



CURSO DE MEDICINA

CAMILA VERÔNICA SOUZA FREIRE

**INFODEMIA DE COVID19: METÁFORA CONTAGIOSA OU REALIDADE
PREOCUPANTE? UMA EXPLORAÇÃO ANALÍTICA DAS TAXAS DE
PUBLICAÇÃO DE ARTIGOS E PRÉ-PRINTS NOS PRIMEIROS MESES DA
PANDEMIA**

SALVADOR

2021

Camila Verônica Souza Freire

**INFODEMIA DE COVID19: METÁFORA CONTAGIOSA OU REALIDADE
PREOCUPANTE? UMA EXPLORAÇÃO ANALÍTICA DAS TAXAS DE
PUBLICAÇÃO DE ARTIGOS E PRÉ-PRINTS NOS PRIMEIROS MESES DA
PANDEMIA**

Trabalho de Conclusão de Curso apresentado ao Curso de Graduação em Medicina da Escola Bahiana de Medicina e Saúde Pública como requisito parcial para aprovação no 4º ano de Medicina.

Orientador: Prof. Luís Cláudio Correia.

Co-orientador: Prof. João de Deus Barreto Segundo.

Salvador

2021

Camila Verônica Souza Freire

**INFODEMIA DE COVID19: METÁFORA CONTAGIOSA OU REALIDADE
PREOCUPANTE? UMA EXPLORAÇÃO ANALÍTICA DAS TAXAS DE
PUBLICAÇÃO DE ARTIGOS E PRÉ-PRINTS NOS PRIMEIROS MESES DA
PANDEMIA**

Trabalho de Conclusão de Curso apresentado
ao Curso de Graduação em Medicina da Escola
Bahiana de Medicina e Saúde Pública como
requisito parcial para aprovação no 4º ano de
Medicina.

Salvador, ____ de _____ de ____.

BANCA EXAMINADORA

Prof. Dr. Luís Cláudio Correia
Orientador
Medicina - EBMSP

Nome do Examinador
Medicina - EBMSP

Nome do Examinador
Medicina – EBMSP

AGRADECIMENTOS

É dado o momento de agradecer a todas as pessoas que me auxiliaram para que pudesse concluir mais uma etapa da minha formação.

Ao prof. Luís Correia, que representa minha maior referência e fonte de aprendizado científico, médico e humano.

À Gabriel de Araújo Grisi, pela orientação e por ser meu maior parceiro na vida.

Ao prof. João de Deus Segundo, pela orientação e apoio constante no desenvolvimento científico.

Ao Grupo Brasileiro de Metaciência e Medicina Baseada em Evidências, pela experiência, aprendizado e oportunidade de iniciar minha carreira científica.

Ao prof. Juarez Pereira Dias, pelo apoio e acompanhamento constante no desenvolvimento desse trabalho.

À Lindalva Tenório de Souza e Paulo José de Souza, meus avós, pelo amor e base.

Ao Diretório Acadêmico Pirajá da Silva, pela oportunidade do desenvolvimento da consciência sociopolítica dentro da universidade.

À Escola Bahiana de Medicina e Saúde Pública, principalmente, em nome da Dra Maria Luisa Soliani, Luiza Ribeiro e Emanuel Souza, pelo incentivo e fundamental apoio nessa jornada.

RESUMO

Introdução: Durante o primeiro ano da pandemia de COVID19, o neologismo “infodemia” foi amplamente utilizado para caracterizar um excesso de publicações de COVID19 de baixa qualidade. O valor dessa metáfora epidemiológica se respalda no significado comportamental de um fenômeno informacional contagioso com um padrão exponencial de crescimento subjacente. Ou seja, o sucesso percebido por meio da publicação de um grupo de autores poderia ter pressionado outros autores a publicarem independentemente da competência sobre o assunto ou da qualidade do artigo. **Objetivo:** Testar a hipótese de que as publicações científicas sobre a pandemia COVID19 seguiram um padrão epidêmico, avaliando se o padrão de crescimento foi exponencial ou não. **Metodologia:** Variando de janeiro a julho de 2020, estudos sobre COVID19 e seus metadados foram recuperados do PubMed, bioRxiv e medRxiv. Os estudos de Influenza publicados durante o mesmo intervalo um ano antes, juntamente com seus metadados, foram empregados como um grupo de controle. A incidência semanal de publicação foi plotada e analisada para COVID19 e Influenza separadamente. O crescimento exponencial das publicações foi testado por regressão linear transformada em log, tendo o número de publicações como variável dependente. Além disso, o número de publicações semanais (R_p) foi definido como uma divisão da incidência semanal de publicações pela média das semanas anteriores. Os mesmos procedimentos foram realizados separadamente para artigos revisados por pares e pré-prints. **Resultados:** As publicações sobre COVID19 em periódicos aumentaram semanalmente, atingindo 97.781 documentos, um aumento de quatorze vezes em relação à produção do Influenza um ano antes, que acumulou 6.649 documentos no intervalo equivalente. A incidência de artigos de periódicos sobre COVID19 aumentou até a 20ª semana do período analisado. Observados por semana, eles exibiram um crescimento exponencial entre a 6ª e a 20ª semana. Um padrão semelhante foi observado para seus pré-prints entre a 8ª e a 20ª semana. Os artigos de periódicos sobre COVID19 apresentaram $R_p = 3,17 \pm 0,72$, enquanto o controle apresentou $R_p = 0,97 \pm 0,12$. Seus pré-prints exibiram $R_p = 2,18 \pm 0,54$ e $0,97 \pm 0,27$, respectivamente. Assim, nenhuma evidência de crescimento exponencial foi observada no grupo de controle. **Conclusão:** As publicações sobre COVID19 apresentaram um padrão epidêmico. Uma vez que o padrão dos artigos revisados por pares foi semelhante ao padrão dos pré-prints, é plausível que não apenas os autores, mas também os editores tenham sido suscetíveis ao efeito da infodemia.

Protocolo do estudo: URL: <https://osf.io/z53tc>.

Palavras-chave: Comunicação científica. Publicações científicas: periódicos. Publicações científicas: preprints. COVID-19. Influenza.

ABSTRACT

Background: During the first year of the COVID19 pandemic, the neologism infodemic has been employed to characterize a deluge of low quality COVID19 publications. The value of this epidemiologic metaphor could rely on the behavioral meaning of a contagious informational phenomenon with an underlying exponential pattern of growth. That is, the perceived success by means of publication of a group of authors could pressure other authors to join in the conversation regardless of competence on the matter or quality of their paper. **Objective:** To test the hypothesis that scientific publications on the COVID19 pandemic followed an epidemic pattern by assessing whether the growth pattern was exponential or not. **Methodology:** Ranging from January to July 2020, studies on COVID19 as primary subjects and their metadata were retrieved from PubMed, bioRxiv and medRxiv. Influenza studies published during the same interval a year prior along with their metadata were employed as a control group. Weekly incidence of publication was plotted and analyzed for COVID19 and Influenza separately. Exponential growth of publications was tested by log-transformed linear regression, having number of publications as the dependent variable. In addition, weekly reproduction number of publications (R_p) was defined as a division of the week's incidence of publications by the average of previous weeks. Same procedures were performed separately for peer-reviewed papers and preprints. **Results:** COVID19 output in journals increased weekly, reaching 97,781 documents, a fourteen-fold increase compared to the Influenza output, which accumulated 6,649 documents in the equivalent interval. COVID19 journal articles incidence increased until the 20th week of the analyzed timeframe. Plotted per week, they displayed exponential growth between the 6th and the 20th week. A similar pattern was observed for their preprints between the 8th and the 20th week. COVID19 journal articles displayed $R_p = 3.17 \pm 0.72$, meanwhile control presented $R_p = 0.97 \pm 0.12$. Their preprints exhibited $R_p = 2.18 \pm 0.54$ and 0.97 ± 0.27 respectively. Thus, no evidence of exponential growth was observed in the control group. **Conclusion:** COVID19 publication displayed an epidemic pattern. Since no pattern distinction was observed between peer-reviewed and preprint outputs and journals were responsible for the bulk in publication, it would appear that not only authors but also editors were susceptible to the infodemic bug.

Study Protocol: URL: <https://osf.io/z53tc>.

Keywords: Scientific communication. Scientific publications: periodicals. Scientific publications: preprints. COVID-19. Influenza.

SUMÁRIO

1 INTRODUÇÃO	5
2 OBJETIVOS	7
3 RACIONAL TEÓRICO.....	7
4 METODOLOGIA.....	9
4.1 Desenho de estudo	9
4.2 Estratégia de busca e pesquisa	9
4.3 Elegibilidade	10
4.5 Extração de dados e avaliação da validade interna	10
4.6 Comparação.....	11
4.7 Avaliação das funções	12
4.8 Número de reprodução de publicações.....	12
4.9 Teste de Hipóteses	13
4.10 Considerações éticas.....	13
5 RESULTADOS	13
5.1 Incidência de publicações	13
5.2 Função de modelagem de publicações ao longo do tempo	15
5.3 Análise de reprodução de publicação	16
6 DISCUSSÃO	18
7 CONCLUSÃO.....	21
REFERÊNCIAS.....	22
MATERIAL SUPLEMENTAR.....	25

1 INTRODUÇÃO

A produção científica global tem dobrado a cada nove anos¹, deixando a comunidade científica lutando para se manter atualizada e separar o joio do trigo²⁻⁴. A publicação excessiva é uma reclamação antiga da comunidade científica, o que levou à criação da revista científica para organizar e validar pesquisas e definir o que é (ou não) conhecimento científico⁵⁻⁸. Não obstante, as queixas sobre a qualidade, englobadas no conceito de desperdício evitável de pesquisa, constituem um debate recente que vem crescendo nos últimos trinta anos⁹⁻¹¹, dando origem à noção de que talvez os incentivos do ecossistema científico pudessem estar distorcendo as publicações em direção da novidade e quantidade e para longe da confiabilidade e qualidade^{2-3, 11}. Se as pressões para publicar novidades são realmente um forte impulso por trás da produção científica atual e a qualidade metodológica média da pesquisa biomédica já era considerada questionável antes da pandemia^{2-4,9-11}, o que aconteceria se o ecossistema científico se encontrasse sob pressão ainda maior causada por uma crise sanitária mundial?

Evidências acumulativas de estudos independentes têm sugerido que o ecossistema de publicação parece estar apressando a publicação de artigos relacionados ao COVID19, de qualidade inferior, em periódicos, em comparação com tempos pré-pandêmicos: o tempo desde a submissão até a publicação de artigos relacionados ao COVID19 em periódicos acelerou¹²⁻¹⁷, enquanto há evidências de que a qualidade metodológica geral dos estudos diminuiu^{15,18}. O tempo médio para a aceitação final foi oito vezes mais rápido para artigos revisados por pares relacionados ao COVID19 em comparação com artigos sobre outras questões^{15-16,19}. Em uma amostra de periódicos indexados no PubMed, 10% dos estudos relacionados ao COVID19 foram aceitos dois dias após a submissão¹⁶. Um estudo sobre revisão por pares aberta de duas revistas de boa reputação internacional encontrou distinções qualitativas na revisão de artigos relacionados ao COVID19, que geralmente passaram por uma única rodada de revisão por pares, exibiram menos propensão para solicitar mais dados e mais propensão para solicitar aos autores que suavizassem afirmações e conclusões, forçando-os a reconhecerem mais incerteza no texto do relato¹⁹.

Também foi verificado que, em 2020, 27% dos autores ativos da base de dados SCOPUS publicaram pesquisas sobre COVID19 em uma disciplina de subcampo que não estava entre as três principais disciplinas de subcampo em que publicaram com mais frequência durante sua carreira conforme registro naquele indexador. Aproximadamente, 2,8% da literatura científica indexada ao SCOPUS publicada em 2020 estava relacionada ao COVID19 e mais de 4% de

todos os cientistas que publicaram naquele período estavam publicando sobre COVID19. Ou seja, aproximadamente um em cada sete cientistas ativos que publicam em periódicos de nível médio e alto nível de língua inglesa ajustaram rapidamente o seu portfólio para incluir artigos sobre COVID19. Esses autores foram encontrados em 21 áreas principais do SCOPUS e até mesmo pesquisadores especializados em pesca, ornitologia, entomologia e arquitetura publicaram sobre COVID19 em 2020 em revistas científicas indexadas no SCOPUS²⁰. Por outro lado, os manuscritos do COVID19 foram carregados como pré-prints simultaneamente à sua submissão a periódicos, o que indica que os autores, em média, não buscavam especificamente o feedback da pré-submissão ao publicar sua pesquisa como pré-print²¹.

Esses achados apontam para um fenômeno informacional que talvez pudesse ser contagioso como outros fenômenos informacionais têm se mostrado: um exemplo famoso é o Efeito Mateus nos estudos métricos de avaliação da produção científica, em que citações causam mais citações^{22,23}. Essa aceleração na publicação começada na pandemia tem sido discutida pela comunidade acadêmica sob o conceito de infodemia^{12,24}. Mas, poderia o excesso de artigos em periódicos e em servidores de pré-prints ser enquadrado como uma infodemia na qual as publicações do COVID19 teriam causado ainda mais publicações do COVID19 de forma exponencial? Em comparação com qual grupo de controle e em qual período de tempo? E, em caso afirmativo, qual teria sido a participação de periódicos e servidores de pré-prints no fenômeno? As taxas de publicação relacionadas ao COVID19 têm seguido algum tipo de padrão epidêmico ou o termo infodemia é apenas uma metáfora contagiosa? A culpa foi dos pré-prints? E em caso afirmativo, em que medida?

À luz deste raciocínio, este estudo teve como objetivo avaliar se as publicações relacionadas ao COVID19 seguem/seguiram um padrão epidêmico. Além disso, objetivou avaliar o comportamento das taxas de publicação de pré-prints em comparação com as de publicação em periódicos. Até onde sabemos, essa comparação de taxas ainda não havia sido feita e os pré-prints provaram ser uma parte relevante do ecossistema da publicação científica²⁶⁻²⁸. Os autores entendem a gravidade da crise atual e seu objetivo foi comparar o ecossistema de publicação em circunstâncias normais com ele sob as pressões pandêmicas atuais e, se possível, caracterizar criticamente e discutir uma eventual infodemia por meio de um experimento mental (*thought experiment*). Assim, as publicações sobre Influenza foram escolhidas como grupo controle para testar a hipótese, pois tanto o Influenza quanto o COVID19, embora muito diferentes no que diz respeito à gravidade e morbidade, são pandemias em andamento.

Publicações sobre MERS-CoV, doença causada pelo vírus Ebola ou vírus Zika foram consideradas inadequadas para grupo controle porque seus respectivos surtos foram autolimitados no tempo e em regiões específicas do mundo, portanto, incapazes de exercer pressão contínua sobre o ecossistema de publicação como um todo.

2 OBJETIVOS

Descrever e analisar se o comportamento de publicações científicas relacionadas ao COVID19 seguiu um padrão epidêmico.

3 RACIONAL TEÓRICO

Pesquisas sobre a eficácia de intervenções biomédicas podem ser seguidas por estudos que chegam a conclusões extremas opostas ou que sugerem que as alegações originais eram muito mais fortes ou muito mais fracas, com tamanhos de efeito muito insuflados ou muito subestimados²⁹⁻³¹. Pode-se tratar desses temas em um emergente campo de estudo, a metaciência. Esse campo mantém interface com diversos outros campos e disciplinas, avaliando métodos, transparência e reprodutibilidade, sistemas de incentivos e outros temas que envolvam o comportamento dos pesquisadores e do ecossistema de publicação científica no intuito de melhorar a eficiência do processo de produção e comunicação da ciência³². Incluem-se na metaciência investigações teóricas e empíricas com o objetivo de avaliar e melhorar as práticas de pesquisa, em especial da pesquisa biomédica³². Neste racional teórico, o presente trabalho foi desenvolvido.

Perguntas de investigação de baixa probabilidade, métodos ineficientes e, paralelo a isso, o sistema de incentivos enviesado em prol de quantidade³³ em detrimento da qualidade^{2,3,11} podem estar causando desalinhamento entre conclusões científicas e fenômenos do mundo real¹⁰. Com isto, estudos pequenos podem acabar sendo refutados por estudos posteriores mais bem dimensionados para detecção de um dado fenômeno biomédico. Há ainda uma confusão entre estudos para detectar sinal (há efeito ou não) e estudos para quantificar magnitude de efeitos³⁴. Exemplo disso é que estudos epidemiológicos e ensaios randomizados podem chegar em resultados muito distintos sobre uma mesma questão de pesquisa²⁹⁻³¹. Existe, naturalmente, a possibilidade de que estudos subsequentes sejam, na verdade, os incorretos e/ou que outras variáveis possam estar em jogo, como o espectro das manifestações clínicas da doença, critérios

de elegibilidade, intervenções concomitantes, manipulação analítica e/ou genuína diversidade populacional. Contudo, evidências sugerem que quando se trata de pesquisas recentes, mesmo ensaios clínicos randomizados, os achados devem ser interpretados com cautela, especialmente quando a pesquisa de um problema está em seus estágios iniciais^{29,31,35}, como é o caso da pesquisa sobre a pandemia de COVID19 e sobre o seu agente etiológico.

Este fenômeno foi nomeado Proteus após uma divindade helênica que representava a mutabilidade do mar ou a fluidez da água: quanto mais extremo um resultado relatado, mais provável é que ele tenha sido publicado devido ao viés de publicação³⁷ e ao *time lag bias*^{29,31,35}. Se e quando mais estudos forem realizados sobre o mesmo tema, seus resultados provavelmente serão, em média, menos extremos do que os resultados do primeiro estudo estipulam o Fenômeno Proteus^{29,31,35}. Alguns corolários que o compõem são os seguintes: (a) a diversidade mais extrema entre estudos que abordam a mesma questão de pesquisa geralmente tendem a ocorrer muito cedo durante o acúmulo de pesquisa publicada; (b) Esse fenômeno pode ocorrer quando um grande pool de análises potenciais está disponível e muitas equipes em todo o mundo podem gerar e disseminar rapidamente dados. Desta forma, a pandemia de COVID19, tensionando o ecossistema de publicação em busca de soluções urgentes para a grave crise em que o mundo se encontra desde dezembro de 2019, poderia ter gerado condições para a tempestade perfeita de estudos extremos em excesso, os quais possivelmente relatavam resultados superdimensionados ou com falsos positivos. Tendo essa consequência grave em mente, se faz necessário entender o comportamento de publicação em revistas e em pré-prints durante a pandemia, se de fato foi excessivo ou não e em comparação com o que.

Apesar de suas muitas problemáticas – variando de oligopólio de editoras comerciais⁸, preconceito contra autores de países pobres ou em desenvolvimento³⁷, baixa qualidade metodológica apesar de prestígio¹¹ –, a publicação em periódicos tem sido o principal espaço através do qual o conhecimento científico circula na comunidade científica. Isso permaneceu incontestável durante a maior parte de aproximadamente trezentos anos de existência do periódico científico⁵⁻⁸. Até o surgimento dos servidores de pré-prints eletrônicos como um meio alternativo de circulação do conhecimento científico^{25,26}.

A publicação de pré-prints eletrônicos foi fundada com base na premissa de que postar publicamente pré-prints permitiria aos autores receber *feedback* e enviar artigos aprimorados para periódicos, ainda que tenha sido usada como uma forma de estabelecer prioridade e

contornar políticas de direitos autorais estritas de editoras comerciais^{21,26}, resultando em uma rota mais rápida de disseminação de descobertas científicas do que a maneira tradicional. A publicação de pré-prints existe pelo menos desde a década de 1990, quando o arXiv começou a distribuir uma coleção com curadoria de artigos que ainda não haviam passado por revisão por pares³⁸. A prática ganhou força em disciplinas como física, matemática, astronomia e ciência da computação. Nas ciências biológicas, o bioRxiv se estabeleceu como o servidor de pré-prints de referência após seu lançamento em 2013, sendo mantido e operado pelo Cold Spring Harbor Laboratory, que também é responsável pelo medRxiv, um projeto irmão mais novo fundado em 2019 dedicado às ciências médicas^{25,26}.

Embora a publicação de pré-prints tenha sido encarada com suspeita pela comunidade de pesquisa médica nos últimos anos, a pandemia de COVID19 provocou um aumento na publicação sobre o tema tanto em periódicos revisados por pares quanto em servidores de pré-prints, com mais de 16.000 artigos relacionados a COVID19 dentro de 4 meses do primeiro caso confirmado, dos quais pelo menos 6.000 foram hospedados por servidores de pré-prints²⁸. Os pré-prints de COVID19 se mostraram mais curtos, com menos referências, geralmente de autoria única e acessados e baixados pelo menos 15 vezes mais do que os pré-prints de outros temas que não a COVID19, recebendo mais atenção logo após terem sido postados online^{21,28}.

O possível excesso de publicações em revistas e em servidores de pré-prints suscitou a preocupação de que uma epidemia de artigos e de informação – “paper infodemic”, “paperdemic” ou “infodemia” – estaria em andamento, chamando atenção e críticas às publicações de pré-prints, apesar de a suspeita de infodemia ter afetado também os periódicos^{12-17,21,24}.

4 METODOLOGIA

4.1 Desenho de estudo

Cientométrico, metacientífico, descritivo e analítico.

4.2 Estratégia de busca e pesquisa

Para artigos de revistas indexadas no PubMed, o aplicativo Entrez Direct (EDirect) NIH / NLMⁱ foi empregado para recuperar metadados de publicação do MEDLINE por E-utilities API. A recuperação de metadados MEDLINE incluiu artigos de 1º de dezembro de 2019 a 6 de fevereiro de 2020 para COVID19. O mesmo intervalo, um ano antes, foi considerado para artigos sobre Influenza porque presumiu-se que as publicações sobre esse tópico foram suprimidas em 2020 devido ao foco dos especialistas na pandemia de COVID19. Assim, este estudo objetivou uma abordagem mais conservadora para não superestimar os achados. As estratégias de busca foram detalhadas nas informações complementares.

Os metadados de preprints foram recuperados dos servidores bioRxiv e medRxivⁱⁱ, presumidos como os mais proeminentes entre os servidores de ciências da vida até hoje^{17,21,25-28}. As ferramentas de pesquisa avançada nativa conectando bancos de dados medRxiv e bioRxiv foram empregadas para consultar preprints.ⁱⁱⁱ O intervalo de recuperação de metadados de preprints para COVID19 variou de 13 de janeiro de 2020 (primeira ocorrência) a 28 de julho de 2020. O mesmo intervalo de um ano antes foi considerado para preprints de Influenza.

Para descartar a possibilidade de ausência de metadados nos arquivos de saída do EDirect, os resultados de COVID19 no EDirect foram comparados aos resultados do PubMed^{iv}. A comparação não revelou perda de artigos através da extração automática empregando EDirect. Assim, para reduzir a possibilidade de a coleta manual de metadados induzir viés, erro ou perda de metadados, optou-se pelo uso do software EDirect.

4.3 Elegibilidade

Os metadados de todos os artigos de periódicos extraídos pelo EDirect foram incluídos. Foram excluídos os metadados de preprints que se misturaram nos resultados da pesquisa e os artigos que não vieram com uma data completa de publicação no momento da extração. Todos os resultados da pesquisa bioRxiv e medRxiv para preprints foram incluídos. Nenhuma curadoria de metadados foi executada neste ponto para qualquer grupo. Foi realizada uma simulação, excluindo cartas, editoriais e artigos de opinião da produção de artigos publicados, não ocorrendo alteração expressiva na quantidade de artigos publicados.

4.5 Extração de dados e avaliação da validade interna

Os metadados do EDirect de artigos de periódicos sobre COVID19 e Influenza foram baixados em um documento legível do Microsoft Excel (Microsoft Corporation, Estados Unidos da América). Os metadados de preprints sobre Influenza foram exportados para um arquivo Extensible Markup Language (XML) a partir da página de resultados de pesquisa de preprints. Os preprints sobre COVID19 foram coletados da página dedicada ao COVID19 conectando os bancos de dados dos servidores, acessada em 28 de julho de 2020.^v

Para avaliar a validade do emprego de um método não curado e automático de extração de metadados, a estratégia de busca de artigos de periódicos foi repetida em diferentes momentos, nos quais a produção do PubMed foi comparada à do NCBI LitCovid. Para mais detalhes, por favor, consulte o material complementar. Devido a essa avaliação, foi verificado que o PubMed tinha 1,17 mais publicações no COVID19 do que o LitCovid. Assim, decidiu-se que o PubMed era mais confiável para uma cobertura precisa. Foi considerado desnecessária a confirmação dos resultados dos metadados de Influenza devido ao resultado ser oito vezes menor do que o output análogo em COVID19.

As variáveis PubmedPubDate @ pubmed, uma data que reflete a inclusão na base de dados PubMed, e PubDate, uma data que reflete a data de emissão da publicação, foram extraídas do PubMeb como controle de qualidade de ArticleDate, uma variável que corresponde à data em que o periódico publicou o artigo online⁶. PubmedPubDate @ pubmed PubDate foi percebido como apresentando perda de precisão, com artigos que pareciam ter sido adicionados à coleção do NLM em bloco dias ou semanas após a data da primeira publicação online nos sites dedicados dos periódicos. Assim, ArticleDate foi escolhido como variável por se mostrar equivalente ou ligeiramente mais matizado que PubmedPubDate @ pubmed e PubDate, apresentando-se então como a escolha mais acurada e dotada de nuances.

Para possibilitar a criação de uma timeline de incidência de publicação, as semanas foram numeradas automaticamente em Excel, colocando-se a semana 1 em janeiro para todos os grupos. Detalhes adicionais da função de data também foram disponibilizados no material suplementar.

4.6 Comparação

Os artigos de periódicos indexados e preprints foram agrupados por tema de publicação: COVID-19 ou Influenza, e uma comparação entre as espécies de publicações foi realizada.

4.7 Avaliação das funções

Como toda função exponencial torna-se linear quando é transformada em logaritmo, para permitir o teste da hipótese de crescimento exponencial, foi realizada uma transformação logarítmica nos dados de incidência cumulativa. Portanto, pode ser avaliado por regressão linear. Foi analisado o padrão de crescimento da curva e quanto mais acentuada a inclinação, mais pronunciado foi o crescimento, o que caracterizou o efeito multiplicador da variável independente do tempo como a potência da função exponencial. Os intervalos de tempo para análise foram posteriormente definidos após avaliação do comportamento da curva e melhor ajuste da função.

4.8 Número de reprodução de publicações

O número de reprodução efetiva (R_t) ajuda a projetar os padrões de crescimento da epidemia, pois é um indicador da contagiosidade de um agente etiológico dentro de um intervalo temporal e contexto específicos. Matematicamente, foi estabelecido como o número de casos secundários que um único caso em determinado momento geraria³⁹. Diante disso, um parâmetro inspirado

$$R_p = \frac{1}{n} \sum_i^n s_i$$

conceitualmente no número de reprodução efetiva foi concebido para estruturar o presente experimento mental: o número de reprodutibilidade da publicação (R_p). Este parâmetro consistia na divisão da incidência semanal de publicações pela média dos casos semanais anteriores, onde i é a primeira semana e n é o número total de semanas no intervalo de tempo analisado, em que s representa a semana.

No contexto de uma infodemia, novos casos deveriam ser entendidos como artigos ou preprints adicionais publicados. Esse conceito foi concebido partindo do pressuposto de que o desejo de publicar estudos relacionados ao COVID19 poderia ter sido ampliado pela quantidade de literatura do COVID19 já publicada e pela atenção que essa literatura recebeu, o que é apoiado

pelas evidências independentes sobre o comportamento da publicação na pandemia^{12-21,24,27,28} e sobre o sistema de recompensa defeituoso da publicação científica^{2,3}, com a possível consequência de que esse excesso de estudos sobre COVID19 esteja em descompasso com a realidade^{29,31,35} por conta do Fenômeno Proteus.

Conforme a metáfora de infodemia, para definir melhor uma epidemia de artigos publicados e preprints, o R_p forneceria um quadro mais claro onde a simples comparação da quantidade de artigos publicados e a produção de preprints não seria eficiente para análise. Portanto, dentro de um modelo putativo em que cada artigo teve influência perpétua sobre a geração subsequente de mais artigos, os autores influenciariam - ou, metaforicamente, contaminariam - outros autores com um desejo de gerar artigos adicionais do COVID19. Portanto, de acordo com esse raciocínio, a maneira como o R_p havia crescido nos primeiros meses da pandemia COVID19 poderia informar o quanto a incidência média de novas publicações semanais estaria aumentando ou diminuindo, apoiando ou refutando a alegação de uma infodemia em curso.

4.9 Teste de Hipóteses

Artigos de periódicos e preprints foram comparados por tópico de publicação (COVID19 ou Influenza) e por tipos de publicação (preprints ou artigos de periódicos). A diferença entre os valores de COVID19 e Influenza R_p foi comparada por teste t independente. A análise estatística e as estimativas das curvas foram executadas no SPSS Statistics versão 14 (International Business Machines Corporation, Estados Unidos da América).

4.10 Considerações éticas

Este estudo não necessita de autorização ética por sua natureza documental e pública, lidando com metadados abertos de artigos publicados e bases de dados nas quais estão indexados.

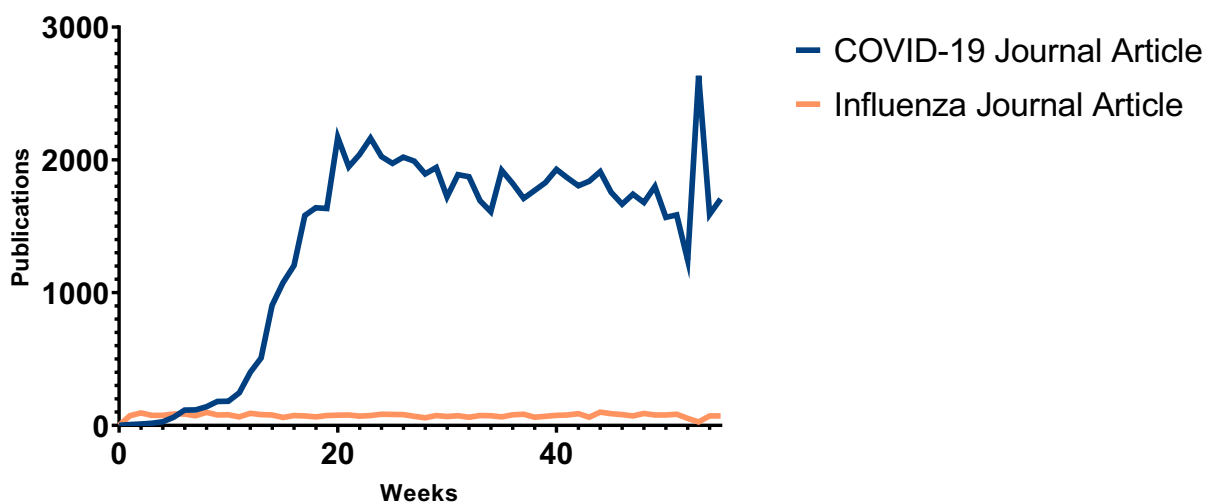
5 RESULTADOS

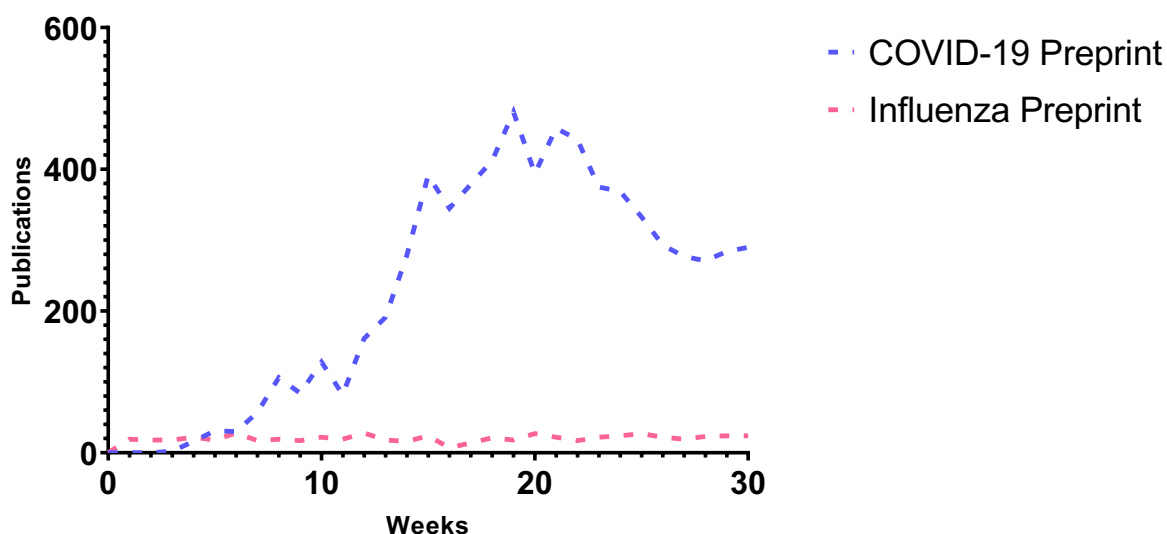
5.1 Incidência de publicações

De janeiro de 2020, mês da primeira publicação do COVID19, a 6 de fevereiro de 2021, a produção do COVID19 aumentou semanalmente para um total de 97.781 documentos, um

aumento de quatorze vezes em relação à produção do Influenza, que acumulou 6.649 documentos no intervalo equivalente. A incidência de artigos de periódicos COVID19 aumentou até a 20ª semana de 2020, semanalmente, de janeiro a junho, enquanto o output de Influenza permaneceu estável ao longo do tempo um ano antes (Figura 1). Da 20ª à 52ª semana, o número de artigos de periódicos COVID19 se estabilizou em torno de 1.820 (DP +/- 233,8). (Figuras 3 e 4 do material complementar).

Da 1ª à 30ª semana de 2020, os preprints de COVID19 foram responsáveis por 6.960 documentos e os de Influenza por 614 (Figura 1a). A incidência de preprints de COVID19 também aumentou significativamente até a semana 19, seguida por um número de incidência decrescente. O controle, por outro lado, exibiu estabilidade em todo o período observado.





Figuras 1, 1a. Incidência de publicação semanal de COVID19 e Influenza.

5.2 Função de modelagem de publicações ao longo do tempo

Tanto para os artigos de periódicos quanto para preprints sobre COVID19, a transformação do log do número cumulativo semanal de publicações exibiu um comportamento exponencial (Figura 2). As regressões lineares descreveram os dados neste lapso de tempo através de $R^2 = 99\%$ para artigos de periódicos e $R^2 = 96\%$ para preprints, através das semanas (t) com: $f(t) = 0.29t + 3.9$ e $f(t) = 0.26x + 3.3$ respectivamente.

Foram encontrados os seguintes coeficientes exponenciais de 0,29 (IC 95%, 0,28-0,29) e 0,26 (IC 95%, 0,23-0,29) para artigos de periódicos e preprints COVID19 respectivamente. Os modelos exponenciais para publicações (p) por tempo (t) foram $p(t) = 206.29e^{0.284(t)}$ para artigos de periódicos e $p(t) = 100.32e^{0.2584t}$ para preprints.

O grupo de controle Influenza apresentou um padrão de crescimento linear estável. Na função semi-log cumulativa, artigos de periódicos e preprints apresentaram uma inclinação de 0,08 ($R^2 = 96\%$, IC 0,07 - 0,09). Assim, se analisarmos a função de incidência cumulativa ao longo do tempo (sem qualquer transformação logarítmica), um padrão linear com uma inclinação de 75,84 (IC 74 - 77,6) para artigos de periódicos e 18,8 (IC 18,1 - 19,5) para preprints foi encontrado, cada um com a $R^2 = 99,8\%$. Conseqüentemente, a melhor função para representar o grupo controle não foi a exponencial.

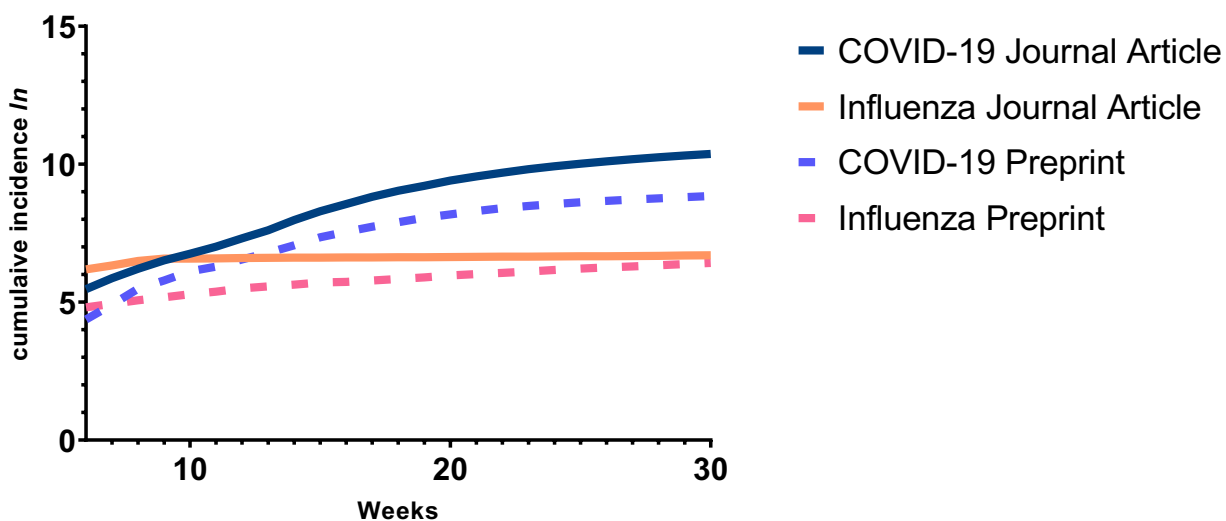


Figura 2. Publicações cumulativas semanais transformadas em log.

5.3 Análise de reprodução de publicação

O R_p para artigos de periódicos da 6^a à 20^a semana apresentou diferença estatisticamente significativa entre COVID19 ($3,17 \pm 0,72$) e Influenza ($0,97 \pm 0,12$), $p < 0,01$. Também foi encontrada diferença para preprints: COVID19 ($2,18 \pm 0,54$) e Influenza ($0,97 \pm 0,27$), $p < 0,01$. Dentro dos grupos COVID19 e Influenza, R_p de preprints e artigos de jornal não diferiu: $p = 0,15$ para publicações COVID e $p = 0,95$ para Influenza (Figura 3).

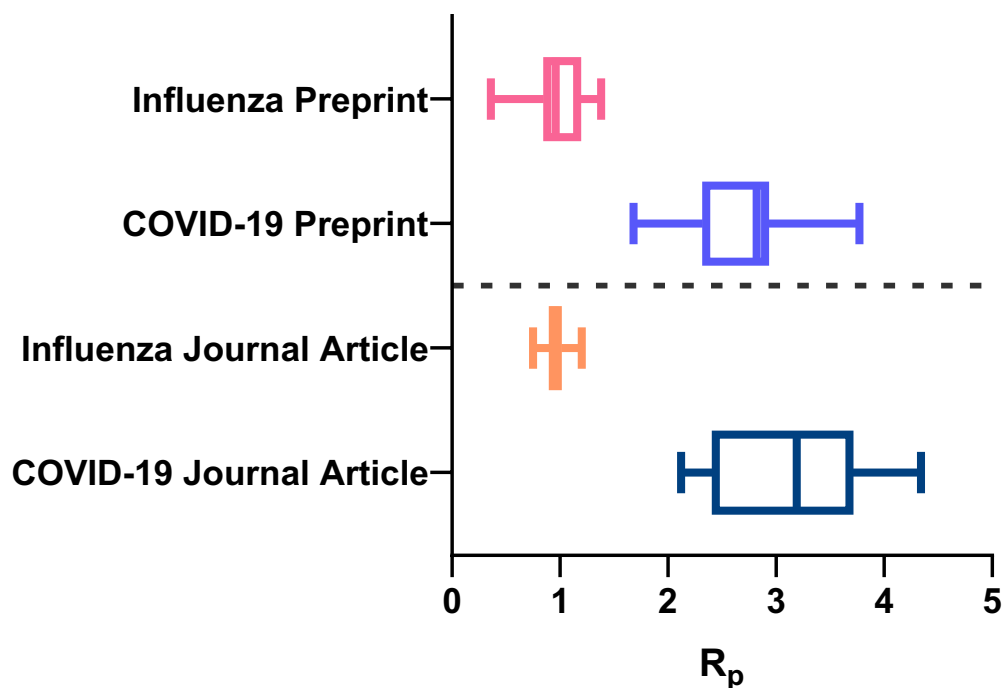


Figura 3. Boxplot da 6^a à 20^a semanas. $P < 0,01$ para artigos de periódicos e pré-prints.

Com relação à R_p para artigos de periódicos e pré-prints, dentro de um modelo putativo em que cada artigo tem influência perpétua sobre os outros, o resultado maior que 1 representou a transmissão como se uma média de 3,17 artigos de periódicos secundários e 2,18 pré-prints secundários tivessem sido gerados de seus predecessores. Comparativamente, o grupo controle do Influenza não apresentou padrões de transmissão, pois seu R_p foi menor que 1 (Figura 4). Após a 20^a semana, o R_p para artigos de periódicos COVID19 aproximou-se de 1 (Figura 5, material suplementar).

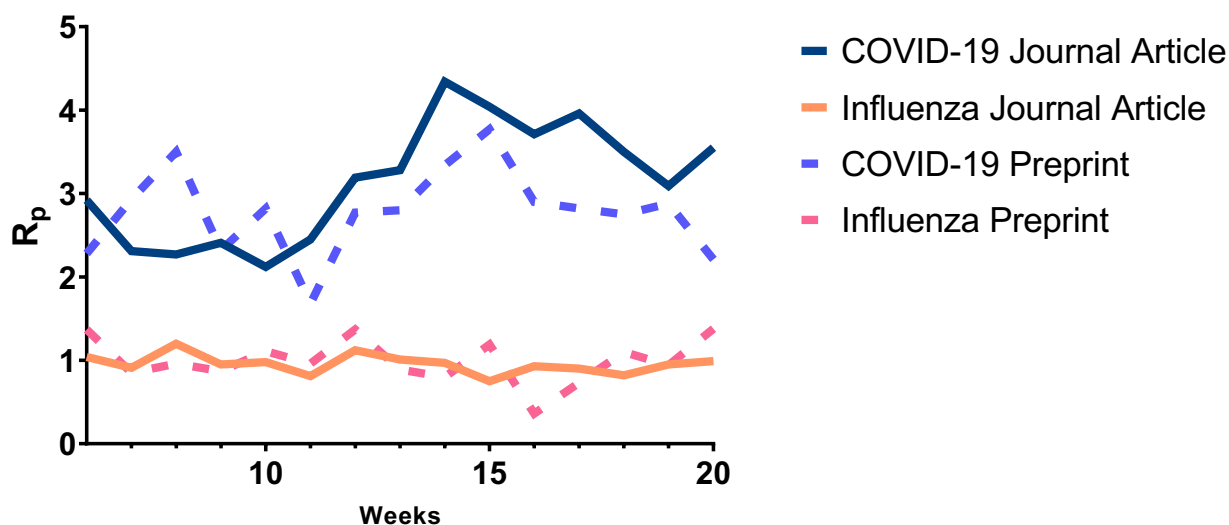


Figura 4. Padrões de comportamento do COVID-19 e Influenza R_p .

6 DISCUSSÃO

Epidemias resultam de uma combinação de transmissibilidade de um agente etiológico, facilitadores ambientais e suscetibilidade de uma população. Nossos resultados sugerem que o excesso de artigos de periódicos e pré-prints relacionados ao COVID19 poderia de fato ser enquadrado como uma infodemia em comparação com o grupo de controle, resultando de (a) um agente muito transmissível, (b) indivíduos suscetíveis expostos em (c) um ambiente em que transmissão foi facilitada por meio de pressão por publicação e reconhecimento, junto com a urgência natural de encontrar soluções para a crise. À luz dos presentes achados, torna-se provável que, nos primeiros meses da pandemia, quanto mais artigos foram publicados, mais os autores alimentaram o ciclo de transmissão publicando ainda mais artigos tanto em periódicos tradicionais quanto em servidores de pré-print, expondo outros à infecção na forma de pressão para publicar.

É interessante notar, no entanto, que tanto as publicações de pré-prints quanto as publicações em periódicos exibiram padrões de crescimento equivalentes, sugerindo que os veículos de publicação podem ter puramente reagido às pressões da comunidade (autores, editores, moderadores). Em outras palavras, do mesmo modo como os artigos sobre Influenza foram publicados em periódicos, eles foram publicados em servidores de pré-prints em uma frequência proporcional, exibida pelas duas curvas de crescimento sobrepostas de padrões de

incidência análogos (Figura 4). Da mesma forma como os artigos relacionados ao COVID19 foram publicados em periódicos, também o foram publicados em servidores de pré-prints, embora seguindo uma curva de crescimento mais acentuada em ambos os locais.

Diante disso, faríamos muito bem em ter em mente que, como uma maçaneta contaminada não é mais culpada na transmissão do que a mão suja que tocou a maçaneta e depois o rosto, os veículos de publicação parecem ter se comportado como a comunidade científica moldou-os para se comportarem. No entanto, como os periódicos se posicionaram nos últimos duzentos anos como curadores oficiais da ciência e, de acordo com nossos achados, eles foram responsáveis pela maior parte da publicação do COVID19, os periódicos poderiam ser compreendidos como os principais culpados por trás da infodemia. Assim, pareceria que o efeito da infodemia contaminou não apenas os autores, mas também editores de revistas indexadas no PubMed.

Pode-se especular, talvez, que os editores de periódicos possam ter decidido acelerar o processo de publicação, publicar mais e, talvez, esperar que a comunidade julgue mais tarde aquele conteúdo, colocando muita pressão na revisão por pares pós-publicação, o que pode ou não ter acontecido e é um tópico para investigação posterior em outros estudos, estando fora do escopo deste experimento mental. Os servidores de pré-prints das ciências da vida e biomédica, por outro lado, desempenharam apenas um papel muito pequeno na infodemia, enquanto os periódicos atuaram como os principais locais de espalhamento, pois são a infraestrutura principal que até hoje domina o ecossistema de publicação científica.

Por outro lado, um maior volume de artigos de periódicos não implica necessariamente que tenham recebido maior atenção, tanto na imprensa quanto no meio científico, para, por si só, responderem por desinformação durante o primeiro ano da pandemia. Nem é possível, dentro do escopo deste experimento, argumentar que os pré-prints não foram responsáveis por mais desinformação do que os artigos em periódicos. Até onde sabemos, não há evidências empíricas fortes que influenciem o argumento para um lado ou para o outro. À luz disto, uma análise futura deve considerar a atenção recebida por artigos publicados e pré-prints na forma de citação, menções nas mídias sociais e cobertura pela imprensa. Uma futura investigação de infodemia faria bem em avaliar e comparar Almetria e desempenho citação para artigos de periódicos e pré-prints empregando também um grupo de controle. Considerando que os pré-prints são facilmente acessíveis, é plausível que as pré-prints também possam ter desempenhado

um papel na disseminação da infodemia por meio das mídias sociais e cobertura da imprensa, mas o debate está longe de ser resolvido e precisa de mais evidências.

A esta altura, sabe-se que mais publicação não implica obrigatoriamente mais conhecimento científico de qualidade. Uma vez que foi encontrada qualidade inferior dos artigos de revistas durante a pandemia por estudos independentes e, devido a isso, a qualidade da revisão por pares deve ser considerada deficiente, a infodemia pode muito bem ter causado uma onda secundária de riscos, como fadiga cognitiva e a possibilidade de boas pesquisas, relevantes e úteis, sobre COVID19 terem sido enterradas sob um excedente de relatos de pesquisa de baixa qualidade. Não afirmamos com isto que todas as pesquisas da era COVID19 resultaram de um desejo de reconhecimento penalizando qualidade, mas nossos achados, juntamente com os achados de outros grupos sobre a qualidade da publicação da pandemia, mostram que a busca por reconhecimento pode ter desempenhado um papel muito forte, fortalecendo a tese da infodemia como um fenômeno informacional comportamental contagioso real entre autores e editores. Não é uma metáfora apenas.

O efeito da infodemia ainda está ativo, não há ainda imunização contra impulsos e pressões para publicar e ser reconhecido. Contudo, o fenômeno não é mais exponencial como foi nos primeiros meses de 2020, o que é um bom sinal. Já que esta não é a primeira vez em que a comunidade científica está em desacordo com a quantidade de publicação que gera e os próprios periódicos foram criados para que os cientistas pudessem lidar com a quantidade de literatura publicada no século XVII, pode-se argumentar que a sobrecarga cognitiva após a explosão bibliográfica exerceu uma pressão adaptativa sobre o ecossistema da publicação científica em direção à profissionalização. A novidade da infodemia de COVID19 é o uso generalizado de servidores de pré-prints, que pode funcionar como outra pressão adaptativa no ecossistema da publicação científica em direção a mais transparência, *accountability* e revisão por pares pós-publicação. Como as evidências mostraram, se os periódicos publicaram muito mais e com pior qualidade do que em tempos pré-pandêmicos, talvez precisemos também de mais ceticismo em relação à revisão por pares pré-publicação e maiores discussões sobre como torná-la mais eficiente.

Ainda assim, à luz dos presentes achados, parece que os veículos de publicação serão o que os autores e editores os moldarão para ser. Logo, não adianta culpar estruturas e tecnologias e não atacar as verdadeiras questões: pressão para publicação e reconhecimento para ascensão na

carreira em detrimento de rigor metodológico e de qualidade como o debate em torno do desperdício evitável em pesquisa tem apontado nestes últimos trinta anos. Assim, como esperamos que nosso experimento mental tenha mostrado, para que os periódicos e os servidores de pré-prints permitam um debate científico público frutífero de alta qualidade por meio de pesquisas confiáveis, tanto os autores quanto os editores precisam colocar qualidade e transparência acima de velocidade e novidade e resistir ativamente ao efeito da infodemia. O mais rápido e o mais original não é necessariamente o de mais qualidade e rigor.

7 CONCLUSÃO

As publicações em COVID19 apresentaram um padrão epidêmico durante os meses iniciais de 2020. Uma vez que o padrão dos artigos revisados por pares se mostrou semelhante ao padrão dos pré-prints, parece que não apenas os autores, mas também os editores foram suscetíveis ao efeito da infodemia. Não foi observado padrão exponencial no grupo controle. Há forte indícios de fenômenos informacionais de origem comportamental tensionando o ecossistema de publicação para longe da racionalidade científica e da eficiência na alocação de recursos informacionais, materiais e humanos. A contribuição deste estudo, assim, para a ciência biomédica é também reforçar a demanda por uma interface desta ciência com outros arcabouços teórico-epistemológicos oferecidos pela metaciência e suas interfaces para melhor compreensão dos fenômenos do campo biomédico e do comportamento de seus agentes.

REFERÊNCIAS

1. Bornmann, Lutz; Mutz, Rüdiger. Growth rates of modern science: A bibliometric analysis based on the number of publications and cited references [Internet]. *Journal of the Association for Information Science and Technology*; 2015. [cited 2020 Out 22]; 6(11): 2215–2222. Available from: doi 10.1002/asi.23329
2. Macleod, Malcolm R. et al. Biomedical research: increasing value, reducing waste [Internet]. *The Lancet*; 2014. [cited 2020 Out 22]; 383(9912): 101-104. Available from: doi: 10.1016/S0140-6736(13)62329-6
3. Ioannidis, John et al. Increasing value and reducing waste in research design, conduct, and analysis [Internet]. *The Lancet*; 2014. [cited 2020 Out 22]; 383(9912):166-175. Available from: doi: 10.1016/S0140-6736(13)62227-8
4. Bouter, Lex; Riet, Gerben. Empirical research must be replicated before its findings can be trusted [Internet]. *Journal of Clinical Epidemiology*; 2021. [cited 2020 Out 22]; 129:188-190. Available from: doi: 10.1016/j.jclinepi.2020.09.032
5. Meadows AJ. *A comunicação científica*. Brasília: Briquet de Lemos; 1998. 268 p.
6. Price DJ de S. *Little science, big science... and beyond*. New York City: Columbia University Press; 1986. 301 p.
7. Guédon, Jean Claude. In *Oldenburg's Long Shadow - librarians, research scientists, publishers, and the control of scientific publishing*. Washington, DC: Association of Research Libraries, 2001. 69p.
8. Larivière V, Haustein S, Mongeon P. The Oligopoly of Academic Publishers in the Digital Era [Internet]. *PLoS ONE*; 2015. [cited 2020 Out 22]; 10(6): e0127502. Available from: doi: 10.1371/journal.pone.0127502
9. Altman, Douglas. The scandal of poor medical research [Internet]. *BMJ*; 1994. [cited 2020 Out 22]; 308:283. Available from: doi: 10.1136/bmj.308.6924.283
10. Ioannidis, John. Why Most Published Research Findings Are False [Internet]. *PLoS Med*; 2005. [cited 2020 Out 22]; 2(8): e124. Available from: doi.org/10.1371/journal.pmed.0020124
11. Brembs B. Prestigious Science Journals Struggle to Reach Even Average Reliability [Internet]. *Front. Hum. Neurosci*; 2018. [cited 2020 Out 22]; 12:37. Available from: doi: 10.3389/fnhum.2018.00037
12. Correia LC, Barreto Segundo JD. An immunization program against the COVID-19 infodemic [Internet]. *J Evid-Based Healthc*; 2020. [cited 2020 Out 22]; 2(1):7-9. Available from: doi: 10.17267/2675-021Xevidence.v2i1.3124

13. Horbach SPJM. Pandemic Publishing: Medical journals strongly speed up their publication process for Covid-19. *Quantitative Science Studies* [Internet]; 2020. [cited 2020 Out 22]; 1-16. Available from: doi: 10.1162/qss_a_00076
14. Helliwell JA, Bolton WS, Burke JR, Tiernan JP, Jayne DG, Chapman SJ. Global academic response to COVID - 19: Cross-sectional study [Internet]. *Learned Publishing*; 2020. [cited 2020 Out 22]. Available from: doi: 10.1002/leap.1317
15. Jung RG, Santo PD, Clifford C, Prospero-Porta G, Skanes S, Hung A et al. Methodological Rigor in COVID-19 Clinical Research: A Systematic Review and Case-Control Analysis [Internet]. *Nat Commun*; 2021. [cited 2020 Out 22]; 12:943. Available from: doi: 10.1038/s41467-021-21220-5
16. Barakat AF, Shokr M, Ibrahim J, Mandrola J, Elgendy IY. Timeline from receipt to online publication of COVID-19 original research articles [Internet]. *medRxiv* 2020.06.22.20137653. [cited 2020 Out 22]. Available from: doi: 10.1101/2020.06.22.20137653
17. Palayew A, Norgaard O, Safreed-Harmon K, Andersen TH, Rasmussen LN, Lazarus JV. Pandemic publishing poses a new COVID-19 challenge [Internet]. *Nat Hum Behav*; 2020. [cited 2020 Out 22]. Available from: doi: 10.1038/s41562-020-0911-0
18. Janiaud P, Axfors C, van't Hooft J et al. The worldwide clinical trial research response to the COVID-19 pandemic - the first 100 days [version 1; peer review: awaiting peer review] [Internet]. *F1000Research*; 2020. [cited 2020 Out 22]. 9:1193. Available from: doi: 10.12688/f1000research.26707.1)
19. Horbach SPJM. No time for that now! Qualitative changes in manuscript peer review during the Covid-19 pandemic [Internet]. *Research Evaluation*; 2021. [cited 2020 Out 22]; rvaa037:1-9. Available from: <https://doi.org/10.1093/reseval/rvaa037>
20. Ioannidis JPA, Salholz-Hillel M, Boyack KW, Baas J. The rapid, massive infection of the scientific literature and authors by COVID-19 [Internet]. 2020. [cited 2020 Out 22]. Available from: doi: 10.1101/2020.12.15.422900
21. Sevryugina YV, Dicks AJ. Publication practices during the COVID-19 pandemic: Biomedical preprints and peer-reviewed literature [Internet]. *bioRxiv*; Posted January 21, 2021. [cited 2020 Out 22]. Available from: doi: 10.1101/2021.01.21.427563
22. Merton RK. The Matthew Effect in Science: The reward and communication systems of science are considered. *Science* [Internet]. 1968 Jan 5. [cited 2020 Out 22]; 159(3810):56–63. Available from: doi 10.1126/science.159.3810.56
23. Price DDS. A general theory of bibliometric and other cumulative advantage processes [Internet]. *Journal of the American Society for Information Science*; 1976. [cited 2020 Out 22]; 27(5):292–306. Available from: doi.org/10.1002/asi.4630270505

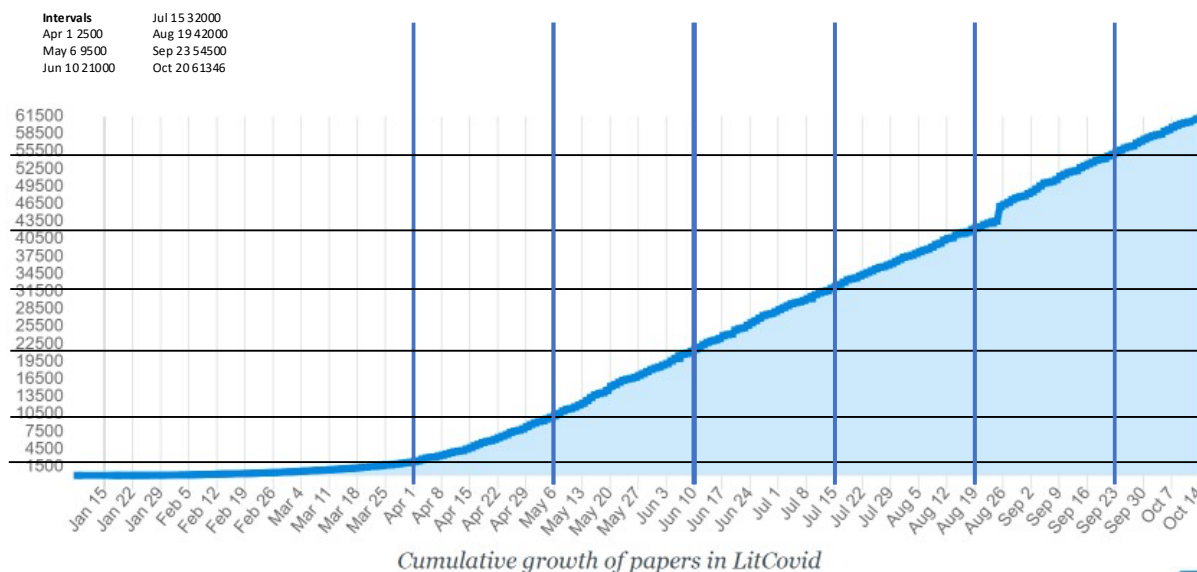
24. Dinis-Oliveira, RJ. COVID-19 research: pandemic versus “paperdemic”, integrity, values and risks of the “speed science” [Internet]. *Forensic Sciences Research*; 2020. [cited 2020 Out 22]; 5(2): 174-187. Available from: doi: 10.1080/20961790.2020.1767754
25. Rawlinson C, Bloom T. New preprint server for medical research [Internet]. *BMJ*; 2019. [cited 2020 Out 22]; 365:l2301. Available from: doi: 10.1136/bmj.l2301
26. Anderson, K.R. bioRxiv: Trends and analysis of five years of preprints. *Learned Publishing* [Internet]. 2020. [cited 2020 Out 22]; 33:104-109. Available from: doi: 10.1002/leap.1265
27. Else, Holly. How a torrent of COVID science changed research publishing — in seven charts [Internet]. *Nature*; 2020. [cited 2020 Out 22]; 588, 553. Available from: doi: 10.1038/d41586-020-03564-y
28. Fraser N et al. Preprinting the COVID-19 pandemic [Internet]. *bioRxiv*. Posted February 05, 2021; [cited 2020 Out 22]; Available from: doi 10.1101/2020.05.22.111294
29. Pfeiffer, Thomas; Bertram, Lars; Ioannidis, John (2011). "Quantifying selective reporting and the Proteus phenomenon for multiple datasets with similar bias" [Internet]. *PLoS ONE*. [cited 2020 Out 22]; 6 (3): e18362. Available from: doi: 10.1371/journal.pone.0018362.
30. Button, Katherine S.; et al. (2013). "Power failure: why small sample size undermines the reliability of neuroscience" [Internet]. *Nature Reviews Neuroscience*. [cited 2020 Out 22]; 14 (5): 365–376. Available from: doi:10.1038/nrn3475. PMID 23571845.
31. Ioannidis, JP; Trikalinos, TA (2005). "Early extreme contradictory estimates may appear in published research: The Proteus phenomenon in molecular genetics research and randomised trials" [Internet]. *Journal of Clinical Epidemiology*. [cited 2020 Out 22]; 58: 543–549. Available from: doi: 10.1016/j.jclinepi.2004.10.019.
32. Peterson D, Panofsky A. Metascience as a scientific social movement [Internet]. *SocArXiv*; 2020. [cited 2020 Out 22]; Available from: osf.io/preprints/socarxiv/4dsqa.
33. Bianchetti, L., Zuin, A. A. S., & Ferraz, O. Publique, apareça ou pereça: produtivismo acadêmico, pesquisa administrativa e plágio nos tempos da cultura digital. In *Publique, apareça ou pereça: produtivismo acadêmico, pesquisa administrativa e plágio nos tempos da cultura digital*. EDUFBA; 2018. [cited 2020 Out 22]. Available from: <https://doi.org/10.7476/9788523220181>.
34. Neves, Kleber & Tan, Pedro Batista & Amaral, Olavo Bohrer, 2021. "Are Most Published Research Findings False In A Continuous Universe? [Internet]," *MetaArXiv jk7sa*; Center for Open Science; [cited 2020 Out 22]. Available from: doi 10.31219/osf.io/jk7sa

35. John P. A. Ioannidis. Contradicted and Initially Stronger Effects in Highly Cited Clinical Research [Internet]. JAMA J Am Med Assoc; 2005. [cited 2020 Out 22]; 294(21):2695–2695. Available from: doi 10.1001/jama.294.2.218
36. Nissen SB, Magidson T, Gross K, Bergstrom CT. Publication bias and the canonization of false facts [Internet]. Elife; 2016. [cited 2020 Out 22]; 5: e21451. Published 2016 Dec 20. Available from: doi:10.7554/eLife.21451
37. King, D. The scientific impact of nations [Internet]. Nature **430**; 2004. [cited 2020 Out 22]; 311–316. Available from: <https://doi.org/10.1038/430311a>.
38. OLIVEIRA, É. B. P. M. Periódicos científicos eletrônicos: definições e histórico. **Inf&Soc**, v.18, n. 2, p. 69-77, 2008.
39. Dietz K. The estimation of the basic reproduction number for infectious diseases [Internet]. Stat Methods Med Res; 1993. [cited 2020 Out 22]. 2:23–41. Available from: doi: 10.1177/096228029300200103.

MATERIAL SUPLEMENTAR

Figura 1 – Incidência cumulativa plotada online em:

<https://www.ncbi.nlm.nih.gov/research/coronavirus/faq> – Acessado em 14 de outubro, 2020.



A interseção acima foi transposta para a tabela abaixo, preenchida com os dados dos mesmos pontos pesquisados no PubMed, esses dados também podem ser encontrados na planilha do conjunto de dados na seção suplementar online:

Tabela 1 – Comparação de dados de número de publicação por plataforma. A diferença absoluta e a razão de comparação média por lapso de tempo foram calculadas

Date	LitCovid	PubMed	Diference	Mean
01-Apr	2500	5957	3457	2.3828
06-May	9500	14162	4662	1.490736842
10-Jun	21000	26320	5320	1.253333333
15-Jul	32000	38258	6258	1.1955625
19-Aug	42000	48708	6708	1.159714286
23-Sep	54500	59771	5271	1.096715596
20-Oct	61346	68212	6866	1.111922538
Total	222846	261388	5506	1.384397871

Number of publications registered. 1.172953519

Gabriel Grisi:
This is the avarege difference.

Gabriel Grisi:
This is the avarege a PubMed/LitCovid ratio. It is how many times, in average, PubMed publications number are higher than LitCovid results.

Gabriel Grisi:
Average of total number of PubMed/LitCovid Publications.

Figura 2 – Gráfico de incidência semanal por data de busca.

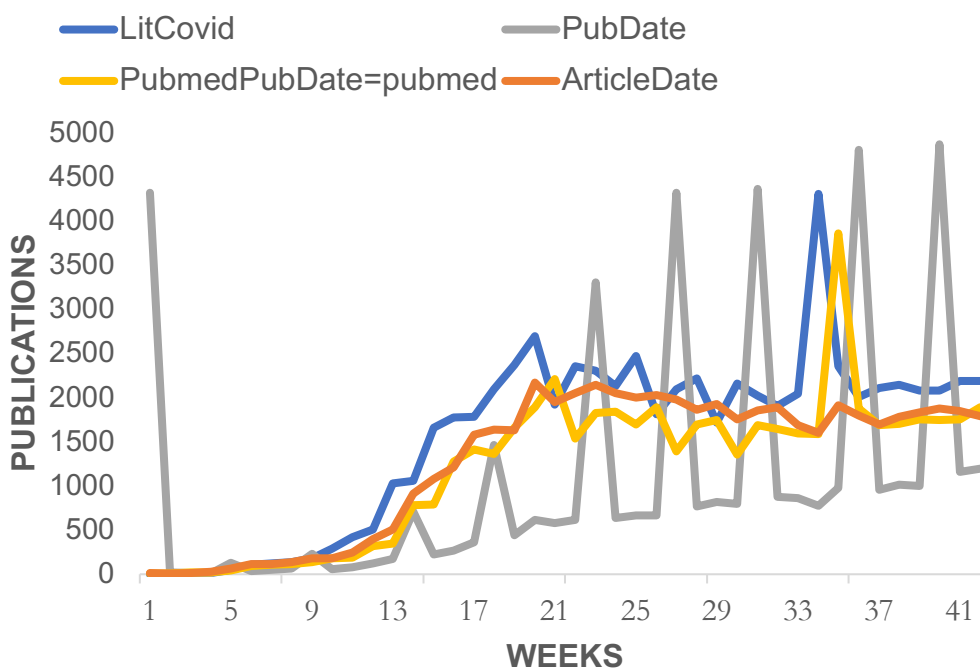


Figura 3 – 20^a a 56^a incidência semanal.

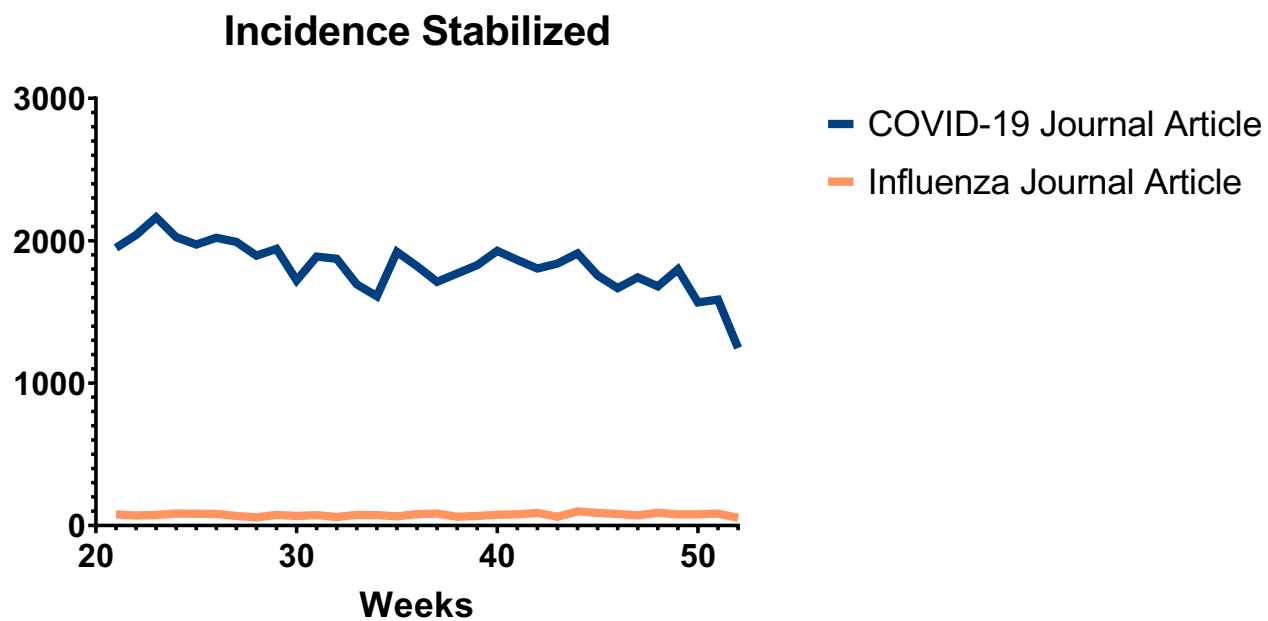


Figura 4 – Número cumulativo das semanas 20 a 56. Min. = 1246; máx. 2134; SD 233,8

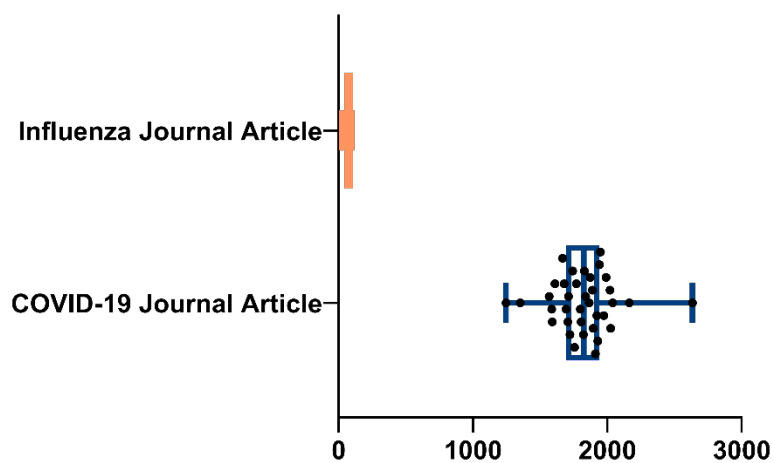


Figura 5 – Rp das semanas 20 a 56

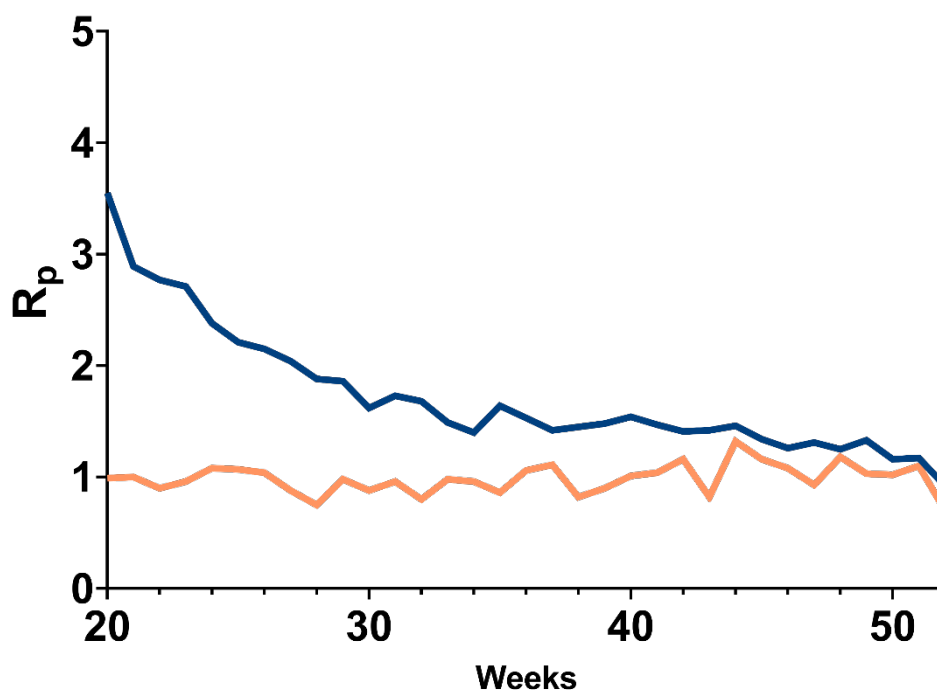


Figura 6 – Testes estatísticos de incidência e Poder

COVID JOURNAL ARTICLE Model Summary and Parameter Estimates

Dependent Variable: Cumulative Incidence

Equation	R Square	Model Summary				Sig.	Parameter Estimates			
		F	df1	df2	Constant		b1	b2	b3	
Quadratic	.997	1827.603	2	12	.000	5886.067	-1371.651	83.839		
Cubic	.999	3978.827	3	11	.000	678.724	27.372	-31.368	2.954	
Power	.985	830.274	1	13	.000	.437	3.369			
Growth	.995	2776.838	1	13	.000	3.907	.284			

The independent variable is Week.

COVID PREPRINT Model Summary and Parameter Estimates

Dependent Variable: Cumulative Incidence

Equation	R Square	Model Summary				Sig.	Parameter Estimates			
		F	df1	df2	Constant		b1	b2	b3	
Quadratic	.998	3443.829	2	12	.000	876.848	-234.738	18.586		
Cubic	.998	2182.445	3	11	.000	625.811	-167.293	13.032	.142	
Power	.997	4394.614	1	13	.000	.321	3.121			
Growth	.969	409.566	1	13	.000	3.316	.258			

The independent variable is Week.

ANOVA - Power

	Sum of Squares	df	Mean Square	F	Sig.
Regression	19.230	1	19.230	4394.614	.000
Residual	.057	13	.004		
Total	19.287	14			

The independent variable is Week.