para uma construção teórica e prática que permita ao pesquisador construir um trabalho científico, dando suporte teóricos e práticos para um aprofundamento do tema escolhido, dando ao pesquisador a possibilidade de escolher a técnica mais adequada ao seu projeto, seja ele, científico, ou de outra natureza, os descritores usados para a busca do material foram; Probabilidade. estatística, pacientes, óbitos.

2. REVISÃO DA LITERATURA

O estudo se baseia na retrospectiva e descritivo de características clínicas e epidemiológicas dos casos por covid-19 notificados e registrados pelo Data SUS no período de 03 de janeiro de 2020 até o dia 19 de abril de 2021 e utilizando o *software Jasp* para realizar a análise dos dados através da regressão logística. Foi escolhido o método de análise a regressão logística binária, uma vez que ela traz uma probabilidade para duas respostas, no caso, cura ou óbito, para os casos de covid-19, de acordo com as variáveis independentes onde pode influenciar o resultado.

2.1. A covid-19 e os fatores de risco

Segundo dados apresentados pela Fundação Oswaldo Cruz (Fiocruz), vários estudos apontam que a influência das comorbidades nos óbitos de pacientes que foram contagiados pelo coronavírus chega a uma média de 56,6%, a mesma pesquisa ainda indica que a desigualdade social também influencia no tratamento dos pacientes [11].

Corroborando com estes dados, o Ministério da saúde, aponta que a maioria das mortes pela covid-19 pode ser associada a doenças preexistentes, como; diabetes, coração, hipertensão entre outras, e, aponta que por ser uma doença que ataca o sistema imunológico, o risco de um paciente vir a óbito em relação a um paciente considerado saudável é bem maior. Para resguardar estes pacientes é que a vacinação foi dada aos mesmo em critério de prioridade [1].

De acordo com a OMS, outro fator além de pacientes com doenças preexistente é a questão da idade, uma vez que é do conhecimento de todos que quando mais velho o ser humano seja, menos defesa os organismos têm se comparado com um outro mais jovem. A probabilidade de uma pessoa de 84 anos vir a falecer de covid-19 em proporção a um jovem saudável é de 95%.

De acordo com os registros do Governo do Estado de Santa Catarina, existe uma associação dos óbitos pela COVID-19, na Tabela 1, com comorbidades preexistentes dos pacientes [12].

Tabela 1. Comorbidades associadas aos óbitos por COVID-19 no período compreendido entre 01 de janeiro de 2020 a 15 de agosto de 2020, em Santa Catarina, Brasil.

VARIÁVEL	N (%)
Comorbidades	1452 (82,17)
Doença Cardiovascular Crônica	876 (49,58)

VARIÁVEL	N (%)
Diabetes	626 (35,43)
Hipertensão	461 (26,09)
Doença Pulmonar Crônica	225 (12,73)
Obesidade	208 (11,77)
Doença Neurológica Crônica	158 (8,94)
Doença Renal Crônica	115 (6,51)
Imunodepressão	112 (6,34)
Asma	63 (3,57)
Câncer	58 (3,28)
Doença Hepática Crônica	33 (1,87)
Doença Hematológica Crônica	20 (1,13)
Síndrome de Down	3 (0,17)
Sem Comorbidades	315 (17,83)

Na Tabela 2, mostra os sintomas associados aos óbitos pela COVID-19 e sua frequência registrados em um período em Santa Catarina [12].

Tabela 2. Prevalência dos sintomas associados aos óbitos por COVID-19 no período compreendido entre 01 de janeiro de 2020 a 15 de agosto de 2020, em Santa Catarina, Brasil. (Ref. Governo de Santa Catarina)

SINTOMAS	N (%)
Dispneia	1.480 (83,76)
Desconforto Respiratório	1.289 (72,95)
Queda na Saturação de O ₂	1.273 (72,04)
Tosse	1.251 (70,80)
Febre	1.103 (62,42)
Dor de Garganta	330 (18,68)
Diarreia	267 (15,11)
Mialgia	127 (7,19)
Vômito	115 (6,51)
Cefaleia	99 (5,60)
Cansaço	61 (3,45)
Coriza	44 (2,49)
Dor no Corpo	31 (1,75)

2.2. Regressão logística

A regressão logística é uma forma de regressão que é formulada para prever e explicar uma variável não-métrica e com duas ou mais categorias por meio de variáveis independentes métricas ou não-métricas. O objetivo da regressão logística é gerar uma função matemática cuja resposta possa estabelecer a probabilidade de uma observação pertencer a um determinado grupo previamente determinado, em razão do comportamento das variáveis independentes. Na medicina, procura-se investigar se a pessoa tem ou não tem determinada doença, ou quais as chances de cura daquele paciente, se ele deve seguir um tratamento ou melhor seria partir para uma intervenção maior como uma cirurgia de acordo com preditores como idade, sexo, sintomas ou doenças pré-existent. Nesses casos, faremos o uso da regressão logística binária, onde estamos interessados em somente conhecer se determinado evento pode ocorrer ou não. Poderíamos também categorizar em função das variáveis,

estimando a probabilidade para cada grupo, neste caso, a variável dependente também é qualitativa, neste caso, deve-se utilizar da regressão logística multinomial. De acordo com Fávero e Belfiore [13], mostrarei de forma dedutiva como chegar em uma equação matemática para a regressão logística, e, conseqüentemente, a probabilidade de ocorrência de um fato.

Então, é possível prever ou saber qual a probabilidade de um fato ocorrer através da regressão logística. Assim, também pode ser utilizado em outras áreas como no Marketing, por exemplo, qual a probabilidade de vender um carro novo no fim ou início do ano, na Educação, para saber se um aluno é aprovado ou não em determinado teste, ou em qualquer área em que se deseja uma probabilidade de um fato ocorra.

Desta forma, pode-se escrever em forma de vetor as variáveis explicativas de acordo com a Equação (1):

$$Z_i = \alpha + \beta_1 \cdot x_{1i} + \beta_2 \cdot x_{2i} + \dots + \beta_k \cdot x_{ki} \quad (1)$$

Em que Z é conhecido com *logito*, α é a constante, β_j ($j = 1, 2, \dots, k$), são os parâmetros de cada variável explicativa, x_i são as variáveis explicativas (métricas ou binárias) e o subscrito i representa cada observação de amostra ($i = 1, 2, \dots, n$, em que n é o tamanho da amostra). Sabendo que Z não representa a variável dependente e o objeto é determinar a probabilidade p_i de ocorrência do evento em função dos parâmetros estimados de cada variável explicativa, deve-se definir o conceito de chance de ocorrência de um evento, conhecido como odds, ilustrada na Equação (2):

$$Chance(odds)_{Y_{i=1}} = \frac{p_i}{1 - p_i} \quad (2)$$

Supõe-se que, por exemplo, a probabilidade de cura de uma doença seja de 90%, a sua chance de cura é de 9 para 1 ($0,9 / 0,1 = 9$). Se a probabilidade para outra doença mais mortal for de 50%, a sua chance de cura é de 1 para 1 ($0,5 / 0,5 = 1$). Tem-se então que o conceito para chance é totalmente diferente de probabilidade, apesar de pensar que são sinônimos.

A regressão logística binária define o logito de Z como o logaritmo natural da chance, ilustrado na Equação (3):

$$\ln(chance Y_i = 1) = Z_i \quad (3)$$

De onde vem a Equação (4):

$$\ln\left(\frac{p_i}{1 - p_i}\right) = Z_i \quad (4)$$

Definindo a expressão para a probabilidade de ocorrência do evento em função do logito, pode-se isolar o p_i , a partir da Expressão (4), da seguinte maneira nas Equações (5), (6) e (7):

$$\frac{p_i}{1 - p_i} = e^{z_i} \quad (5)$$

$$p_i = (1 - p_i) \cdot e^{z_i} \quad (6)$$

$$p_i \cdot (1 + e^{z_i}) = e^{z_i} \quad (7)$$

E, finalmente, se tem:

Probabilidade de ocorrência do evento, Equação (8):

$$p_i = \frac{e^{z_i}}{1 + e^{z_i}} = \frac{1}{1 + e^{-z_i}} \quad (8)$$

Probabilidade de ocorrência do não evento, Equação (9):

$$1 - p_i = 1 - \frac{e^{z_i}}{1 + e^{z_i}} = \frac{1}{1 + e^{z_i}} \quad (9)$$

A partir das Equações (1) e (8), pode-se escrever a equação geral da probabilidade estimada de ocorrência de um evento que se apresenta na forma dicotômica para uma observação i da seguinte forma na Equação (10):

$$p_i = \frac{1}{1 + e^{-(\alpha + \beta_1 \cdot X_{1i} + \beta_2 \cdot X_{2i} + \dots + \beta_k \cdot X_{ki})}} \quad (10)$$

O que se tem é a probabilidade de fato ocorrer para cada observação, de acordo com os valores atribuídos para cada variável independente.

3. MATERIAIS E MÉTODOS

A análise foi realizada a partir de um banco de dados usando a técnica de regressão logística, que é utilizada em pesquisas em que, a partir de um conjunto de observações, se deseja obter um modelo matemático possível de mostrar a probabilidade da ocorrência de um resultado esperado em decorrência de um conjunto de variáveis independentes, que no caso estudado, seria a cura ou óbito do paciente com COVID-19. Logo, seria a probabilidade de um fato ocorrer a partir das informações obtidas de dados coletados anteriormente e modelado em uma expressão matemática.

3.1. Coleta de dados

O Ministério da Saúde, por meio da Secretaria de Vigilância em Saúde acompanha os casos da Síndrome Respiratória Aguda Grave no Brasil, desde a pandemia de Influenza A (H1N1) em 2009, e desde então a vigilância vem acompanhando casos de influenza e outros vírus respiratórios, que anteriormente atuava apenas com a sentinela de Síndrome Gripal [14].

Em 2020, a vigilância da COVID-19, infecção causada pelo novo coronavírus, foi incorporada na rede de vigilância da Influenza e outros vírus respiratórios.

Os casos registrados são inseridos no Sistema de Informação da Vigilância Epidemiológica da Gripe (Sivepe-Gripe) do Governo Brasileiro, que foi implementado para registro de casos de Síndrome Respiratória e são utilizados para registrar novos casos do Coronavírus. No Guia de Vigilância Epidemiológica Emergência de Saúde Pública de Importância Nacional pela Doença pelo Coronavírus 2019 estão disponíveis informações sobre definições de casos, critérios de confirmação e encerramento dos casos, dentre outros.

Os dados são sujeitos a alterações decorrente a investigação pela equipe de vigilância epidemiológica que desenvolve o serviço nas três esferas de gestão.

3.2. Tratamento dos dados

O Banco de Dados possui 31 variáveis, para otimizar a seleção das informações para os resultados obtidos. Foi utilizado o período de 01 de fevereiro de 2020 até 19 de abril de 2021, no qual possui mais de 605 mil registros. Esse banco de informações é extraído no formato CSV (separado por vírgulas) e pode ser utilizado para editar no Microsoft Excel ou qualquer software de planilha eletrônica.

As informações selecionadas são: idade, sexo, gestante, febre, tosse, dor de garganta, dispnéia, desconforto respiratório, saturação de oxigênio menor que 95%, diarreia, vômito, dor abdominais, fadiga, perda de olfato e/ou paladar, fator de risco, puerpéra, doença cardiovascular crônica, doença hematológica crônica, síndrome de down, doença hepática crônica, asma, diabetes, doença neurológica crônica, pneumopatia crônica, imunodeficiência, doença renal crônica, obesidade, internação, internação em UTI e uso de suporte ventilatório.

As técnicas estatísticas empregadas no presente estudo foram a análise descritiva e as análises de regressão logística simples e múltipla.

3.2.1. Caracterização do estudo

O estudo compreende os registros de pessoas com casos de Síndrome Respiratória Aguda Grave no Brasil obtidos através do Sistema de Informação da Vigilância Epidemiológica da Gripe (Sivepe-Gripe) do Governo Federal Brasileiro, que são utilizados também para registrar casos de Coronavírus.

Foram considerados nesse estudo 606.043 registros no período de fevereiro de 2020 a abril de 2021, permitindo que as probabilidades de morte pudessem ser calculadas em função da presença ou ausência de alguns fatores de risco e sintomas apresentados.

3.2.2. Obtenção dos dados

A identificação para os casos foi feita através de testes RT-PCR positivos para a COVID-19 registrados no sistema do DATA-SUS, extraído em formato de planilha eletrônica, que foi realizado o tratamento dos dados no *Microsoft Excel* e exportado em formato .csv para leitura no programa JASP.

3.2.3. Ajustes no banco de dados

Foram considerados no estudo somente casos positivos para COVID-19. Na seção 3.2 pode ser observado as variáveis selecionadas, onde para facilitar o estudo, foram considerado dois resultados possíveis: sim (1) e não (0) para não, tornando as variáveis dicotômicas, com exceção para a idade dos pacientes.

3.2.4. Variável dependente

Neste estudo a variável dependente (desfecho dicotômico) é a ocorrência ou não de óbito em pacientes infectados pela COVID-19.

3.2.5. Variáveis independentes ou covariáveis

Foram consideradas como covariáveis aquelas com potenciais de serem fatores de riscos associados à morte do paciente acometido pelo Coronavírus sendo escolhido como: sexo, idade, sintomas e comorbidades preexistentes.

Informações irrelevantes que não são objetos de estudo e não iria alterar o resultado, foram dispensadas como município, código IBGE do município, unidade de saúde de atendimento, escolaridade, CEP, se o paciente recentemente fez viagem internacional, quais exames foram realizados, os tipos de teste para COVID-19, entre outras informações julgadas indiferentes para o resultado da pesquisa.

As covariáveis relativa às doenças preexistentes do paciente bem como os sintomas apresentados, são dicotômicas, sendo as respostas “sim” para existentes e “não” para ausentes.

3.2.6. Modelagem de dados

Para definir a Equação (12) foi necessário fazer o tratamento das informações da planilha para facilitar o seu processamento no programa que irá fazer a regressão logística, como mostra na Figura 1. Foi necessário filtrar somente os casos com covid-19 confirmados e padronizar as informações para 0, sendo “não”, e 1 para “sim”, nas variáveis independentes contidas na planilha eletrônica.

Após a adequação, o arquivo é salvo em formato .csv (separado por vírgulas). Esse arquivo é importado para o *software* livre JASP, onde pode fazer diversas análises. Ao ser importado, o programa faz a análise dos dados em regressão logística binária.



Figura 1. Organograma modelagem. (Autoria própria)

4. RESULTADOS

A fim de definir o modelo mais adequado para os dados, as seguintes etapas foram estabelecidas para testar e eliminar variáveis estatisticamente insignificantes, mais comum em pesquisas de ciências sociais [15]:

1º) Seleção das variáveis candidatas ao modelo de regressão múltipla: foi realizada a análise de regressão logística uni variada de cada variável independente com o desfecho. Para a seleção das variáveis que entrarão no modelo, pode-se utilizar o método *enter*. A covariável que apresentou o *p*-valor menor ou igual a 25% no teste uni variado foi incluída no modelo múltiplo, ou seja, se $p\text{-valor} \leq 0,25$, a covariável foi significativa e houve evidência de associação desta covariável com o desfecho. Se $p\text{-valor} > 0,25$, a covariável foi desconsiderada para ser testada juntamente com as outras no modelo múltiplo. A regressão logística uni variada possibilitou a estimação das razões de chance (OR bruto) e os respectivos intervalos de confiança de 95% (IC 95%);

2º) Para a análise de regressão logística múltipla, foi utilizada a estratégia de análise não-condicional que incluiu, simultaneamente no modelo, todas as covariáveis independentes significativas da análise de regressão logística uni variada, método *enter* e nível de significância de 0,25, verificados pelo teste de *p*-valor;

3º) Foram eliminadas do modelo múltiplo, uma a uma, as covariáveis não significativas ($p\text{-valor} > 0,05$), verificando-se o efeito da saída de cada covariável nos coeficientes (β_i) que permaneciam no modelo. A cada retirada de uma covariável (com maior valor de *p*), o modelo era processado para que os coeficientes fossem reajustados para aquele novo conjunto de covariáveis.

As análises estatísticas foram realizadas com o auxílio do aplicativo computacional JASP 0.14.1.

Com a modelagem foi retirada duas covariáveis por apresentar $p > 0,05$ no teste multivariado (Puérpera e dor abdominal).

Os coeficientes (β_i 's) da equação correspondem aos valores do ln (OR ajustado) da Tabela 3, para cada categoria das covariáveis é descrito na Equação (11).

$$\begin{aligned} \ln \text{odds} = & 3,589 - 0,051 * X_{\text{idade}} - 0,196 * X_{\text{sexo}} + 0,249 * X_{\text{gestante}} + 0,059 * X_{\text{febre}} + 0,266 * \\ & X_{\text{tosse}} + 0,043 * X_{\text{garganta}} - 0,24 * X_{\text{dispneia}} - 0,248 * X_{\text{desc_respiratorio}} - 0,155 * X_{\text{saturação}} + \\ & 0,158 * X_{\text{cardiopatia}} + 0,286 * X_{\text{asma}} - 0,085 * X_{\text{diabetes}} - 0,205 * X_{\text{pneumotapia}} - 0,091 * X_{\text{obesidade}} \\ & - 0,662 * X_{\text{UTI}} - 1,974 * X_{\text{ventilação}} + 1,410 * X_{\text{hospital}} - 0,232 * X_{\text{fatorrisco}} + 0,181 * X_{\text{diarreia}} + \\ & 0,305 * X_{\text{fadiga}} + 0,148 * X_{\text{perd_olfato}} + 0,253 * X_{\text{perd_palad}} - 0,218 * X_{\text{hematologica}} - \\ & 0,419 * X_{\text{sind_down}} - 0,518 * X_{\text{hepatica}} - 0,588 * X_{\text{imunodepre}} - 0,536 * X_{\text{renal}} \end{aligned} \quad (11)$$

Aplicando a equação em situações hipotéticas podemos estimar a probabilidade de ocorrência de óbito em um paciente infectado pela COVID-19.

- 1) Qual a probabilidade de cura para um idoso do sexo masculino com 80 anos de vida com comorbidades de doenças renais e que precisou ser hospitalizado?

$$\ln \text{odds} = 3,589 - 0,051*(80) - 0,196*(1) - 0,536*(1) + 1,410*(1)$$

$$\ln \text{odds} = 0,187$$

$$P\left(\frac{Y}{X}\right) = \frac{1}{1 + e^{-(\ln \text{odds})}}$$

$$P\left(\frac{Y}{X}\right) = \frac{1}{1 + e^{-(0,187)}}$$

$$P\left(\frac{Y}{X}\right) = 0,546 = 54,6\%$$

- 2) Qual a probabilidade de cura para um adolescente do sexo feminino com 16 anos de vida sem comorbidades onde apresentou sintomas leves como febre, tosse, dor de garganta e não precisou ser hospitalizado?

$$\ln \text{odds} = 3,589 - 0,051 * \text{Xidade} + 0,059 * \text{Xfebre} - 0,196 * \text{Xsexo} + 0,266 * \text{Xtosse} + 0,043 * \text{Xgarganta}$$

$$\ln \text{odds} = 3,589 - 0,051 * 1 + 0,059 * 1 - 0,196 * 0 + 0,266 * 1 + 0,043 * 1$$

$$P\left(\frac{Y}{X}\right) = \frac{1}{1 + e^{-(3,906)}}$$

$$P\left(\frac{Y}{X}\right) = 0,9585 = 95,85\%$$

Portanto, corroborando com os estudos apresentados durante a construção deste artigo, a chance de uma pessoa com 80 anos com doenças preexistente ser curada em decorrência de ser positivo com covid-19 em estado grave é de 54,6%. O que atesta que para pessoas com mais idades e com comorbidades são fatores que contribuem para o óbito do enfermo. Em outro caso, pessoas mais jovens e saudáveis estão menos suscetíveis ao óbito, como na segunda simulação, no qual, a pessoa infectada sem comorbidades e sem sintomas graves da doença tem uma probabilidade grande de cura, em torno de 96% de chance.

Pode-se também identificar as variáveis que mais contribuem para o desfecho da doença para óbito, como por exemplo quais comorbidades apresentam uma probabilidade maior de óbito e quem passa por uso de suporte respiratório é um agravante para o paciente.

Apesar de haver outras variáveis que essa pesquisa não engloba e que surgiram durante a pandemia como: aparição de novas variáveis, imunidade por anticorpos, vacinas, tratamentos experimentais etc., a equação se mostra bastante aplicável para a realidade brasileira. Podendo assim ser utilizada como fonte de estudo para probabilidade de evolução da doença nos seus casos mais variados e características do paciente dentre mais de 29 variáveis disponíveis.

5. CONSIDERAÇÕES FINAIS

Considera-se os objetivos propostos nesse estudo alcançados, pois o desenvolvimento desta pesquisa busca o entendimento da técnica aplicada de regressão logística na análise, bem como a compreensão dos procedimentos de cálculo envolvido, assim como a identificação das variáveis que influenciam para óbito dos pacientes da covid-19.

É importante considerar que o resultado obtido, foi através de um banco de dados oficial, com acesso livre e gratuito, é de suma importância para realização de estudos científicos que tenha como objetivo aprofundar e identificar outros fatores que aqui não foram identificados e de outras doenças de relevância para a saúde pública.

Há fatores que devemos levar em consideração nessa pesquisa, como o avanço da vacinação de diversas fabricantes e tipos diferentes, houve uma queda na mortalidade, principalmente nas pessoas mais idosas que foram

as primeiras a serem vacinadas, mas que não influencia tanto no resultado já que foram considerados dados até 19 de abril de 2021, onde a vacinação da população estava em apenas 4,78% da população brasileira [16]. Existe também as cepas do vírus que vão mutando e a taxa de reinfecção pelo coronavírus, que pode criar anticorpos na população. E por isso, pode estar imune ou não, o que merece também ser mais aprofundado.

Outros estudos seriam necessários para avaliar a nova taxa de mortalidade após a vacinação completa da população, bem como para buscar o melhor planejamento do sistema de saúde para que não surja novas epidemias e a volta das atividades normais que foram afetadas devido as ações emergenciais tomadas para evitar o contágio em massa da doença, que por ser um vírus ainda desconhecido no início da epidemia, não se sabia na época, como evitar o contágio e como tratar as pessoas infectadas.

Portanto, este trabalho pode ser utilizado como referência para futuras pesquisas na área das ciências da saúde, podendo também auxiliar e subsidiar outros trabalhos que utilizem a análise de regressão logística.

6. REFERÊNCIAS

- [1] Ministério da Saúde. Disponível em: <https://coronavirus.saude.gov.br/sobre-a-doenca>. Acesso em 05/05/21.
- [2] GRUBER, Arthur. **Covid-19: o que se sabe sobre a origem da doença**. Disponível em: <https://jornal.usp.br/?p=314416>. Acesso em 05/05/21.
- [3] WHO Director-General's opening remarks at the media briefing on COVID-19—11 March 2020» (em inglês). Organização Mundial da Saúde. 11 de março de 2020. Disponível em: <https://www.who.int/dg/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020>. Acesso em 05/05/2021.
- [4] Worldometer. Disponível em: <https://www.worldometers.info/coronavirus/>. Acesso em 30/03/21.
- [5] SILVEIRA, Daniel. **Desemprego diante da pandemia bate recorde no Brasil em setembro, aponta IBGE**. Portal G1. Disponível em: <https://g1.globo.com/economia/noticia/2020/10/23/no-de-desempregados-diante-da-pandemia-aumentou-em-34-milhoes-em-cinco-meses-aponta-ibge.ghtml>. Acesso em 05/05/2021.
- [6] BOTO, Carlota. **A educação e a escola em tempos de coronavírus**. Jornal da USP. Disponível em: <https://jornal.usp.br/?p=313176>. Acesso em 05/05/2021.
- [7] AFP. **O ar está mais puro nos países confinados por coronavírus**. Revista Exame. Disponível em: <https://exame.com/ciencia/o-ar-esta-mais-puro-nos-paises-confinados-por-coronavirus/>. Acesso em 05/05/2021.
- [8] MAGENTA, Matheus. **Quarentenas funcionam para combater o coronavírus? Veja o que dizem os estudos**. BBC News Brasil. Disponível em: <https://www.bbc.com/portuguese/internacional-52830618>. Acesso em 05/05/2021.
- [9] ORGAZ, Cristina J. **Coronavírus: o efeito psicológico da quarentena forçada para tentar conter a covid-19**. BBC News Mundo. Disponível em: <https://www.bbc.com/portuguese/geral-51800944>. Acesso em 05/05/2021.
- [10] MARCONI, M. A.; LAKATOS, E. M. **Fundamentos da metodologia científica**. 6. ed. São Paulo: Atlas, 2009
- [11] Pauline Almeida. **Fiocruz relata perfil das vítimas da Covid-19 e desigualdades na pandemia**. Disponível em: <https://www.cnnbrasil.com.br/saude/pesquisa-da-fiocruz-relata-perfil-das-vitimas-da-covid-19-e-desigualdades-no-ini/>. Acesso em 30/03/2021.
- [12] Governo do Estado de Santa Catarina. **Coronavírus em SC**. Disponível em: <https://www.coronavirus.sc.gov.br/>. Acesso em 30/03/2021.
- [13] FÁVERO, Luiz Paulo; BELFIORI, Patrícia. **Manual de análise de dados**. Elsevier Editora Ltda. Rio de Janeiro. 1ª Edição, 2017. p.611 – 614.
- [14] OpenDataSUS. **SRAG 2020 - Banco de Dados de Síndrome Respiratória Aguda Grave - incluindo dados da COVID-19**. Disponível em: <https://opendatasus.saude.gov.br/dataset/bd-srag-2020>. Acesso em 30/04/2021.
- [15] MEZZOMO, Meire. **Estudo da mortalidade infantil – um modelo de Regressão logística múltipla**. p.37. Monografia (Curso de Especialização em Estatística e Modelagem Quantitativa) – Universidade Federal de Santa Maria, Santa Maria, RS, Brasil, 2000. Disponível em: https://repositorio.ufsm.br/bitstream/handle/1/2992/Mezzomo_Meire.pdf. Acesso em: 30/04/2021.
- [16] G1. **Brasil aplicou ao menos uma dose de vacina contra covid em 26,6 milhões de pessoas, aponta consórcio de veículos de imprensa**. Disponível em: <https://g1.globo.com/bemestar/vacina/noticia/2021/04/19/brasil-aplicou-ao-menos-uma-dose-de-vacina-contracovid-em-266-milhoes-de-pessoas-aponta-consorcio-de-veiculos-de-imprensa.ghtml>. Acesso em : 21/10/2021.