

**PROCEDIMENTOS PARA ESTIMATIVA DA VELOCIDADE DE DISSEMINAÇÃO DE DOENÇAS –  
UM ESTUDO DE CASO COM DADOS DA COVID-19 NO ESTADO DO PARANÁ**

**PROCEDURES FOR ESTIMATING THE SPREAD VELOCITY OF DISEASES - A CASE STUDY  
WITH COVID-19 DATA IN THE STATE OF PARANÁ**

**Tony Vinicius Moreira Sampaio**

Universidade Federal do Paraná, PR, Brasil  
[tony2sampaio@gmail.com](mailto:tony2sampaio@gmail.com)

**Emanuele Cristina Gustani Buss**

Universidade Estadual do Centro-Oeste, PR, Brasil  
[manugustani@gmail.com](mailto:manugustani@gmail.com)

**Carla Luiza da Silva**

Universidade Estadual de Ponta Grossa, PR, Brasil  
[clsilva.2187@gmail.com](mailto:clsilva.2187@gmail.com)

**Daniela Frizon Alfieri**

Universidade Estadual de Londrina, PR, Brasil  
[frizon.alfieri@uel.br](mailto:frizon.alfieri@uel.br)

**Glauco Nonose Negrão**

Universidade Estadual do Centro-Oeste, PR, Brasil  
[gnegrao@unicentro.br](mailto:gnegrao@unicentro.br)

**Marcos Aurélio Pelegrina**

Universidade Estadual do Centro-Oeste, PR, Brasil  
[marcospelegrina@gmail.com](mailto:marcospelegrina@gmail.com)

**Pedro Augusto Breda Fontão**

Universidade Federal do Paraná, PR, Brasil  
[pedrofontao@ufpr.br](mailto:pedrofontao@ufpr.br)

**RESUMO**

Caracterizar e quantificar a velocidade com a qual uma doença se espalha em uma região é fundamental para entender e comparar o processo de disseminação em diferentes localidades e intervalos de tempo. Durante a pandemia de COVID-19, pesquisadores do Brasil e do mundo engendraram esforços para entender e caracterizar a dinâmica de contaminação e espalhamento da doença. A presente pesquisa tem como objetivo analisar diferentes possibilidades de caracterização da velocidade de surgimento de novos casos e dispersão do COVID-19 no estado do Paraná. A pesquisa utilizou dados de 205328 casos (teste RT-PCR) de 396 municípios, relativos a 244 dias do período pandêmico para cálculo da média móvel, de seu valor comparado com o observado 14 dias antes (viés) e de outros cinco diferentes procedimentos de cálculo baseados na inclinação da reta que caracteriza a frequência acumulada de ocorrência de novos casos a cada mil habitantes (valor de beta, inclinação/tangente do período analisado, máximo valor de inclinação observado nos intervalos quintílico e quartílico e a soma dos valores de inclinação diários observados). A pesquisa identificou que no Paraná o município que apresentou maior velocidade de espalhamento do COVID-19 foi o de Itambé e que 37 localidades podem ser caracterizadas como com ocorrência de surto para o período analisado. Ainda, identificou que os valores de beta e inclinação/tangente do período analisado podem ser utilizados para caracterizar a velocidade de dispersão da doença em escala municipal, bem como que valores superiores a 0,15 e 0,3 podem ser utilizados como indicativos de surto para intervalos maiores de tempo e para a semana epidemiológica, respectivamente.

**Palavras-chave:** Epidemiologia. Surto. COVID-19.

## ABSTRACT

Characterizing and quantifying the speed at which a disease spreads in a region is crucial for understanding and comparing the dissemination process in different locations and time intervals. During the COVID-19 pandemic, researchers from Brazil and around the world made efforts to understand and characterize the dynamics of infection and the spread of the disease. The present study aims to analyze different methods of characterizing the speed of new case emergence and COVID-19 spread in the state of Paraná. The research used data from 205,328 cases (RT-PCR tests) in 396 municipalities over a 244-day pandemic period for calculating the moving average, its value compared to what was observed 14 days earlier (bias), and five different calculation procedures based on the slope of the line characterizing the cumulative frequency of new cases per thousand inhabitants (beta value, slope/tangent of the analyzed period, maximum observed slope value in quintile and quartile intervals, and the sum of observed daily slope values). The research identified that in Paraná, the municipality with the highest COVID-19 spread rate was Itambé, and 37 locations can be characterized as having an outbreak during the analyzed period. Furthermore, it found that the beta values and the slope/tangent of the analyzed period can be used to characterize the disease's spread rate at the municipal level, and values exceeding 0.15 and 0.3 can be used as indicators of outbreaks for longer time intervals and for the epidemiological week, respectively.

**Keywords:** Epidemiology. Outbreak. COVID-19.

## INTRODUÇÃO

A velocidade com a qual uma doença se espalha em uma região está associada a diferentes fatores como a exemplo as características da população (ALBERCA et al., 2021; SEYEDALINAGHI et al., 2021) e a adoção ou não de medidas de restrição (SILVA; FILHO; FERNANDES, 2020; WAMOYI; RANGANATHAN; STÖCKL, 2021). Caracterizar essa velocidade é importante, pois permite entender e comparar o processo em diferentes localidades e intervalos de tempo.

A rápida propagação da síndrome aguda respiratória (SARS-CoV) ou COVID-19, impulsionada pela mobilidade, falta de conhecimento e forma de disseminação do vírus, resultou em uma pandemia global, declarada pela Organização Mundial da Saúde (OMS) em março de 2020 e em milhões de mortes em todo o mundo (GRISÓTTI, 2020).

Amplamente utilizada para caracterização de epidemias e pandemias, a Média Móvel (MM) fornece para cada dia de ocorrência da doença um valor de referência que reflete a intensidade de novos casos. Por ser uma média, atenua flutuações decorrentes de fatores diversos observados nos registros diários (MEROW; URBAN, 2020). A comparação do valor diário da MM com o observado 14 dias antes é comumente utilizado como indicativo de “viés” diário da doença (alta, redução ou estabilidade).

A MM pode ser calculada com diferentes intervalos de tempo (7, 14, 28 e 49 dias), os quais afetam a percepção do fenômeno (JOSÉ et al., 2021; MEROW; URBAN, 2020; SADOWSKI et al., 2021). Também é utilizada em modelos de previsão de novos casos e óbitos, como a exemplo do ARMA (*Autoregressive Moving Average*) e do ARIMA (*Autoregressive Integrated Moving Average*) (LYNCH; GORE, 2021).

Outros valores amplamente utilizados para caracterizar a velocidade de espalhamento de doenças são o  $R(0)$ ,  $R(e)$  /  $R(t)$ . Enquanto o  $R(0)$  indica o número inicial de reprodução do vírus no início da pandemia, o  $R(e)$  /  $R(t)$  indicam o número efetivo em um determinado tempo (LÓPEZ-FLORES et al., 2021).

De forma mais específica, a inclinação das linhas que representam o aumento no número de novos casos pode ser interpretada (CLIFF; HAGGETT, 1982), como a taxa média de crescimento por unidade de tempo, ou seja, velocidade. Conforme os autores, em processos de mensuração e estimação por regressão linear o coeficiente angular (inclinação) está diretamente associado a intensidade com a qual se deu o processo de disseminação (LEIGHTNER; INOUE; MICHEAUX, 2021; MEDINA, 2020) e, portanto, quanto menor o intervalo de tempo para que uma localidade atinja uma determinada

quantidade de casos por habitante, maior a inclinação e mais rápido foi o processo de disseminação da doença neste local.

Tanto o coeficiente angular quanto a média móvel e  $R(t)$ , podem apresentar valores negativos e positivos ao longo do tempo. Assim, variações nos intervalos de tempo (de ocorrência da doença e de cálculo) e a presença de valores negativos e positivos dificultam a análise e comparação da disseminação de doenças em diferentes localidades a partir do uso direto desses dados (CHAUHAN; KUMAR; JAMDAGNI, 2020).

Outros procedimentos podem ser utilizados para caracterizar e representar o processo de espalhamento de doenças, como o uso de modelagem matemática associada a análise do comportamento populacional, modelos com integração com fractais, como o FDC Fractional Differential Calculus (TUAN; MOHAMMADI; REZAPOUR, 2020), modelos para simular a propagação (da COVID) em escala global e, finalmente, o número básico de reprodução ou  $R_0$  (DUTRA, 2020).

A multiplicidade de procedimentos e suas particularidades demonstram a relevância científica de se conhecer/calcular a velocidade de espalhamento e, os diferentes problemas apontados pelos autores citados revelam a dificuldade de aplicação e limitações metodológicas, como a exemplo, os demonstrados por (DELAMATER et al., 2019).

Cabe enfatizar que a maior parte dos modelos citados apresentam ênfase na projeção/simulação de cenários futuros em escala de país/estado ou global, sendo adequados para caracterização da velocidade ou comportamento momentâneo do processo de disseminação.

Considerando que o conhecimento e a caracterização da velocidade de espalhamento de doenças, em especial infectocontagiosas, são fundamentais para comparar, planejar e a adotar medidas diretas sobre locais e populações atingidas e susceptíveis, o objetivo desta pesquisa foi o de analisar diferentes procedimentos que podem ser utilizados para quantificar o processo de dispersão, bem como os efeitos das flutuações no número de casos observadas ao longo do tempo sobre esses procedimentos. Para tanto, a pesquisa utilizou dados dos casos de COVID-19 registrados no estado do Paraná no ano de 2020.

## MATERIAIS E MÉTODOS

### *Dados e procedimentos metodológicos utilizados*

Para a análise e caracterização da velocidade de disseminação da COVID-19 foram utilizados os dados relativos aos casos identificados via teste RT-PCR obtidos diretamente na base de dados da Secretaria de Estado da Saúde do Paraná (SESA-PR). Foram utilizados inicialmente 205328 casos relativos a 396 dos 399 municípios do Estado. O intervalo de tempo utilizado para coleta de dados para a pesquisa foi de 01 de março de 2020 até o dia 31 de outubro de 2020, perfazendo 244 dias.

Para evitar problemas decorrentes da análise de dados em intervalos reduzidos de tempo, associados ao registro esporádico de novos casos (como a exemplo o registro de dados acumulados por mais de 20 dias), os quais resultariam em abrupta inclinação e efeitos sobre o coeficiente angular, foram considerados somente os municípios que apresentavam mais de 60 casos, no período analisado, e com mais de 60 dias de série temporal.

A pesquisa analisou e comparou cinco procedimentos de cálculo que podem ser utilizados para caracterizar a velocidade de espalhamento da doença/vírus: valor de beta (obtido via Regressão Linear), inclinação/tangente do período analisado, máximo valor de inclinação observado nos intervalos quintílico e quartílico e a soma dos valores de inclinação diários observados. Todos os procedimentos são baseados na inclinação da reta que caracteriza a frequência acumulada de ocorrência de novos casos a cada mil habitantes e serão detalhados a seguir.

O Projeto apresenta aprovação do comitê de ética em pesquisa sob número CAAE 39108020.0.0000.0107, registrada na Plataforma Brasil – Comissão Nacional de Ética em Pesquisa (<https://plataformabrasil.saude.gov.br>).

### **Cálculo da velocidade de disseminação dos casos da COVID-19 – uso de dados em escala municipal**

Para realização dos cálculos os dados foram organizados e sistematizados por município possibilitando a caracterização, cálculo e representação gráfica da disseminação do vírus ao longo do tempo.

O número de casos a cada 1000 habitantes levou em consideração a população estimada para o ano de 2020 (conforme IBGE, 2020). Os dados (casos de COVID) foram organizados por data e frequência acumulada para obtenção da média móvel (7 dias) e demais procedimentos de cálculo. O uso do número de casos normalizado pelo número de habitantes, se justifica por permitir a comparação entre as diferentes localidades, uma vez que os municípios apresentam quantitativos populacionais distintos.

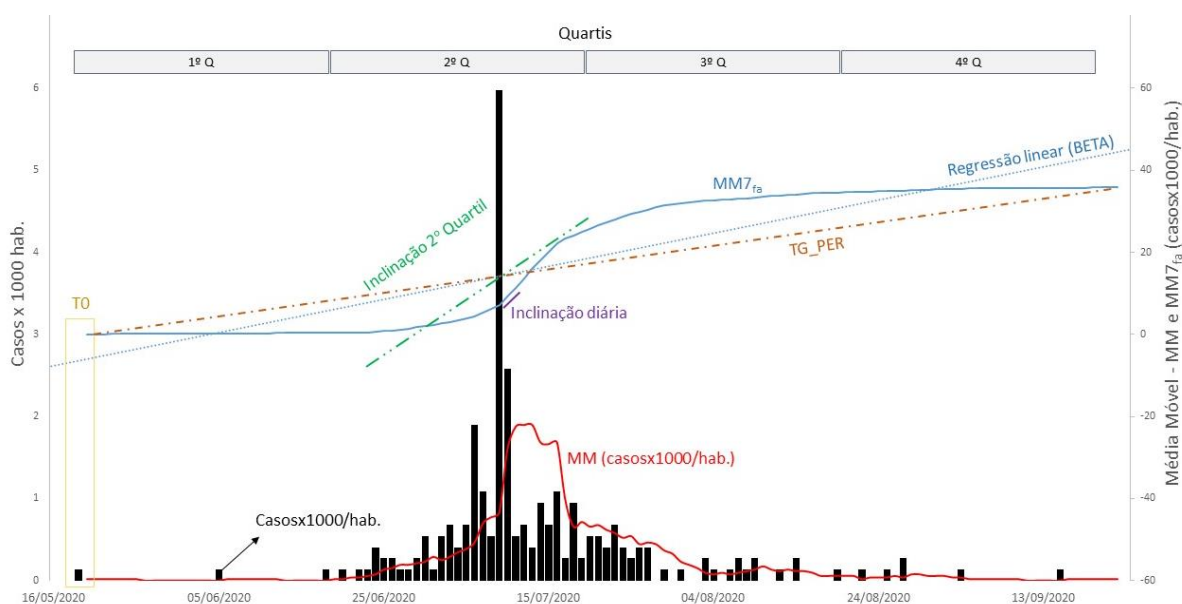
#### **Procedimentos utilizados no cálculo da velocidade de dispersão**

Além da obtenção e análise dos valores de viés (comparação dos valores de MM diários com o valor da MM de 14 dias antes), a velocidade de dispersão foi calculada por outros quatro procedimentos os quais objetivaram avaliar os efeitos de flutuações no número de casos ao longo do tempo sobre a estimativa da velocidade de disseminação da doença.

Os quatro procedimentos de cálculo foram baseados nos valores de inclinação e utilizam, como referência, os valores de média móvel (7 dias) a partir dos valores de frequência acumulada do número de novos casos (casos x 1000/hab.). O uso da média móvel calculada utilizando os valores de frequência acumulada do número de novos casos por habitante é utilizada também em estudos que objetivam identificar surtos (OVERTON et al., 2020) e se dá em função da necessidade de obtenção de valores positivos para uso nos procedimentos de cálculo (PINTO et al., 2022). Isto porque, o registro direto da média móvel do número diário de casos por habitante pode gerar tanto valores de inclinação positivos quanto negativos ao longo do tempo. A presença de valores de inclinação positivos (aceleração) e negativos (desaceleração) ao longo do tempo pode resultar na caracterização de locais com distintas velocidades de disseminação como sendo similares e vice-versa.

A Figura 1 a mostra a diferença observada entre o registro do número de novos casos (colunas - cor preto), em relação aos valores de média móvel (7 dias) de novos casos por habitante (linha vermelha) e média móvel da frequência acumulada de novos casos por habitante - MM7<sub>fa</sub> (linha azul).

Figura 1 – Procedimentos de cálculo baseados na inclinação



Fonte: Elaborado pelos autores.

Para cada município o período de análise adotou como T0 (tempo zero – Figura 1) a data do primeiro registro (RT-PCR) naquele município. A adoção do T0 objetivou permitir comparar os valores de inclinação entre as diferentes localidades pesquisadas, uma vez que o registro do primeiro caso de COVID-19 se deu em diferentes momentos nos municípios do Paraná.

O primeiro procedimento empregado baseado na inclinação das linhas que representam o número de novos casos foi o cálculo do coeficiente angular (BETA) via modelo de Regressão Linear (RLM – Regression Linear Model). Conforme apontado por Cliff & Haggett (CLIFF; HAGGETT, 1982) e Chauhan et al. (CHAUHAN; KUMAR; JAMDAGNI, 2020), a inclinação que indica o número de novos casos está associada a velocidade de disseminação. Enquanto inclinações abruptas indicam “processo de dispersão acelerado” (aumento acentuado no número de novos casos por habitante), segmentos de linhas horizontais ou levemente inclinados indicam períodos de estabilidade ou “relativa estabilidade”. A Figura 1 **Erro! Fonte de referência não encontrada.**(linha tracejada azul) ilustra a aplicação do método.

Regressões lineares apresentam como vantagens a possibilidade de cálculo para distintos intervalos de tempo, a obtenção dos valores de significância (*p-value*) e do coeficiente de determinação, os quais servem como indicadores da significância dos valores observados.

O segundo procedimento proposto para o estudo foi o cálculo direto da inclinação ou tangente do período de análise (TG\_PER). Esse procedimento possibilita a obtenção dos valores de inclinação de forma direta e simplificada, sem a necessidade de aplicação de regressões lineares. A TG\_PER (Figura 1 – linha marrom tracejada) foi calculada para cada município a partir da divisão do valor final da média móvel do número de casos (MM<sub>fa</sub>) observado no último dia do período de análise (intervalo correspondente a  $\Delta Y$  total), pelo número de dias de casos (correspondente a  $\Delta X$  ou intervalo de tempo de casos registrados).

$$TG\_PER = \frac{\Delta y}{\Delta x} \quad (1)$$

Sendo:

$\Delta y$  diferença entre o valor final e o valor inicial da MM<sub>fa</sub> e;

$\Delta x$  intervalo de tempo (número de dias) entre o registro do primeiro caso e o último do período em análise.

Com intuito de analisar o efeito das variações observadas na velocidade de surgimento e registro de novos casos ao longo do tempo sobre os valores de inclinação e sobre os valores BETA, outros três procedimentos de cálculo foram propostos e testados: inclinação máxima observada nos intervalos quartílicos (MAXQUARTIL) e quintílicos (MAXQUINTIL) e, inclinação diária acumulada (INCL\_ACUM). A inclinação máxima observada nos intervalos quartílicos e quintílicos foi calculada a partir da inclinação direta da reta que representa a frequência acumulada de novos casos em cada intervalo de tempo correspondente (quartílico ou quintílico) (na Figura 1 é representada pela linha tracejada verde).

O total acumulado dos valores diários de inclinação no período de monitoramento (INCL\_ACUM) foi calculado a partir da soma das inclinações diárias observadas (Figura 1 linha roxa). A inclinação diária corresponde a relação entre o número de novos casos por habitante do dia (MM<sub>7fa</sub>) e o valor de MM<sub>7fa</sub> do dia anterior (Figura 1 d).

$$INCL\_ACUM = \sum_{i=0}^n \frac{\Delta y}{\Delta x} \quad (2)$$

Sendo:

$\Delta y$  valor de MM<sub>7fa</sub> do dia menos o valor MM<sub>7fa</sub> do dia anterior

$\Delta x$  o intervalo de um dia

Os softwares utilizados para as análises foram o Microsoft Excel para organização e filtragem dos dados, bem como confecção de gráficos e os programas R Studio e Python (Jupyter Notebook) para tratamento estatístico e demais procedimentos de cálculo.

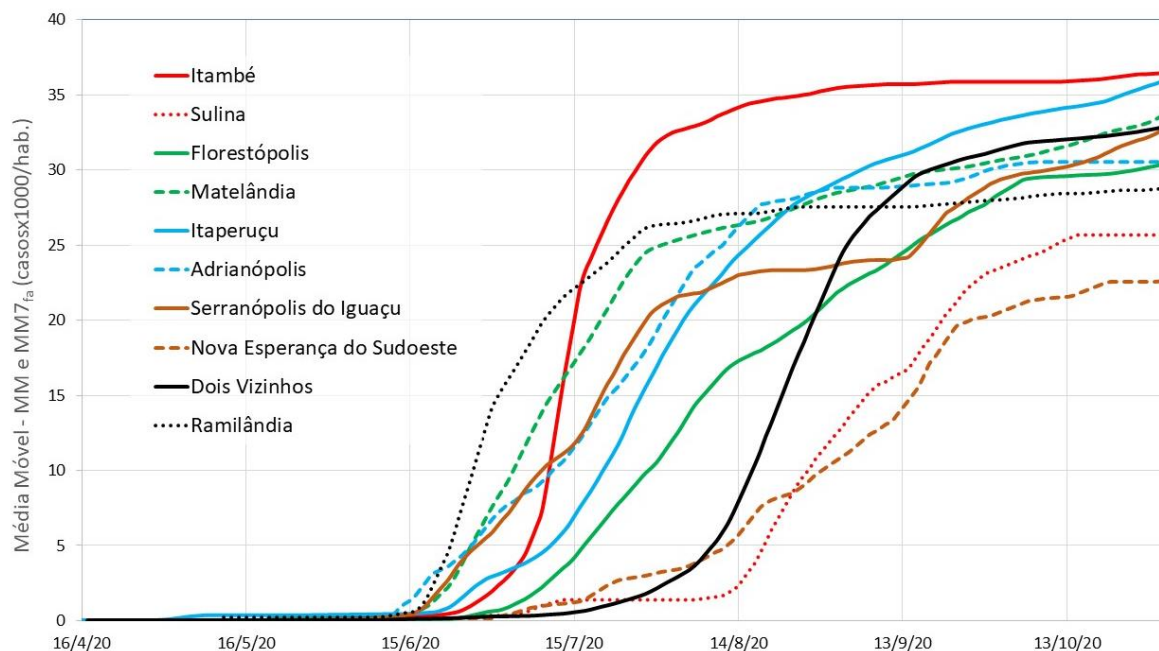
## RESULTADOS E DISCUSSÃO

### **Análise do coeficiente angular – BETA (velocidade de dispersão).**

Dos 399 municípios paranaenses, 235 (59%) apresentaram no período de estudo mais de 60 casos de COVID-19 registrados em período contínuo de mais de 60 dias, totalizando 200245 casos. Para estes, todas regressões lineares calculadas apresentaram *p-value* inferior à 0,001 e coeficientes de determinação ( $R^2$ ) acima de 0,685, indicando bom ajustamento em relação aos dados.

Os dez municípios com maiores valores de BETA para o período de análise foram Itambé (0,289), Sulina (0,266), Florestópolis (0,241), Matelândia (0,240), Itaperuçu (0,238), Adrianópolis (0,229), Serranópolis do Iguaçu (0,224), Nova Esperança do Sudoeste (0,216), Dois Vizinhos (0,214) e Ramilândia (0,201). As regressões lineares para estas localidades apresentaram coeficiente de determinação  $R^2$  entre 0,775 e 0,961.

Figura 2 – Municípios com maiores valores de BETA no estado do Paraná. Fonte dos dados: Secretaria de Estado da Saúde do Paraná (SESA-PR)



Fonte: Elaborado pelos autores.

Apesar de apresentarem diferentes datas para o início dos primeiros casos de COVID-19, os municípios com maiores valores de BETA (Figura 2) apresentaram comportamento similar em relação a forma de ocorrência de novos casos, sendo os momentos iniciais e finais com menores velocidades e o período intermediário (entre final de junho e final de setembro de 2020) o de maior inclinação. O intervalo de tempo (*delay*) entre o registro dos primeiros casos e o período de maior velocidade de dispersão nestes municípios foi de aproximadamente 30 dias.

A figura 2 mostra que entre junho e setembro de 2020 houve aceleração no número de novos casos de COVID-19 ampliando sua distribuição espacial. Esses dados contrastam com o  $R(t)$  calculado para o Paraná, que variou entre 1,22 (01/07/2020) e 0,84 (02/10/2020), com valores, na maior parte desse período, abaixo de 1, indicando redução na velocidade de espalhamento do vírus (Calculadora

Epidemiológica, UNB - <https://sds.unb.br/calculadora-epidemiologica/>). Isto porque, apesar de ser utilizado como indicador da velocidade com a qual o vírus se espalha numa determinada população, Delamater et al (GUERRA et al., 2017) e Guerra et al. (GUERRA et al., 2017) advertiram para diferentes tipos de inconsistência vinculados ao valor de  $R(t)$ .

Nicolelis et al. (NICOLELIS et al., 2021) corroboram com as discrepâncias observadas uma vez que associam a aceleração no processo de dispersão a ocorrência de processos multifatoriais. Neste sentido, é importante destacar que, apesar de amplamente utilizado, o valor de  $R(t)$  isolado não se mostra adequado para caracterizar a velocidade de dispersão do vírus em um determinado território.

No Paraná, o município de Itambé (a ser utilizada como exemplo para algumas das análises a seguir) foi o que apresentou maior valor de BETA no período de análise ( $BETA = 0,289$ ,  $p\text{-value} = 0,000$ ,  $R^2 = 0,792$ ). No período compreendido entre os dias 29/06/2020 e 29/07/2020 o número de casos em Itambé passou de 1,77 para 31,37 casos ( $MM7_{fa}$ ) e, para a semana de 09/07/2020 a 16/07/2020 o BETA foi de 2,151 (maior valor observado no município para o intervalo de 01 semana).

Apesar do Ministério da Saúde (BRASIL. MINISTÉRIO DA SAÚDE, 2022) (MS) não definir valores ou procedimentos de cálculo específicos para configurar a ocorrência de “surto epidemiológico”, o MS conceitua surto como correspondente a “situação em que há aumento acima do esperado na ocorrência de casos de eventos ou doença em uma área ou entre um grupo específico de pessoas, em determinado período”.

Considerando que no Paraná 90% dos municípios apresentaram valores de BETA abaixo de 0,148 para o período de análise, valores superiores à 0,15 podem ser considerados como locais nos quais houve, em algum momento durante o período pesquisado, ocorrência de surto. Assim, dos 235 municípios analisados 37 se enquadram nesta condição.

Quadro 1 – Municípios do Paraná nos quais se observou valores de BETA acima de 0,15 (surto) no período de estudo

Adrianópolis	Imbaú	Porecatu
Antonina	Iracema do Oeste	Ramilândia
Atalaia	Itambé	Rio Branco do Sul
Capanema	Itaperuçu	Sabáudia
Corbélia	Jandaia do Sul	Santo Inácio
Coronel Domingos Soares	Jataizinho	São Jorge d'Oeste
Cruzeiro do Iguaçu	Jesuítas	Serranópolis do Iguaçu
Dois Vizinhos	Matelândia	Sulina
Enéas Marques	Nova Esperança do Sudoeste	Telêmaco Borba
Florestópolis	Nova Santa Bárbara	Toledo
Foz do Iguaçu	Ortigueira	Vera Cruz do Oeste
Foz do Jordão	Palotina	
Francisco Alves	Paranaguá	

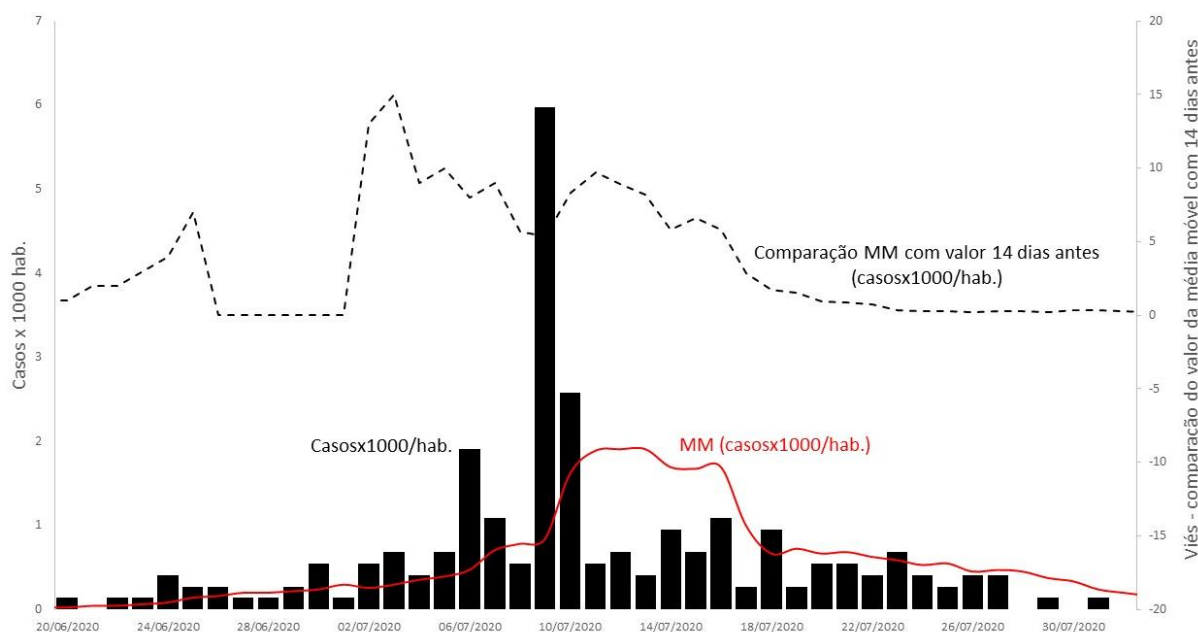
Fonte: Elaborado pelos autores.

Para o intervalo de tempo relativo à semana epidemiológica, os valores de BETA observados foram maiores. Assim, conforme conceituado pelo Ministério da Saúde (BRASIL. MINISTÉRIO DA SAÚDE, 2022) e considerando os 37 municípios com maiores valores de BETA, sugere-se como referência para caracterização de ocorrência de surto de COVID-19 no intervalo de tempo correspondente a semana epidemiológica, valores de BETA acima de 0,3.

### **Análise da média móvel e do viés (comparação da média móvel com o valor de MM observado 14 dias antes)**

Utilizando, ainda, dados do município de Itambé como exemplo, é possível observar que a média móvel (linha vermelha), de fato atenuou as variações observadas nos dados brutos registrados ao longo do tempo, decorrentes de problemas diversos como por exemplo, em função do fluxo de alimentação da base de dados, conforme já observado por Merow e Urban (MEROW; URBAN, 2020).

Figura 3 – Ocorrência de novos casos de COVID-19 (casos x 1000/hab.), Média Móvel e viés. Fonte dos dados: Secretaria de Estado da Saúde do Paraná (SESA-PR)



Fonte: Elaborado pelos autores.

Enquanto se observou um aumento no número de casos de 05 a 06/07 e de 08 a 09/07, a média móvel registrou um curto período de aumento e posterior estabilidade. Isto porque, a partir do momento em que o número de novos casos aumenta e se mantém alto, tanto os valores de MM quanto os de viés tendem a apresentar valores de estabilidade.

Quanto aos valores de viés (linha tracejada - Figura 3), é possível observar que estes fornecem, em determinados momentos, informação discordante sobre a forma como se encontrava o processo de surgimento de novos casos. Neste sentido, valores de viés elevados, baixos e negativos em geral discordam da forma como novos casos ocorrem localmente. Ainda, é importante destacar que os valores de viés observados em Itambé - PR (município com maior valor de BETA no Paraná no período analisado) apresentaram variação inferior a 15%, ou seja, considerados como viés de estabilidade (entre +15% e -15%).

Neste sentido, no que se refere à tomada de decisão e adoção de medidas de combate à doença, tanto os valores de média móvel, quanto de viés, se mostraram pouco representativos em relação a intensidade de propagação da doença e indicadores da necessidade de ações de intervenção direta.

Esse descompasso temporal em relação ao surgimento de novos casos nos valores de MM e de viés (observados em diferentes momentos), dificulta o planejamento e à tomada de decisão, conforme já observado por He e Yunting *et al.* (HE, YUNTING *et al.*, 2020; HE, YUNTING *et al.*, 2020) ao proporem modificações no intervalo de tempo e procedimento de cálculo da média móvel. Isto porque, no momento que o número de casos está aumentando ou diminuindo, a média móvel e os valores de viés



podem indicar diferentes contextos/situações (conforme demonstrado anteriormente). Essas divergências podem dificultar a percepção do processo de dispersão da doença e inviabilizam a comparação entre localidades.

### **Análise dos diferentes métodos de cálculo da velocidade**

Conforme citado anteriormente, quatro procedimentos de cálculo foram aplicados objetivando avaliar a obtenção do valor de inclinação de forma rápida e simplificada, bem como os efeitos das variações no tempo do número de novos casos sobre os valores observados. Observou-se que os quatro procedimentos apresentaram forte correlação (Pearson), com valores acima de 0,853. Os menores valores de Correlação de Pearson observados foram entre os procedimentos MAXQUINTIL e BETA (0,859) e, MAXQUINTIL e TG\_PER (0,853) e, o maior valor de correlação (0,968) foi observado entre o coeficiente angular (BETA) e a tangente do período (TG\_PER) (Tabela 1).

Tabela 1 – correlação entre diferentes métodos de cálculo da velocidade de disseminação

	BETA	TG_PER	MAXQUARTIL	MAXQUINTIL
TG_PER	0,968			
MAXQUARTIL	0,875	0,865		
MAXQUINTIL	0,859	0,853	0,954	
INCL_ACUM	0,926	0,924	0,927	0,908

Fonte: Elaborado pelos autores.

As maiores divergências entre MAXQUINTIL e BETA e, MAXQUINTIL e TG\_PER, eram esperadas uma vez que a regressão linear e a inclinação total adotam como referência todo o período de análise (do primeiro ao último dia) enquanto o MAXQUINTIL analisa as inclinações em cinco intervalos de tempo distintos. Esse fato também se reflete nos valores de correlação entre MAXQUARTIL e BETA e MAXQUARTIL e TG\_PER.

Em função dos valores de correlação observados e das vantagens já observadas em relação aos procedimentos de cálculo de BETA e TG\_PER, as análises a seguir apresentam foco apenas nestes dois.

Por se tratar de um procedimento simplificado de obtenção e similar a BETA, a TG\_PER (conforme proposto neste estudo) é útil para caracterização do processo de disseminação em semanas epidemiológicas, podendo auxiliar a tomada de decisões sobre a ocorrência de surto e necessidade de adoção de medidas. Isto porque o valor de inclinação baseado na  $MM7_{fa}$  apresenta menor oscilação que a MM, que os valores brutos do registro diário e que valores obtidos pelo cálculo da média móvel e de sua comparação com seu valor 14 dias antes.

Assim, a título de exemplo, no período de 29/06/2020 e 29/07/2020 (mês de maior aumento no número de novos casos em Itambé), enquanto BETA e TG\_PER indicavam os valores de 0,289 e 0,219 (caracterizados aqui como surto para os casos de COVID-19 no Paraná), os valores de  $\text{casos} \times 1000 / \text{hab.}$  variaram de 0 a 5,98, enquanto os valores de média de móvel apresentaram variação entre 0,23 e 1,90 e, de viés entre 0 e +15.

Valores mais elevados foram observados para o período relativo à semana epidemiológica de maior aumento no número de novos casos (de 09/07/2020 a 16/07/2020). Neste período, os valores de BETA e TG\_PER foram de 2,151 e 2,124, enquanto o número de novos casos oscilou entre 0,41 ( $\text{casos} \times 1000 / \text{hab.}$ ) a 5,98 novos casos ( $\text{casos} \times 1000 / \text{hab.}$ ). Já os valores de média móvel apresentaram variação entre 0,83 a 1,90 e, de viés entre 5,38 a 9,7, indicando viés de estabilidade.

Esses resultados indicam que o cálculo direto da inclinação (TG\_PER) pode ser utilizado para caracterizar o processo de dispersão da doença, tanto em relação a períodos maiores de análise (como por exemplo para o intervalo total de tempo de ocorrência de casos), quanto para intervalos de tempo menores (como para a semana epidemiológica).

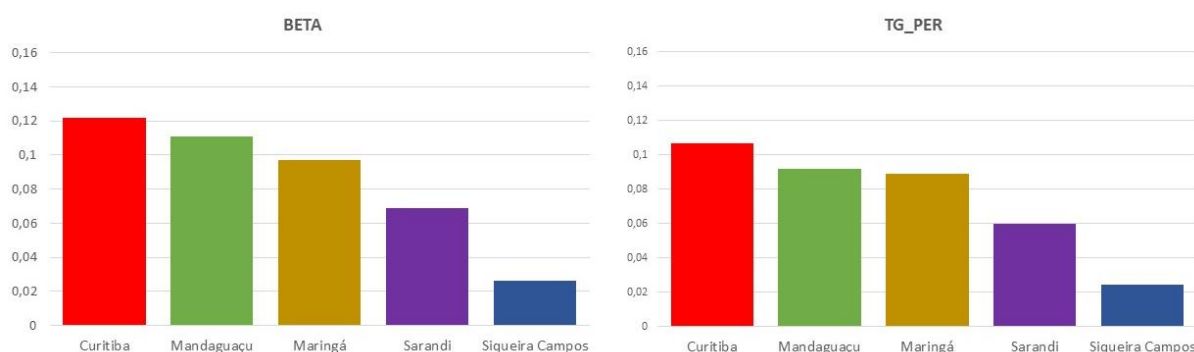
Se por um lado BETA possibilita a obtenção do  $p$ -value e  $R^2$ , e é de fácil implementação computacional, TG\_PER pode ser feito sem a necessidade de construção de regressões, via divisão do valor final do número de casos (calculado utilizando a  $MM7_{fa}$ ) pelo intervalo de tempo. Assim, TG\_PER pode ser utilizada de forma análoga a BETA para comparações rápidas entre localidades com similares intervalos de tempo. A título de primeira aproximação entre os valores de BETA e TG\_PER, a regressão linear entre os resultados obtidos pelos dois procedimentos de cálculo indica, de forma simplificada, que  $TG\_PER = 0,79 * BETA$  ou  $BETA = TG\_PER / 0,79$ .

Cabe destacar que o valor de BETA obtido a partir de regressões lineares é usual em estudos que objetivam avaliar e projetar a dispersão e contaminação da COVID (CHAUHAN; KUMAR; JAMDAGNI, 2020; MEDINA, 2020) e, é diferente do Beta utilizado em EDO (Equações Diferenciais Ordinárias) (LÓPEZ-FLORES et al., 2021).

#### **Efeitos dos procedimentos de cálculo BETA e TG\_PER sobre os valores observados**

No que se refere aos efeitos dos dois principais procedimentos de cálculo sobre os valores observados para o conjunto dos municípios pesquisados, observou-se que os resultados são similares em relação a indicação dos locais com maiores e menores valores de estimativa da velocidade de dispersão (conforme se observa na Figura 4).

Figura 4 – Relação entre BETA e TG\_PER na caracterização dos municípios



Fonte: Elaborado pelos autores.

A Figura 4 mostra um exemplo dos efeitos dos procedimentos BETA e TG\_PER sobre a estimativa da velocidade de dispersão, utilizando dados de casos de cinco municípios do Paraná.

Apesar de diferentes, no geral os procedimentos adotados indicam as mesmas localidades como as que apresentaram maior e menor velocidade de disseminação da COVID-19. No exemplo acima, é possível observar que Siqueira Campos e Curitiba apresentam os menores e maiores valores de inclinação (respectivamente). A diferença entre os procedimentos de cálculo é observada principalmente para as localidades com valores intermediários, os quais podem ser tratados como mais “similares”, como no caso dos municípios de Mandaguauçu e Maringá (via cálculo do valor da TG\_PER), ou “distintos” (cálculo via do valor de BETA).

Dos sete municípios com maiores valores de aceleração em BETA e TG\_PER, seis são comuns aos dois procedimentos de cálculo: Itambé, Sulina, Florestópolis, Matelândia, Adrianópolis e Serranópolis do Iguauçu, com mudanças em relação a posição hierárquica. Itambé, município com maior valor para BETA (0,289), apresentou segundo maior valor em relação a TG\_PER. O coeficiente de correlação indica que os dois procedimentos se encontram fortemente correlacionados (0,972), contudo, em geral os valores de BETA se mostram superiores aos valores de TG\_PER.

Por fim, cabe destacar que 75% dos municípios pesquisados apresentaram valores de BETA e TG\_PER abaixo de 0,125 e 0,107 (respectivamente), e 90% apresentaram valores inferiores a 0,148 e

0,128. Neste sentido, valores superiores aos observados podem ser considerados como uma primeira aproximação indicativa de localidades que apresentaram surto da doença.

Tendo em vista a importância desse tema para a Saúde Pública, a análise experimental de modelos e procedimentos quantitativos é algo necessário para melhor compreender o espalhamento de doenças, e os fatores que influenciam a velocidade e propagação de epidemias em diferentes escalas (MASUM et al., 2022).

## CONSIDERAÇÕES FINAIS

A pesquisa analisou diferentes procedimentos para estimativa da velocidade de dispersão do COVID-19, utilizando dados agrupados em escala municipal, e os efeitos das flutuações no registro dos dados e número de casos sobre os valores estimados.

A compreensão das diferentes possibilidades de acompanhamento do processo de dispersão do vírus auxilia, de forma complementar, às técnicas e métricas existentes. Neste sentido, a pesquisa mostrou que é possível observar mensurar de forma rápida e simples, em escala municipal, possíveis surtos ou eventuais áreas que demandam mais atenção à população (pelos valores de BETA acima de 0,15, por exemplo).

Observou-se que os diferentes procedimentos testados apresentaram forte correlação (valores superiores a 0,853 e, que os procedimentos BETA e TG\_PER apresentaram resultados similares independente do período de análise (correlação de 0,968). Enquanto o cálculo de BETA permite a obtenção dos valores de *p-value* e  $R^2$ , o que possibilita uma avaliação mais robusta de sua representatividade, o cálculo da TG\_PER possui a vantagem de poder ser facilmente empregado em análises rápidas.

O uso dos dados organizados por frequência acumulada e calculados via uso da média móvel com intervalo de tempo de 7 dias ( $MM7_{fa}$ ) possibilitou a obtenção dos coeficientes angulares positivos, bem como a aplicação dos diferentes procedimentos de cálculo testados. Este procedimento mostra relevância para o acompanhamento visual (de forma gráfica) da evolução dos casos de COVID-19 uma vez que, em função do tempo de incubação do vírus e de internação dos pacientes, a permanência do processo de disseminação e de ocorrência de novos casos, ainda que em valores baixos (baixos valores de BETA/inclinação), podem resultar no colapso dos sistemas de saúde.

Os resultados obtidos indicam que tanto BETA como TG\_PER se mostraram viáveis para representar o contexto de disseminação da doença, possibilitando a obtenção de valores de referência para identificação de surtos em diferentes escalas espaciais de análise e tempo, como a exemplo as semanas epidemiológicas.

Ainda, observou-se que tanto a média móvel quanto sua comparação com seu valor 14 dias antes e o valor de  $R(t)$ , podem não resultar em valores indicativos da velocidade de dispersão da doença em determinados intervalos de tempo.

Os valores de BETA indicativos de ocorrência de surto apontados pela pesquisa, se aplicam aos casos de COVID-19 e ao estado do Paraná, cabendo para outras doenças e localidades a análise particular dos dados.

Os procedimentos aplicados mostraram-se promissores, podendo ser aplicados mesmo em séries históricas com limitações nos dados, sendo pertinentes para serem utilizados em novos estudos na área da Geografia da Saúde. Conforme sugerido por Hasan et al (HASAN et al., 2023), técnicas como as utilizadas neste trabalho podem ser empregadas em sistemas de gestão de saúde.

## AGRADECIMENTOS

A Fundação Araucária – financiamento via convênio nº 30/2021 PDI, Protocolo nº 17.788.242-9.

## REFERÊNCIAS

ALBERCA, R. W. et al. Obesity as a risk factor for COVID-19: an overview. **Critical Reviews in Food Science and Nutrition**, v. 61, n. 13, p. 2262–2276, 2021. <https://doi.org/10.1080/10408398.2020.1775546>

BRASIL. MINISTÉRIO DA SAÚDE. **Guia para Investigações de Surtos ou Epidemias**. 2. ed. Brasília: Ministério da Saúde, 2022.

CHAUHAN, P.; KUMAR, A.; JAMDAGNI, P. Regression Analysis of COVID-19 Spread in India and its Different States. **medRxiv**, p. 2020.05.29.20117069, 29 maio 2020. <https://doi.org/10.1101/2020.05.29.20117069>

CLIFF, A.; HAGGETT, P. Methods for the measurement of epidemic velocity from time-series data. **International Journal of Epidemiology**, v. 11, n. 1, p. 82–89, 1982. <https://doi.org/10.1093/ije/11.1.82>

DELAMATER, P. L. et al. Complexity of the basic reproduction number (R0). **Emerging Infectious Diseases**, v. 25, n. 1, p. 1–4, 2019. <https://doi.org/10.3201/eid2501.171901>

DUTRA, C. Estimativa do número básico de reprodução R0 do COVID-19 nos países da América do Sul. **InterAmerican Journal of Medicine and Health**, v. 3, p. 1–7, 2020. <https://doi.org/10.31005/iajmh.v3i0.127>

GRISOTTI, M. Pandemia de covid-19: Agenda de pesquisas em contextos de incertezas e contribuições das ciências sociais. **Physis: Revista de Saúde Coletiva**. Instituto de Medicina Social da UERJ, 2020. <https://doi.org/10.1590/s0103-73312020300202>

GUERRA, F. M. et al. The basic reproduction number (R0) of measles: a systematic review. **The Lancet Infectious Diseases**, v. 17, n. 12, p. e420–e428, 2017. [https://doi.org/10.1016/S1473-3099\(17\)30307-9](https://doi.org/10.1016/S1473-3099(17)30307-9)

HASAN, I. et al. Data analytics and knowledge management approach for COVID-19 prediction and control. **International journal of information technology: an official journal of Bharati Vidyapeeth's Institute of Computer Applications and Management**, v. 15, n. 2, p. 937–954, 2023. <https://doi.org/10.1007/s41870-022-00967-0>

HE, Y. et al. Moving Average Based Index for Judging the Peak of the COVID-19 Epidemic. **International Journal of Environmental Research and Public Health**, v. 17, n. 15, 2020a. <https://doi.org/10.3390/ijerph17155288>

HE, Y.-T. et al. Moving-average based index to timely evaluate the current epidemic situation after COVID-19 outbreak. **medRxiv**, p. 2020.03.24.20027730, 30 mar. 2020b.

JOSÉ, L. et al. O uso da média móvel como indicador de tendência: do mercado financeiro ao acompanhamento dos casos de covid-19. **Negócios em Projeção**, v. 12, n. 2, p. 73–81, 26 nov. 2021.

LEIGHTNER, J.; INOUE, T.; MICHEAUX, P. L. DE. Variable Slope Forecasting Methods and COVID-19 Risk. **Journal of Risk and Financial Management** 2021. Vol. 14, Page 467, v. 14, n. 10, p. 467, 3 out. 2021. <https://doi.org/10.3390/jrfm14100467>

LÓPEZ-FLORES, M. M. et al. **Equações diferenciais e modelos epidemiológicos**. Rio de Janeiro: Instituto de Matemática Pura e Aplicada, 2021.

LYNCH, C. J.; GORE, R. Application of one-, three-, and seven-day forecasts during early onset on the COVID-19 epidemic dataset using moving average, autoregressive, autoregressive moving average, autoregressive integrated moving average, and naïve forecasting methods. **Data in Brief**, v. 35, p. 106759, 1 abr. 2021. <https://doi.org/10.1016/j.dib.2021.106759>

MASUM, M. et al. Comparative study of a mathematical epidemic model, statistical modeling, and deep learning for COVID-19 forecasting and management. **Socio-Economic Planning Sciences**, v. 80, p. 101249, 2022. <https://doi.org/10.1016/j.seps.2022.101249>

MEDINA, M. A. Preliminary Estimate of COVID-19 Case Fatality Rate in the Philippines using Linear Regression Analysis. **SSRN Electronic Journal**, 7 abr. 2020. <https://doi.org/10.2139/ssrn.3569248>

MEROW, C.; URBAN, M. C. Seasonality and uncertainty in global COVID-19 growth rates. **Proceedings of the National Academy of Sciences of the United States of America**, v. 117, n. 44, p. 27456–27464, 3 nov. 2020. <https://doi.org/10.1073/pnas.2008590117>

NICOLELIS, M. A. L. et al. The impact of super-spreader cities, highways, and intensive care availability in the early stages of the COVID-19 epidemic in Brazil. **Scientific Reports**, v. 11, n. 1, p. 13001, 2021. <https://doi.org/10.1038/s41598-021-92263-3>

OVERTON, C. E. et al. Using statistics and mathematical modelling to understand infectious disease outbreaks: COVID-19 as an example. **Infectious Disease Modelling**, v. 5, p. 409–441, 1 jan. 2020. <https://doi.org/10.1016/j.idm.2020.06.008>

PINTO, A. DE S. et al. COVID-19 epidemic curve in Brazil: a sum of multiple epidemics, whose inequality and population density in the states are correlated with growth rate and daily acceleration. An ecological study. **Revista da Sociedade Brasileira de Medicina Tropical**, v. 55, 2022. <https://doi.org/10.1590/0037-8682-0118-2021>

SADOWSKI, A. et al. Big data insight on global mobility during the Covid-19 pandemic lockdown. **Journal of Big Data**, v. 8, n. 1, p. 1–33, 1 dez. 2021. <https://doi.org/10.1186/s40537-021-00474-2>

SEYEDALINAGHI, S. A. et al. Genetic susceptibility of COVID-19: a systematic review of current evidence. **European Journal of Medical Research**, v. 26, n. 1, p. 1–12, 1 dez. 2021. <https://doi.org/10.1186/s40001-021-00516-8>

SILVA, L.; FILHO, D. F.; FERNANDES, A. The effect of lockdown on the COVID-19 epidemic in Brazil: evidence from an interrupted time series design. **Cadernos de Saúde Pública**, v. 36, n. 10, p. e00213920, 19 out. 2020. <https://doi.org/10.1590/0102-311x00213920>

TUAN, N. H.; MOHAMMADI, H.; REZAPOUR, S. A mathematical model for COVID-19 transmission by using the Caputo fractional derivative. **Chaos, Solitons and Fractals**, v. 140, 1 nov. 2020. <https://doi.org/10.1016/j.chaos.2020.110107>

WAMOYI, J.; RANGANATHAN, M.; STÖCKL, H. COVID-19 social distancing measures and informal urban settlements. **Bulletin of the World Health Organization**, v. 99, n. 6, p. 475, 6 jun. 2021. <https://doi.org/10.2471/BLT.20.265942>