



MODELAGEM DE TÓPICOS DE PESQUISA SOBRE O NOVO CORONAVÍRUS: APLICAÇÃO DO LATENT DIRICHLET ALLOCATION

Tatiane Barateri*
Maicon Henrique Lentsck**
Clérito Kaveski Peres***
Erica de Brito Pitilin****

RESUMO

Objetivo: analisar o desenvolvimento global de publicações científicas e construir tópicos abordados em pesquisas sobre o novo coronavírus/COVID-19. **Métodos:** estudo de revisão de resumos de publicações sobre o novo coronavírus/COVID-19, com busca nas bases de dados *Embase* e *Web of Science*, em 03 de novembro de 2020. A análise dos dados se deu pela modelagem dos tópicos de pesquisa pelo método *Latent Dirichlet Allocation* (LDA). **Resultados:** analisaram-se 31.769 publicações, com envolvimento de pesquisadores de 182 países. Os estudos foram publicados em 5.336 periódicos diferentes. A análise identificou 20 tópicos principais de pesquisa sobre o novo coronavírus/COVID-19, sendo que mais da metade (54,5%) das publicações foram contempladas nos temas de impactos e desafios decorrentes da pandemia (11,3%), resposta imunológica (9,6%), potenciais terapêuticas (9,0%), epidemiologia (7,8%), saúde mental (6,3%), recomendações para assistência a outras necessidades de saúde (5,3%) e complicações clínicas. O tema da telemedicina e telessaúde foi evidenciado em 4,2% das publicações. **Conclusões:** as publicações científicas estavam distribuídas dentre 20 principais temas, caracterizando sua importância e relevância no entendimento do agravamento em questão para traçar estratégias de controle, prevenção, tratamento e atenção aos agravos/implicações decorrentes da pandemia. Identificaram-se como principal lacuna de conhecimento estudos sobre cuidados primários/atenção primária à saúde.

Palavras-chave: Infecções por Coronavírus. Pandemias. COVID-19. Revisão.

INTRODUÇÃO

As infecções virais têm sido devastadoras ao longo dos anos. Em 2003, a epidemia grave da síndrome respiratória aguda (SARS) afetou cerca de 8.000 pessoas, matou 780 e causou uma enorme crise social e econômica. Em 2006, um novo H5N1 aviário, em 2009 uma nova ameaça de pandemia de influenza H1N1 'suína' causou ansiedade global⁽¹⁾. Situações como poliomielite (2014), ebola na África Ocidental (2014), zika (2016) e ebola na República Democrática do Congo (2019) foram consideradas emergência de saúde pública de interesse internacional pela Organização Mundial da Saúde (OMS)⁽²⁾.

Entretanto, nenhuma tomou a proporção, em curso, nominada pela OMS como 2019-nCoV: doença de coronavírus (COVID-19)⁽²⁾, que no final de dezembro de 2019 provocou um surto de pneumonia de etiologia desconhecida em Wuhan,

província de Hubei, China, e se espalhou rapidamente por todo o país⁽³⁾. O Centro Chinês de Controle e Prevenção de Doenças identificou um novo beta-coronavírus chamado 2019-nCoV, agora oficialmente conhecido como *coronavírus da síndrome respiratória aguda grave 2* (SARS-CoV-2)⁽³⁾, responsável pela atual pandemia.

Esse foi o terceiro surto de coronavírus zoonótico nas duas primeiras décadas do século XXI, permitindo a transmissão de humano para humano e aumentando as preocupações globais com a saúde⁽⁴⁾, visto que já foram identificados mais de 46 milhões de casos e mais de 1,2 milhão de óbitos até 03 de novembro de 2020⁽⁵⁾.

A pandemia exige para um futuro próximo ações em muitas frentes, uma vez que ela é combatida por várias estratégias, que vão desde a prevenção, fortalecimento do diagnóstico precoce e tratamento com drogas e vacinas eficazes capazes

*Enfermeira. Doutora. Departamento de Enfermagem da Universidade Estadual do Centro-Oeste – UNICENTRO. Guarapuava, Paraná, Brasil. E-mail: barateri.tatiane@gmail.com. ORCID ID: 0000-0002-0270-6395.

**Enfermeiro. Doutor. Departamento de Enfermagem da UNICENTRO. Guarapuava, Paraná, Brasil. E-mail: maiconlentsck@yahoo.com.br. ORCID ID: 0000-0002-8912-8902

***Engenheiro de Produção. Mestre. Doutorando pelo Programa de Pós-Graduação em Engenharia de Produção da Universidade Federal de Santa Catarina – UFSC.

Florianópolis, Santa Catarina, Brasil. E-mail: cleritokp@gmail.com. ORCID ID: 0000-0002-5643-6888.

****Enfermeira. Doutora. Departamento de Enfermagem da Universidade Federal da Fronteira Sul. Chapecó, Santa Catarina, Brasil. E-mail: erica.pitilin@gmail.com. ORCID

ID: 0000-0003-3950-2633.

de neutralizar o vírus. Nesse sentido, a razão e a ciência devem ser o norte. Contudo, à medida que a pandemia aumenta e o vírus se torna endêmico, percebe-se a necessidade de uma abordagem estruturada e bem coordenada em vários setores, como prevenção de contágio e complicações da COVID-19 em populações especiais, como gestantes e doentes crônicos; no âmbito social, com o aumento da pobreza e fome; e no âmbito econômico, forçando a repensar o estilo de vida⁽⁶⁾. Assim, para enfrentar temas no âmbito da saúde, faz-se necessário e urgente avaliar criticamente as evidências disponíveis.

A COVID-19 é agora uma crise global de saúde, de informação e cada vez mais uma crise socioeconômica ampla⁽⁵⁾. Por isso, as autoridades de saúde pública devem continuar monitorando a situação e constantes investimentos devem ser feitos em pesquisas, considerando que quanto maior conhecimento produzido sobre esse novo vírus e seus surtos associados, melhor será a capacidade de resposta⁽²⁾.

A urgência no combate à pandemia e manejo de outros problemas/agraves decorrentes dela exige a busca pelas melhores evidências científicas, necessitando ser sumarizadas e atualizadas para auxiliar na tomada de decisões. Essa necessidade tem gerado um número significativo de pesquisas na área e, conseqüentemente, sua publicação em revistas científicas de diversos países, as quais têm fomentado a divulgação de estudos sobre o tema.

Em razão do grande volume de publicações sobre o novo coronavírus, questiona-se: quais os tópicos mais abordados para a possível identificação de lacunas de conhecimento sobre o novo coronavírus? A fim de responder e evidenciar os aspectos mais importantes considerados na realização de um estudo sobre o tema, objetivou-se analisar o desenvolvimento global de publicações científicas e construir tópicos abordados em pesquisas sobre o novo coronavírus/COVID-19.

MÉTODO

Conjunto de dados

Trata-se de um estudo de revisão do tipo *scoping review* realizado nas bases de dados *Embase* e *Web of Science*. A escolha dessas bases justifica-se por sua abrangência global e por permitirem a coleta de grande volume de dados em

formato .csv, viabilizando a análise proposta pelo presente estudo.

A coleta de dados se deu em 03 de novembro de 2020, pesquisando as palavras-chave: "COVID-19", "COVID 19", "COVID19", "2019 novel coronavirus disease", "novel coronavirus 2019", "SARS-CoV-2", "2019 novel coronavirus infection", "2019-nCoV", "2019 nCoV", "2019nCoV", "coronavirus disease 2019", "coronavirus disease-19", nos títulos das publicações, utilizando o operador booleano OR.

Incluíram-se resumos de estudos publicados em inglês, nos anos de 2019 e 2020, classificados como artigo ou revisão, a fim de abarcar pesquisas originais e de revisão, bem como limitar a inclusão de cartas, editoriais, notas, entre outros. Excluíram-se estudos que apresentaram os campos "endereço do autor" e "resumo" vazios, como também estudos duplicados nas bases.

Realizou-se o *download* dos dados em formato .csv, incluindo as seguintes informações: título da publicação; autores; nome do periódico; ano de publicação; palavras-chave; endereço do autor; e resumo. Dois pesquisadores trabalharam de forma independente para baixar simultaneamente os dados. Realizou-se o *linkage* dos bancos de dados baixados das bases e analisados por meio da linguagem *Python* (versão 3.7.5), via plataforma *Jupyter Notebook*.

Análise dos dados

Inicialmente, realizou-se uma análise descritiva de todos os artigos que atenderam aos critérios de elegibilidade com o intuito de identificar os países com maior número de pesquisadores participantes das publicações e as revistas que mais publicam sobre o tema. A identificação do número de pesquisadores por país se deu por meio do campo de endereço do autor. Para tanto, utilizaram-se as bibliotecas *PyCountry* (<https://pypi.org/project/pycountry/>) e *Plotly* (<https://plotly.com/>), disponíveis na linguagem *Python* (versão 3.7.5).

Em relação à análise de tópicos, o estudo passou pelas etapas de pré-processamento dos dados com os processos de limpeza, transformação e etapa de modelagem de tópicos.

O pré-processamento do texto foi empregado a fim de organizar, limpar e padronizar os dados textuais para ser consumido pelo algoritmo de

processamento. O intuito desse procedimento é remover o conteúdo desnecessário para reduzir o tempo de processamento e aumentar a performance dos modelos⁽⁷⁾. As técnicas utilizadas para o pré-processamento do conjunto de dados foram a remoção de caracteres especiais, conversão de letras maiúsculas em minúsculas, remoção de palavras sem valor semântico (*stopwords*), remoção de outros termos desnecessários e aplicação de tokenização, que quebra o texto em palavras. Essas tarefas foram realizadas por meio da linguagem *Python* (versão 3.7.5) com apoio da biblioteca *Natural Language Toolkit* (NLTK) (<https://www.nltk.org/>).

A modelagem de tópicos foi empregada com a finalidade de descobrir quais temas mais têm mobilizado os cientistas. Para isso, utilizou-se o método *Latent dirichlet allocation* (LDA). Trata-se de um algoritmo para detecção automática de tópicos em textos que pertence à classe de métodos não supervisionados. O algoritmo do LDA se baseia na ideia de que os documentos são compostos por uma mistura de tópicos e busca descobrir padrões latentes (escondidos) para entender a relação entre documentos e palavras, então as palavras que ocorrem em documentos relacionados são agrupadas em tópicos⁽⁸⁾. O método LDA foi implementado por meio do pacote para linguagem *PythonMachine Learning for Language Toolkit* (MALLET) (<http://mallet.cs.umass.edu/>).

Os dados foram transformados em vetores para posterior entrada nos algoritmos utilizando o método de engenharia de atributos *Bag of Words* (BoW). O modelo BoW representa cada documento de texto como um vetor numérico em que cada dimensão é uma palavra específica do *corpus* e o valor é sua frequência no documento⁽⁷⁾.

O modelo foi também ajustado para remover termos com menos de 10 ocorrências em todos os documentos e termos que ocorrem em mais de 60% dos documentos, reduzindo de 36.672 para 8.239 termos. Essa prática aumenta a precisão do modelo e reduz o esforço computacional, já que termos únicos ou raros têm baixa contribuição semântica para o modelo e termos recorrentes interferem na interpretação do contexto⁽⁷⁾.

O número ótimo de tópicos foi definido com base no *coherence score*. Para tanto, geraram-se modelos com um número de tópicos variando de 2 a 50. A Figura 1 mostra a variação do *coherence score* conforme o número de tópicos, evidenciando-se que a maior coerência (0,642) se deu com 33 tópicos, sendo este o valor adotado para este estudo. Dessa forma, todas as sentenças do conjunto de dados foram classificadas em algum dos 29 tópicos revelados pelo LDA, de acordo com a maior probabilidade de pertencerem a cada um dos tópicos.



Figura 1. Variação da coerência do modelo de acordo com o número de tópicos (03 de novembro de 2020)

Para a identificação dos temas abordados nas pesquisas, os tópicos foram rotulados manualmente com base na combinação de julgamento humano com referência à literatura sobre o tema de pesquisa, de modo que os que apresentaram semelhanças foram unidos. Três

tópicos que não apresentaram consistência de palavras foram excluídos da rotulagem, apresentando-se, ao final da análise, 20 tópicos de pesquisa.

RESULTADOS

A pesquisa inicial identificou 57.389 publicações, sendo 34.498 da base de dados *Embase* e 22.888 da *Web of Science*. Excluíram-se 7.940 publicações duplicadas e 17.677 por não

preenchimento do campo “endereço do autor” e/ou “resumo”. Ao final, analisaram-se 31.769 resumos (Figura 2), os quais foram publicados em 5.336 periódicos diferentes.

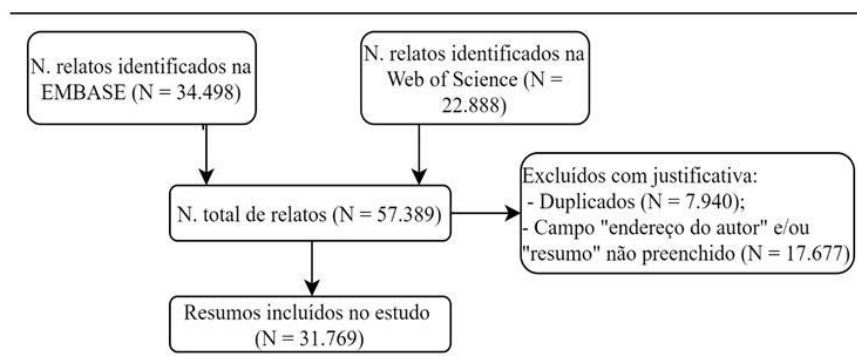


Figura 2. Fluxograma do processo de seleção dos estudos (03 de novembro de 2020)

A Figura 3 apresenta a distribuição dos pesquisadores de acordo com os 182 países envolvidos em estudos sobre o tema analisado. O país com o maior número de pesquisadores

envolvidos foram os Estados Unidos (19,6%), seguido da China (10,6%), Itália (7,4%), Reino Unido (5,9%) e Índia (5,7%). O Brasil ocupou a 11ª posição com 2,2% dos pesquisadores.

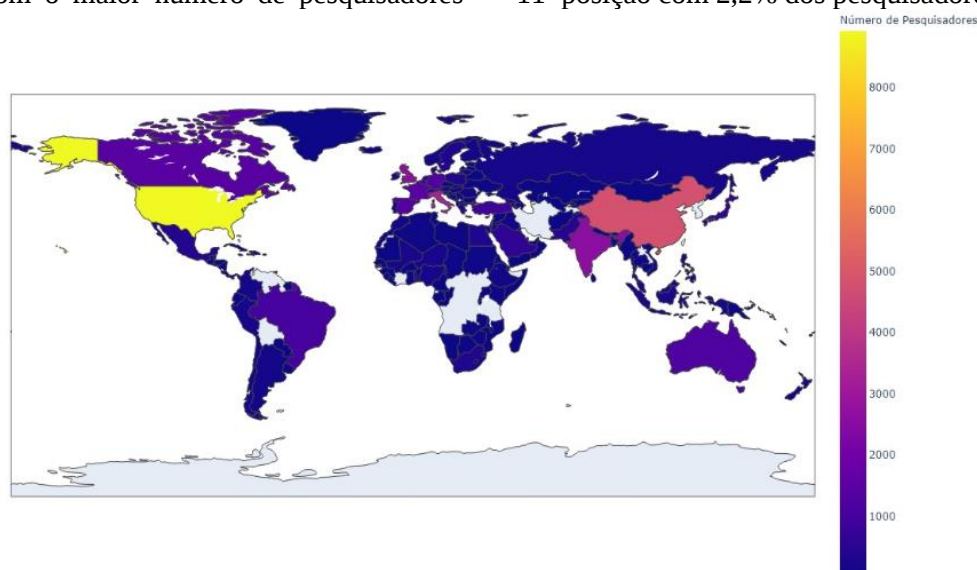


Figura 2. Distribuição do número de pesquisadores de acordo com o país (03 de novembro de 2020)

A Tabela 1 apresenta os 20 tópicos abordados em pesquisas científicas construídos pelo LDA. Considerando a exclusão de três tópicos que não puderam ser rotulados devido à inconsistência do conjunto de palavras, compostos por 1.584 (4,9%) resumos analisados, apresentaram-se os tópicos de pesquisa de 30.185 artigos.

Verifica-se que mais da metade (54,5%) das publicações estão contempladas entre os sete primeiros tópicos de análise. O tema mais abordado foi sobre impactos e desafios

decorrentes da pandemia (11,3%), seguido de publicações sobre temas importantes para o desenvolvimento de mecanismos de prevenção, diagnóstico e tratamento (Tópicos 2 e 3) e dos aspectos epidemiológicos envolvidos (7,8%). Quanto às áreas específicas da atenção à saúde, verificou-se que a saúde mental (Tópico 5) foi a especialidade mais abordada (6,3%). Também cabe destaque ao tema da telemedicina e telessaúde (Tópico 11), que foi evidenciado em 4,2% das publicações analisadas.

Tabela 1. Tópicos abordados em pesquisas sobre o novo coronavírus (03 de novembro de 2020)

Tópico do LDA	N	%	Rótulo do tópico
1	3421	11,3	Impactos e desafios
2	2892	9,6	Resposta imunológica
3	2730	9,0	Potenciais terapêuticas
4	2357	7,8	Epidemiologia
5	1894	6,3	Saúde mental
6	1603	5,3	Recomendações para assistência à saúde integral em tempos de COVID-19
7	1570	5,2	Complicações clínicas da COVID-19
8	1346	4,5	Manifestações clínicas
9	1304	4,3	Segurança dos trabalhadores de saúde e pacientes
10	1300	4,3	Avaliação pulmonar
11	1270	4,2	Uso da telemedicina e telessaúde
12	1185	3,9	Formação de profissionais de saúde
13	1186	3,9	Hospitalização e terapia intensiva
14	1139	3,8	Comorbidades associadas
15	1061	3,5	Mecanismo de ação do vírus
16	1048	3,5	Desenvolvimento de vacinas
17	924	3,1	Atenção à saúde das gestantes
18	836	2,8	Estratégias governamentais para controle da transmissão do vírus
19	627	2,1	Atenção à saúde das crianças
20	492	1,6	Conscientização da população

O Quadro 1 apresenta uma síntese descritiva dos 20 tópicos analisados. dos temas abordados pelos estudos em cada um

Quadro 1. Síntese descritiva dos tópicos (03 de novembro de 2020)

Tópico	Engloba estudos que tratam de:
1	Impactos e desafios, em curto e longo prazo, nos sistemas e diferentes setores da saúde, socioeconômicos, meio ambiente, condições/situações de saúde/doença específicas (por exemplo: hemodiálise, saúde mental, situações de emergência) causados pelo novo coronavírus/COVID-19 e/ou das mudanças decorrentes da pandemia.
2	Análise da resposta imunológica perante o novo coronavírus para controle e tratamento da doença, desenvolvimento de medicamentos e vacinas.
3	Desenvolvimento e testagem de drogas para prevenção e tratamento.
4	Aspectos epidemiológicos do surgimento do novo coronavírus, cenário de disseminação/crescimento, incidência, taxas de transmissão e mortalidade, modelagem de possíveis cenários.
5	Alterações psicológicas e desenvolvimento de estresse, ansiedade e/ou depressão na população e em profissionais de saúde; implicações em pacientes com problemas/doenças mentais preexistentes à pandemia.
6	Diretrizes clínicas para assistência à saúde em situações agudas ou crônicas durante a pandemia, como câncer, atendimento de urgência/emergência, procedimentos cirúrgicos.
7	Complicações em órgãos (rim, fígado, coração...) e sistemas em pacientes com ou sem patologias prévias.
8	Sintomas e manifestações clínicas em pessoas com o novo coronavírus.
9	Necessidade e qualidade de equipamentos de proteção individual em diferentes setores da saúde e estabelecimento de protocolos para cirurgias de emergência e eletivas em pacientes não contaminados pelo novo coronavírus, para proteção e segurança de profissionais de saúde e pacientes.
10	Avaliação das características e achados de exames de imagem de tórax em pacientes com COVID-19.
11	Uso da telemedicina e telessaúde em diferentes cenários/condições como meio de garantir acesso à assistência à saúde e prevenir contaminação.
12	Programas de treinamento e ensino a distância de profissionais de saúde em período de pandemia, principalmente para medicina e enfermagem.
13	Desafios e necessidade de hospitais e de unidades de terapia intensiva; diretrizes clínicas para terapia intensiva; estudos que abordam características e prognóstico de pacientes hospitalizados e/ou em terapia intensiva.
14	Análise de prevalência, desfechos, risco e severidade da COVID-19 em pacientes com comorbidades associadas.
15	Conhecimento sobre os mecanismos de ação do novo coronavírus e suas relações com características clínicas e desenvolvimento de terapêuticas para prevenção e tratamento.

Continua...

Tópico	Engloba estudos que tratam de:
16	Desenvolvimento de vacinas para imunização contra o novo coronavírus.
17	Prevalência, riscos, agravos associados e recomendações para atenção às gestantes; cuidados durante o parto.
18	Implementação de estratégias governamentais para controlar a pandemia, como isolamento social e quarentena.
19	Análise de sinais/sintomas em crianças e assistência em pediatria.
20	Educação e conscientização da população sobre a importância da adoção de medidas individuais de combate ao novo coronavírus.

DISCUSSÃO

Ao analisar publicações científicas sobre o novo coronavírus e modelar seus tópicos de pesquisa, evidenciou-se um número expressivo de publicações em nível global, com maior participação de pesquisadores dos Estados Unidos e China, sendo esses países as duas principais potências econômicas mundiais atualmente. Ademais, China foi o país onde iniciou a pandemia e os Estados Unidos apresentam-se como o país com mais infectados e mortos, fatores esses que podem estar motivando investimentos em pesquisa.

As publicações estiveram distribuídas em periódicos de diversas áreas, o que indica a preocupação de diferentes campos em compreender e encontrar soluções para os problemas decorrentes da pandemia em distintos setores.

Identificaram-se 20 tópicos principais de pesquisa, de modo que o tema mais abordado se refere aos impactos e desafios, considerando que a COVID-19 surgiu como uma ameaça à saúde global e cada nação tem enfrentado desafios únicos. A pandemia levou à necessidade de reestruturação dos sistemas de saúde para oferta de atenção em tempo oportuno às pessoas com COVID-19, afetando, por outro lado, pacientes que requerem assistência por outros agravos, gerando atrasos no atendimento e tratamento, bem como necessidade de reorganização dos serviços para o atendimento às diferentes especialidades⁽⁹⁻¹³⁾.

A pandemia também tem causado intensos impactos econômicos pela necessidade de recursos emergenciais e queda na arrecadação de impostos⁽¹⁴⁾, além de revelar as estreitas interconexões entre desigualdade social/econômica, indicando que políticas que promovem a solidariedade social também poderão melhorar a saúde da população e o desempenho econômico para o melhor

enfrentamento de futuras situações de emergenciais⁽¹⁵⁾.

A abordagem de temas que envolvem impactos e desafios diante da pandemia expressa a preocupação de buscar uma normalidade para a desestruturação sem precedentes em uma sociedade. Epidemias globais não são novas, mas a COVID-19 encontra-se em um mundo cada vez complexo, tecnológico e conectado, em que eventos de risco significam maneiras diferentes e inesperadas de enfrentamento⁽¹⁶⁾.

Os estudos sobre o novo coronavírus também têm abordado de forma relevante os aspectos epidemiológicos da pandemia, desde o seu surgimento até seu alcance global, conhecendo as diferenças e semelhanças nas taxas de transmissão, incidência e mortalidade, para auxiliar na elaboração de estratégias de enfrentamento, como as governamentais (Tópico 18) e conscientização da população (Tópico 20).

Nesse tópico (Tópico 4), ressalta-se a importância da elaboração de modelos preditivos para doenças infecciosas, possibilitando alerta precoce e melhorando significativamente a capacidade de vigilância, controle e prevenção de surtos, particularmente nas regiões em desenvolvimento⁽¹⁷⁾.

As informações epidemiológicas ajudam a definir, conter, prevenir e tratar com sucesso o impacto da doença infecciosa. O uso da epidemiologia e seu arcabouço torna-se capaz de sustentar hipóteses que forneçam a compreensão para conter ou eliminar a doença⁽¹⁸⁾. O desafio da epidemiologia é grande, uma vez que as respostas dos governos ou medidas de saúde pública para o gerenciamento da pandemia podem ser ineficazes à medida que os problemas ocorrem em vários setores da sociedade e se manifestam de maneiras desconhecidas⁽¹⁶⁾.

Outra questão importante evidenciada nos tópicos deste estudo é, que até o momento, não há tratamento antiviral específico recomendado para a pandemia de COVID-19 e as estratégias

terapêuticas para lidar com a infecção são apenas de suporte⁽¹⁹⁾. Estão em andamento muitas tentativas de desenvolvimento de vacinas contra o SARS-CoV-2 (Tópico 16). Assim, em razão das fortes semelhanças genéticas da epidemia SARS-CoV de 2003, há investigações de reaproveitamento de drogas aprovadas, como antivirais, antimaláricos e imunomoduladores⁽²⁰⁾.

Desse modo, a compreensão da resposta imunológica é fundamental para o desenvolvimento de potenciais terapêuticas. Esses temas foram amplamente abordados em 2 dos 20 tópicos analisados no presente estudo, enfatizando sua importância e relevância para controle, tratamento e prevenção da COVID-19.

Além dos esforços para entender a epidemiologia, características clínicas, transmissão e manejo da COVID-19, o tema da saúde mental foi o agravo à saúde mais presente nas publicações. No período de pandemia, as pessoas precisam lidar com questões como o isolamento social, distanciamento, morte, xenofobia e outras situações que têm gerado consequências psicológicas, tais como estresse, ansiedade, medo e depressão, tanto na população geral, em diferentes faixas etárias e classes sociais^(21,22), quanto em grupos vulneráveis e/ou com agravos em saúde mental preexistentes^(23,24). Ainda, há uma preocupação considerável nos estudos identificados sobre os transtornos psicológicos nos profissionais de saúde, em especial àqueles que atuam na linha de frente de combate à COVID-19⁽²⁵⁾.

Como forma de superar alguns desafios de acesso à assistência à saúde, garantir continuidade de cuidados e diminuir o contato entre pacientes e profissionais de saúde, o tema da telemedicina e telessaúde foi outro destaque das publicações analisadas. Essas estratégias vêm tendo seu uso ampliado nas diferentes especialidades em saúde com discussões sobre a qualidade e eficácia dessa modalidade de assistência⁽²⁶⁻²⁸⁾.

No âmbito da COVID-19, a telemedicina surgiu como uma ferramenta para triagem da população, de modo que o monitoramento remoto tem recursos operacionais específicos que são adequados para esse agravo, especialmente a comunicação assíncrona, podendo ser usado para coletar dados de

pandemia e obter *feedback* clínico em tempo real⁽²⁹⁾.

Além dos temas discutidos, a análise pelo LDA permitiu identificar outros tópicos abordados nas pesquisas com a COVID-19, entre eles o desenvolvimento de diretrizes clínicas para atenção em situações eletivas ou de urgência/emergência, equipamentos de proteção individual, estratégias governamentais e conscientização da população para controle da transmissão, aspectos clínicos, e atenção à saúde das gestantes e crianças. Esse amplo comportamento das pesquisas torna-se dinâmico e essencial para o entendimento de muitos aspectos que envolvem a COVID-19.

Apesar da ampla variedade de temas, percebeu-se ainda uma lacuna em relação ao tema da atenção primária à saúde (APS). Esse tópico vem sendo pesquisado, porém não de forma tão expressiva a ponto de aparecer como um tópico dominante de pesquisa na análise pelo LDA, o que leva a questionar se as evidências produzidas até o momento são suficientes para respaldar ações nesse campo.

Mesmo que os sistemas de saúde de muitos países não sejam estruturados por meio da APS, seu fortalecimento é a resposta política mais eficaz, especialmente em países de baixa e média renda, para proteger contra emergências de saúde, alcançar cobertura universal de saúde e promover saúde e bem-estar⁽³⁰⁾. Esse papel da APS tem sido evidenciado durante a pandemia SARSCoV-2019, em que fortes sistemas de APS contribuíram para países responderem com mais rapidez, eficácia e obter melhores resultados de saúde para suas populações⁽³¹⁾.

Mais desafiador ainda podem ser pesquisas que envolvam questões consideradas não prioritárias para a escalada de contágio do vírus que se vive. Para além de conhecer estratégias terapêuticas, podem ser importantes pesquisas socioambientais, que demonstram a adaptação e por que não a forma de sobrevivência da população na escalada da pandemia, no sentido de entender a dinâmica social de grupos, como em favelas, na zona rural, em pequenas cidades, diante de todas as etapas do curso pandêmico, mesmo as ainda não detectadas ou descobertas.

Embora se tenha tentado desenvolver uma estratégia de busca abrangente, é possível que

alguns estudos relevantes tenham sido perdidos. Igualmente, materiais e pesquisas que não foram publicados em periódicos indexados nas bases *Web of Science* e *Embase* não foram pesquisados e poderiam ter fornecido algumas informações adicionais, configurando uma limitação do estudo.

CONCLUSÃO

O presente estudo analisou o desenvolvimento global de publicações científicas e construiu tópicos abordados em pesquisas sobre o novo coronavírus/COVID-19.

Verificaram-se muitas publicações em diferentes periódicos, realizadas por pesquisadores de 182 países. Além disso, identificaram-se os principais temas que têm sido abordados sobre o tema, caracterizando sua importância e relevância no entendimento do agravo em questão para traçar estratégias de controle, prevenção, tratamento e atenção aos agravos/implicações decorrentes da pandemia.

Recomenda-se a ampliação de estudos sobre o novo coronavírus/COVID-19, especialmente em áreas menos contempladas, a exemplo da APS.

MODELING OF RESEARCH TOPICS ON THE NEW CORONAVIRUS: APPLICATION OF LATENT DIRICHLET ALLOCATION

ABSTRACT

Objective: to analyze the global development of scientific publications and build topics covered in research on the new coronavirus/COVID-19. **Methods:** review of abstracts of publications on the new coronavirus/COVID-19 with search in the Embase and Web of Science databases, on November 3, 2020. Data were analyzed by modeling the research topics using Latent Dirichlet Allocation (LDA) method. **Results:** a total of 31,769 publications were analyzed, involving researchers from 182 countries. The studies were published in 5,336 different journals. The analysis identified 20 main research topics on the new coronavirus/COVID-19, with more than half (54.5%) covering topics of impacts and challenges resulting from the pandemic (11.3%), immune response (9.6%), potential therapies (9.0%), epidemiology (7.8%), mental health (6.3%), recommendations for assistance with other health needs (5.3%), and clinical complications. The theme of telemedicine and telehealth was present in 4.2% of publications. **Conclusions:** scientific publications were distributed among 20 main themes, characterizing their importance and relevance in understanding the disease in question to outline strategies for control, prevention, treatment and attention to diseases/implications resulting from the pandemic. Studies on primary care measures/primary health care were identified as the main knowledge gap.

Keywords: Coronavirus infections. Pandemics. COVID-19. Review.

MODELADO DE TÓPICOS DE INVESTIGACIÓN SOBRE EL NUEVO CORONAVIRUS: APLICACIÓN DEL LATENT DIRICHLET ALLOCATION

RESUMEN

Objetivo: analizar el desarrollo global de publicaciones científicas y construir tópicos tratados en investigaciones sobre el nuevo coronavirus/COVID-19. **Métodos:** estudio de revisión de resúmenes de publicaciones sobre el nuevo coronavirus/COVID-19, con busca en las bases de datos *Embase* y *Web of Science*, el 03 de noviembre de 2020. El análisis de los datos se llevó a cabo por el modelado de tópicos de investigación por el método *Latent Dirichlet Allocation* (LDA). **Resultados:** se analizaron 31.769 publicaciones, con involucramiento de investigadores de 182 países. Los estudios fueron publicados en 5.336 periódicos diferentes. El análisis identificó 20 tópicos principales de investigación sobre el nuevo coronavirus/COVID-19, más de la mitad (54,5%) de las publicaciones fue contemplada en los temas de impactos y desafíos resultantes de la pandemia (11,3%), respuesta inmunológica (9,6%), potenciales terapéuticos (9,0%), epidemiología (7,8%), salud mental (6,3%), recomendaciones para asistencia a otras necesidades de salud (5,3%) y complicaciones clínicas. El tema de la telemedicina y telesalud fue evidenciado en 4,2% de las publicaciones. **Conclusiones:** las publicaciones científicas estaban distribuidas entre 20 principales temas, caracterizando su importancia y relevancia en el entendimiento del agravo en cuestión para esbozar estrategias de control, prevención, tratamiento y atención a los agravios/implicaciones resultantes de la pandemia. Se identificaron como principal laguna de conocimiento estudios sobre cuidados primarios/atención primaria a la salud.

Palabras clave: Infecciones por coronavirus. Pandemias. COVID-19. Revisión.

REFERÊNCIAS

1. Jefferson T, Del Mar CB, Dooley L, Ferroni E, Al-Ansary LA, Bawazeer GA, et al. Physical interventions to interrupt or reduce the spread of respiratory viruses. *Cochrane database Syst Rev*. 2011;(7): CD006207. Doi:

<https://doi.org/10.1002/14651858.CD006207.pub4>

2. Lai CC, Shih TP, Ko WC, Tang HJ, Hsueh PR. Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and coronavirus disease-2019 (COVID-19): The epidemic and the challenges. *Int J Antimicrob Agents*. 2020;55(3):105924. Doi: <https://doi.org/10.1016/j.ijantimicag.2020.105924>

3. Gorbalenya AE, Baker SC, Baric RS, Groot RJ, Drosten C, Gulyaeva AA, et al. The species Severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2. *Nat Microbiol.* 2020; 5:536–44. Doi: <https://doi.org/10.1038/s41564-020-0695-z>.
4. Xie M, Chen Q. Insight into 2019 novel coronavirus: An updated interim review and lessons from SARS-CoV and MERS-CoV. *Int J Infect Dis.* 2020; 94:119–124. Doi: <https://doi.org/10.1016/j.ijid.2020.03.071>.
5. World Health Organization. Coronavirus Disease (COVID-19) Situation Reports [on-line]. 2020 [cited 1 Jul. 2020]. Available at: <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports>
6. Amin Tabish S. COVID-19 pandemic: Emerging perspectives and future trends. *Journal of Public Health Research.* 2020. *J Public Health Res.* 2020; 9(1): 1786. Doi: <https://doi.org/10.4081/jphr.2020.1786>
7. Sarkar D. Text Analytics with Python: A Practitioner's Guide to Natural Language Processing. Apress; 2019.
8. Blei DM, Ng AY, Edu JB. Latent Dirichlet Allocation Michael I. Jordan. *Journal of Machine Learning Research.* 2003.
9. Hebbar PB, Sudha A, Dsouza V, Chilgod L, Amin A. Healthcare delivery in India amid the Covid-19 pandemic: Challenges and opportunities. *Indian J Med Ethics.* 2020;05(03):215–218. Doi: <https://doi.org/10.20529/IJME.2020.064>
10. Negrini S, Grabljevec K, Boldrini P, Kiekens C, Moslavac S, Zampolini M, et al. Up to 2.2 million people experiencing disability suffer collateral damage each day of COVID-19 lockdown in Europe. *Eur J Phys Rehabil Med.* 2020; 56(3):361–365. Doi: <https://doi.org/10.23736/S1973-9087.20.06361-3>
11. Möhn N, Pul R, Kleinschnitz C, Prüss H, Witte T, Stangel M, et al. Implications of COVID-19 Outbreak on Immune Therapies in Multiple Sclerosis Patients: Lessons Learned From SARS and MERS. *Front Immunol.* 2020; 11:1059. Doi: <https://doi.org/10.3389/fimmu.2020.01059>
12. Ghosh R, Dubey MJ, Chatterjee S, Dubey S. Impact of COVID-19 on children: Special focus on the psychosocial aspect. *Minerva Pediatrica* 72(3):226-235. Doi: <https://doi.org/10.23736/S0026-4946.20.05887-9>
13. Pichard R, Kopel L, Lejeune Q, Masmoudi R, Masmajeun EH. Impact of the Coronavirus Disease 2019 lockdown on hand and upper limb emergencies: experience of a referred university trauma hand centre in Paris, France. *Int Orthop.* 2020; 44(8):1497–1501. Doi: <https://doi.org/10.1007/s00264-020-04654-2>
14. Bonaccorsi G, Pierri F, Cinelli M, Flori A, Galeazzi A, Porcelli F, et al. Economic and social consequences of human mobility restrictions under COVID-19. *Proc Natl Acad Sci USA.* 2020; 117(27):15530–15535. Doi: <https://doi.org/10.1073/pnas.2007658117>
15. Lynch J. Health Equity, Social Policy, and Promoting Recovery from COVID-19. *J Health Polit Policy Law.* 2020;8641518. Doi: <https://doi.org/10.1215/03616878-8641518>
16. Bryce C, Ring P, Ashby S, Wardman JK. Resilience in the face of uncertainty: early lessons from the COVID-19 pandemic. *J Risk Res.* 2020;1-9. Doi: <https://doi.org/10.1080/13669877.2020.1756379>
17. Bai Z, Gong Y, Tian X, Cao Y, Liu W, Li J. The Rapid Assessment and Early Warning Models for COVID-19. *Viro Sin.* 2020; 35(3):272–279. Doi: <https://doi.org/10.1007/s12250-020-00219-0>
18. Binns C, Low WY, Kyung LM. The COVID-19 Pandemic: Public Health and Epidemiology. *Asia Pac J Public Health.* 2020;32(4):140–144. Doi: <https://doi.org/10.1177/1010539520929223>
19. Bari E, Ferrarotti I, Saracino L, Perteghella S, Torre ML, Corsico AG. Mesenchymal Stromal Cell Secretome for Severe COVID-19 Infections: Premises for the Therapeutic Use. *Cells.* 2020;9(4):924. Doi: <https://doi.org/10.3390/cells9040924>
20. Amawi H, Abu Deiaab GI, Aljabali AA, Dua K, Tambuwala MM. COVID-19 pandemic: an overview of epidemiology, pathogenesis, diagnostics and potential vaccines and therapeutics. *Ther Deliv.* 2020; 11(4):245–68. Doi: <https://doi.org/10.4155/tde-2020-0035>
21. Naqvi HA. Mental health in the aftermath of COVID-19: A new normal. *J Pak Med Assoc.* 2020;70(5):S141–4. Doi: <https://doi.org/10.5455/JPMA.30>
22. Javed B, Sarwer A, Soto EB, Mashwani Z. The coronavirus (COVID-19) pandemic's impact on mental health. *Int J Health Plann Manage.* 2020; 35(5):993–6. Doi: <https://doi.org/10.1002/hpm.3008>
23. French I, Lyne J. Acute exacerbation of OCD symptoms precipitated by media reports of COVID-19. *Ir J Psychol Med.* 2020; 1–4. Doi: <https://doi.org/10.1017/ipm.2020.61>
24. Stefana A, Youngstrom EA, Chen J, Hinshaw S, Maxwell V, Michalak E, et al. The COVID-19 pandemic is a crisis and opportunity for bipolar disorder. *Bipolar Disord.* 2020; 22(6):641–643. Doi: <https://doi.org/10.1111/bdi.12949>
25. Zhu J, Sun L, Zhang L, Wang H, Fan A, Yang B, et al. Prevalence and Influencing Factors of Anxiety and Depression Symptoms in the First-Line Medical Staff Fighting Against COVID-19 in Gansu. *Front Psychiatry.* 2020;11:386. Doi: <https://doi.org/10.3389/fpsy.2020.00386>
26. Said M, Ngo V, Hwang J, Hom DB. Navigating telemedicine for facial trauma during the COVID-19 pandemic. *Laryngoscope Investig Otolaryngol.* 2020;5(4):649–656. Doi: <https://doi.org/10.1002/lto.2.428>
27. Tenforde AS, Iaccarino MA, Borgstrom H, Hefner JE, Silver J, Ahmed M, et al. Telemedicine During COVID-19 for Outpatient Sports and Musculoskeletal Medicine Physicians. *PM&R.* 2020;12(9):926–932. Doi: <https://doi.org/10.1002/pmrj.12422>
28. Boehm K, Ziewers S, Brandt MP, Sparwasser P, Haack M, Willems F, et al. Telemedicine Online Visits in Urology During the COVID-19 Pandemic: Potential, Risk Factors, and Patients' Perspective. *Eur Urol.* 2020; 78(1):16–20. Doi: <https://doi.org/10.1016/j.eururo.2020.04.055>
29. Watson AR, Wah R, Thamman R. The Value of Remote Monitoring for the COVID-19 Pandemic. *Telemed J E Health.* 2020;26(9):1110–1112. Doi: <https://doi.org/10.1089/tmj.2020.0134>
30. Kraef C, Juma P, Kallestrup P, Mucumbitsi J, Ramaiya K, Yonga G. The COVID-19 Pandemic and Non-communicable Diseases: A Wake-up Call for Primary Health Care System Strengthening in Sub-Saharan Africa. *J Prim Care Community Health.* 2020; 11:1–3. Doi: <https://doi.org/10.1177/2150132720946948>
31. Dunlop C, Howe A, Li D, Allen LN. The coronavirus outbreak: The central role of primary care in emergency preparedness and response. *BJGP Open.* 2020;4(1):1–3. Doi: <https://doi.org/10.3399/bjgpopen20X101041>

Endereço para correspondência: Tatiane Baratieri. Rua Alameda Élio Antonio Dalla Vecchia, 838. Departamento de Enfermagem. CEP 85040-167. Bairro Vila Carli, Guarapuava/PR. E-mail: baratieri.tatiane@gmail.com

Data de recebimento: 26/10/2020

Data de aprovação: 19/12/2020