

Análises de classes latentes dos sintomas relacionados à COVID-19 no Brasil: resultados da PNAD-COVID19

Latent class analysis of COVID-19 symptoms in Brazil: results of the PNAD-COVID19 survey

Análisis de clases latentes de síntomas relacionados con la COVID-19 en Brasil: resultados de la PNAD-COVID19

Rafael da Silveira Moreira ^{1,2}

doi: 10.1590/0102-311X00238420

Resumo

A ausência de testagens em massa para o diagnóstico da COVID-19 gera a necessidade de conhecer a dimensão da doença por meio da sua sintomatologia clínica. O objetivo foi investigar o perfil de sintomas relacionados à COVID-19 e aspectos relacionados. Foi analisada a amostra de participantes da Pesquisa Nacional por Amostra de Domicílios (PNAD-COVID19) realizada em maio de 2020. Foi realizada análise de classes latentes (ACL) com covariáveis sociodemográficas sobre 11 sintomas relatados por 346.181 indivíduos. Foram utilizados testes de Rao-Scott e análise de resíduos padronizados para mensurar a associação com o padrão de utilização dos serviços de saúde. Análise espacial de varredura foi realizada para identificar as áreas de risco para os casos de COVID-19. A ACL mostrou seis classes de sintomatologia, segundo o padrão de respostas dos indivíduos analisados: (1) todos os sintomas; (2) prevalência alta dos sintomas; (3) predominância de febre; (4) predominância de tosse/dor de garganta; (5) leves sintomas com predominância de dor de cabeça e (6) ausência de sintomas. Pessoas do sexo feminino, cor parda, provenientes das regiões Norte e Nordeste e em todas as três faixas etárias mais velhas apresentaram maior associação com a classe com todos os sintomas (classe 1). A maioria da procura por serviços também foi realizada por esse grupo de indivíduos, porém com distintos perfis de uso. A análise espacial mostrou sobreposição dessa classe com áreas de maior risco de casos de COVID-19. Os achados sustentam a importância da investigação dos sintomas, servindo para a identificação epidemiológica de possíveis casos em um cenário com baixa taxa de testagem populacional.

COVID-19; Análise de Classes Latentes; Pandemias; Monitoramento Epidemiológico

Correspondência

R. S. Moreira
Instituto Aggeu Magalhães, Fundação Oswaldo Cruz.
Av. Professor Moraes Rego s/n, Cidade Universitária, Recife, PE
50670-420, Brasil.
rafael.moreira@fiocruz.br

¹ Instituto Aggeu Magalhães, Fundação Oswaldo Cruz, Recife, Brasil.

² Faculdade de Medicina, Universidade Federal de Pernambuco, Recife, Brasil.



Introdução

A pandemia causada pelo novo coronavírus, nomeado COVID-19 (*corona virus disease*, identificada em 2019), inicialmente relatado em dezembro de 2019, em Wuhan, China ^{1,2}, surgiu associada a formas graves de pneumonia, com alto poder de transmissão. De modo geral, os pacientes apresentam falta de ar, tosse seca, febre, dor de cabeça e dispneia. Os casos de óbito normalmente ocorrem devido a falência respiratória progressiva, com graves danos pulmonares ^{3,4,5}.

No Brasil, o primeiro caso foi identificado em 26 de fevereiro de 2020. Em 17 de março de 2020, houve o registro do primeiro óbito. Em 5 de agosto de 2020, pouco mais de cinco meses depois do primeiro registro da doença, o total de casos e óbitos registrados no país foi de 2.859.073 e 97.256, respectivamente (Ministério da Saúde. Painel coronavírus. <https://covid.saude.gov.br/>, acessado em 05/Ago/2020). Com modelos exponenciais de progressão, medidas de contenção incluem testagem do maior número de casos, cuidados com higiene e isolamento social ⁶. Embora existam casos em que a doença não expressa sintomas em seus portadores ^{7,8,9}, a demanda por serviços de saúde geralmente é acompanhada pelas queixas sintomatológicas associadas à doença. O período de incubação relatado foi, em média, de 5,2 dias ¹⁰.

Entretanto, a circulação de casos assintomáticos pode representar aumento na taxa de infecção. Considerando a amplitude de tempo desde o momento da manifestação da doença até o óbito, que leva de 6 a 41 dias, com uma mediana de 14 dias ¹¹, a identificação precoce de casos suspeitos representa uma imprescindível janela de oportunidade para o melhor manejo, farmacológico ou não, da epidemia.

A experiência de países que conseguiram êxito no controle da epidemia mostrou alta taxa de testagens da população, acompanhamento dos casos e de contactantes, bem como maior engajamento político e social na manutenção das medidas de distanciamento social. A testagem em massa também permite o conhecimento de infecções ainda em nível subclínico ^{10,12,13}. Países de baixa e média renda, que têm dificuldades de implantação de testes diagnósticos massivos, devem priorizar a triagem baseada na definição de caso clínico ou no diagnóstico presuntivo ¹⁴.

O Brasil apresenta uma taxa de testagem bastante inferior ao que seria esperado para o adequado controle da epidemia, flexibilizações de medidas de distanciamento social por pressões econômicas e instabilidades na coordenação setorial no campo da saúde pública. Como resultado, observamos o avanço da epidemia em diferentes ritmos nas diversas escalas geográficas nacionais. Essa desigualdade é espelho da iniquidade social brasileira, na qual a pandemia encontra já de partida um país debilitado pelo modesto crescimento econômico e por um setor público (ciência, educação e saúde) enfraquecido pela redução dos investimentos em políticas públicas, resultado do modelo neoliberal de austeridade fiscal, realidade compartilhada pelos demais países da América Latina ¹⁵.

Na ausência de testes em massa, o conhecimento dos casos notificados geralmente ocorre no momento da procura pelos serviços de saúde, quando há meios de confirmação dos casos suspeitos. Outra forma de notificação ocorre na confirmação da causa básica do óbito. Em ambas as situações, a notificação se limita aos casos que buscaram atendimento e que tiveram acesso, caracterizando um cenário propício para o colapso dos sistemas de saúde, por não exibir estratégia de vigilância epidemiológica ativa dos casos suspeitos.

Nesse sentido, o Instituto Brasileiro de Geografia e Estatística (IBGE) vem realizando sistematicamente a *Pesquisa Nacional por Amostra de Domicílios* específica para a COVID-19, chamada de PNAD-COVID19 ¹⁶. Os entrevistados respondem a perguntas relacionadas aos sintomas autorreferidos de COVID-19, ao perfil de busca por serviços de saúde e à situação ocupacional no mercado de trabalho. Assim, considerando o cenário brasileiro de baixa taxa de testagem em massa, torna-se fundamental conhecer o padrão sintomatológico relacionado à COVID-19. A forma como se apresenta esse padrão na população, reforçada pelo conhecimento científico sobre os fatores associados, pode gerar melhor suporte para o manejo focalizado de pacientes com maior probabilidade de confirmação positiva da doença. Do mesmo modo, propicia o uso racional dos leitos de UTI e de ventiladores pulmonares epidemiologicamente direcionados ¹⁷, além de permitir melhor detecção, diagnóstico e monitoramento oportuno dos casos suspeitos. Dentro desse escopo, o objetivo do estudo foi conhecer o padrão sintomatológico relacionado à COVID-19 e os principais fatores socioespaciais associados.

Metodologia

Foi realizada uma análise transversal dos dados provenientes da PNAD-COVID19 ocorrida em maio de 2020. Nessa amostra, foram coletadas questões sobre sintomas relacionados à COVID-19, o perfil de uso de serviços de saúde relacionado a esses sintomas e as repercussões no mercado de trabalho. Detalhes sobre o plano amostral podem ser encontrados em IBGE ¹⁶. A PNAD-COVID19 teve como base a amostra de domicílios da PNAD Contínua do primeiro trimestre de 2019. Foi realizado um processo de pareamento para integração com outras bases de dados, a fim de se obterem números de telefone de cada domicílio. Dessa forma, foi possível identificar uma amostra com, ao menos, um telefone disponível em 193.662 domicílios. Esses dados representam cerca de 92% da amostra-base, distribuídos em conjuntos de cerca de 48 mil domicílios por semana. Essa amostra é fixa, no sentido de que os domicílios entrevistados no primeiro mês de coleta de dados permanecerão na amostra dos meses subsequentes até o fim da pesquisa. Todos os moradores residentes no domicílio selecionados foram convidados a responder à pesquisa, realizada por CATI (*computer assisted telephone interviewing*), ou seja, entrevistas por telefone assistidas por computador.

Para este estudo, foram considerados os dados dos sintomas, do perfil sociodemográfico e do uso de serviços de saúde. As perguntas dos sintomas foram referentes à presença, na semana de referência (semana anterior à entrevista), de 11 sintomas: febre, tosse, dor de garganta, dificuldade para respirar, dor de cabeça, dor no peito, náusea, nariz entupido ou escorrendo, fadiga, dor nos olhos e perda de cheiro ou sabor.

Com relação ao perfil do uso de serviços de saúde, foi questionado se houve procura por esses serviços e qual a providência que foi tomada (ficar em casa, ligar para profissional de saúde, automedicar-se, tomar medicação prescrita, receber visita de profissional de saúde do Sistema Único de Saúde – SUS ou privado). Sobre o local da procura, foram utilizadas as seguintes opções: unidade básica de saúde (UBS) ou equipe de saúde da família, unidade de pronto-atendimento (UPA), hospital do SUS, consultório particular, pronto-socorro particular, hospital particular. Finalmente, questionou-se sobre a ocorrência de internação.

Os dados sociodemográficos foram: sexo, cor da pele, escolaridade, idade e região geográfica. Para contextualizar espacialmente os casos de COVID-19 com os sintomas, foi utilizado o número de casos acumulados confirmados até junho de 2020 por 100 mil habitantes, de acordo com Regiões de Saúde e Unidades da Federação. Foram utilizados os casos acumulados até o dia 31 de maio de 2020, disponibilizados pelo Ministério da Saúde (Painel coronavírus. <https://covid.saude.gov.br/>, acessado em 05/Ago/2020).

Para conhecer o perfil dos sintomas relacionados à COVID-19, considerada a variável dependente neste estudo, foi realizada análise de classes latentes (ACL) com covariáveis. Trata-se de um procedimento estatístico que busca agrupar indivíduos segundo padrões semelhantes de respostas, modeladas com covariáveis. Nesse sentido, formam-se classes com maior homogeneidade intraclasse e maior heterogeneidade interclasse. As classes latentes foram geradas a partir dos onze sintomas autorreferidos pelos participantes da PNAD-COVID19. Contudo, no sentido de garantir a independência local entre as variáveis indicadoras, foi necessário realizar três agrupamentos: (1) tosse com dor de garganta; (2) dificuldade de respirar com dor no peito; (3) náusea com fadiga. Dessa forma, passamos a ter oito variáveis indicadoras no lugar das 11 iniciais.

Na construção dessa variável latente, modelos com diferentes números de classes latentes (categorias) foram criados e testados até se encontrar o modelo ideal para descrever essa variável. Para essa escolha, foram observados os seguintes critérios: critério de informação de Akaike (AIC), critério de informação Bayesiano (BIC) e critério de informação Bayesiano ajustado (BIC ajustado), observando-se sempre os menores valores quando comparados o modelo atual com o anterior. Também foi considerado o maior valor da entropia ¹⁸. Além desses critérios, foram realizados outros dois testes estatísticos (Vuong-Lo-Mendell-Rubin; Lo-Mendell-Rubin) para verificar se o número de classes escolhido é melhor em termos de ajuste do modelo, quando comparado com o número de classes do modelo anterior.

A fim de testar a associação das classes latentes com as variáveis independentes socioeconômicas, foram estimados modelos simples e múltiplo de regressão logística multinomial, com a *odds ratio* (OR) como medida de efeito, e seus respectivos intervalos de 95% de confiança (IC95%). A categoria de

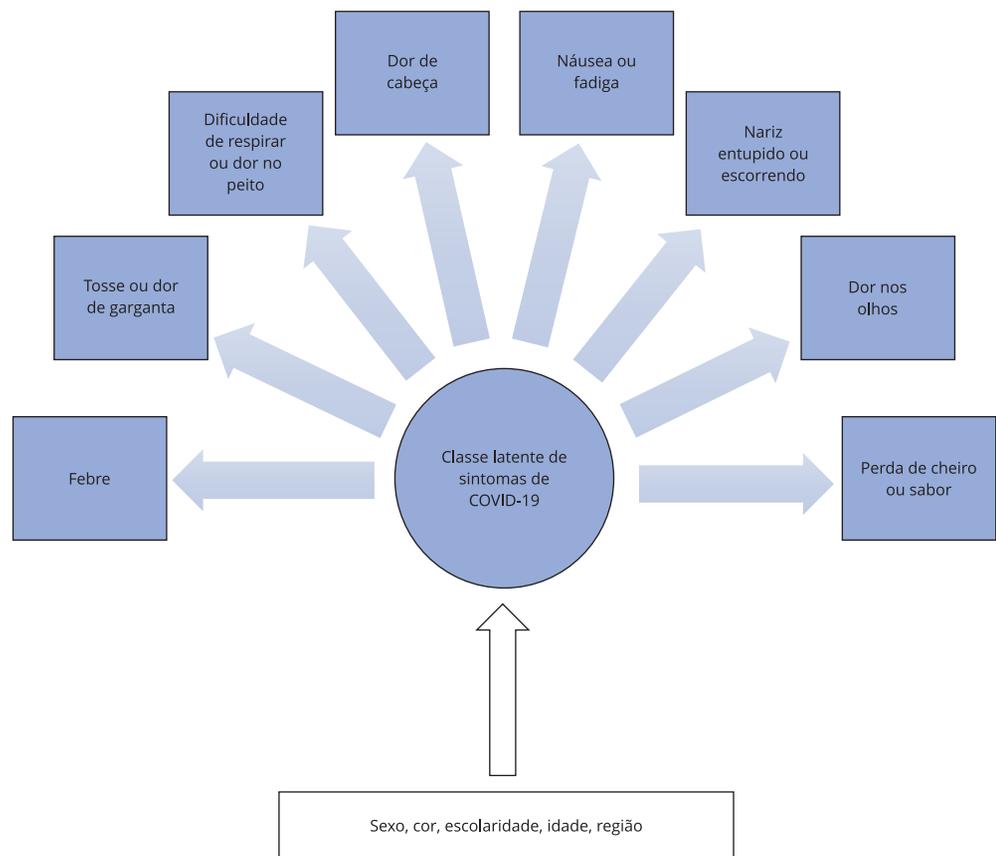
referência foi a classe latente “sem sintomas”. Para a estimação dos modelos de ACL com covariáveis sociodemográficas, foi utilizado o programa Mplus 6.12 (<https://www.statmodel.com/>), considerando as características do desenho amostral complexo. O nível de significância foi de 5%. A Figura 1 ilustra o modelo que representa a ACL com covariáveis.

Para analisar a associação com o perfil do uso de serviços de saúde, foram realizados testes de Rao-Scott¹⁹, como adaptação do teste de qui-quadrado para amostras complexas. Análise de resíduos padronizados (ARP) foi realizada para identificar excessos de observações em cada categoria da variável dependente com as categorias das variáveis independentes. Foram considerados significantes os resíduos padronizados acima de 1,96 em um teste unicaudal com 2,5% de significância. A ARP consiste no exame dos resíduos (diferenças entre observado e esperado em tabelas de contingência) em uma forma padronizada, ou seja, expressos em unidade de desvios-padrão. Nesse sentido, na distribuição de probabilidades de ocorrência, resíduos padronizados maiores que 1,96 ou menores que -1,96 têm pequenas chances de ocorrência ($\pm 2,5\%$)^{20,21}.

Em amostras complexas, os resíduos padronizados podem ser muito grandes, haja vista a tendência de os erros-padrões produzidos sem a correção de Rao-Scott serem menores do que os erros-padrões com a devida correção. Dessa forma, o programa SPSS para Windows versão 20 (<https://www.ibm.com/>) oferece a opção de análise de tabelas cruzadas em amostras complexas. Nessa opção, tanto

Figura 1

Modelo de análise de classes latentes (ACL) com covariáveis.



a correção de Rao-Scott do teste de qui-quadrado quanto o cálculo dos resíduos padronizados são ajustados com base no plano amostral complexo.

A não incorporação das variáveis do perfil do uso dos serviços de saúde na estimação das classes latentes (ACL com covariáveis) pode provocar vieses por erros de mensuração. Entretanto, segundo Wang & Wang ²², o uso da ACL (sem covariáveis) mesclada com outros bancos de dados, com auxílio de outros programas, tal como o SPSS, pode ser executado quando o modelo mais provável de pertencimento de classe é adequado, ou seja, quando a entropia é alta (maior que 0,80). Em nosso caso, a entropia com seis classes foi 0,928. Nesse sentido, mesmo assumindo uma possível limitação pelo viés de mensuração, a alta entropia permite uma análise das classes por meio de alternativas, além do modelo com covariáveis. Outro aspecto importante, devido ao elevado número de indivíduos que não procuraram os serviços de saúde, especialmente pela ausência de sintomas, foi a preferência por realizar apenas uma análise bivariada com tabelas de contingência, pois a inclusão dessas variáveis em um modelo múltiplo de regressão logística multinomial violaria tanto a parcimônia do modelo quanto a ausência de colinearidade entre essas variáveis independentes.

Ressaltamos que em todas as análises foram incluídos os parâmetros da amostra complexa. Por ser uma amostra muito grande, deve-se considerar o desenho amostral na análise da significância estatística dos testes e das estimativas produzidas.

Para a análise espacial, foi realizado o teste de varredura com os casos confirmados de COVID-19. O círculo máximo permitido pelo programa é de 50% da população em risco. Contudo, essa alta proporção incluiria metade da população brasileira, dificultando a identificação de áreas menores de risco. Assim, o valor da janela espacial baseou-se no valor indicado nos resultados do índice de Gini do SaTScan (<https://www.satscan.org/>), disponibilizado após uma primeira varredura espacial padrão (50%), gerando um valor ideal de 2% da população em risco. Quando a análise encontra um círculo com risco relativo maior do que o risco de fora do círculo, são processadas réplicas do tipo Monte Carlo para o cálculo do valor de p, considerando 5% de significância ²³. Dessa forma, foi construído um mapa temático com apresentação do número de casos por 100 mil habitantes segundo quintil, distribuição proporcional das classes latentes e áreas com risco relativo com significância estatística.

Por envolver dados publicamente disponibilizados na Internet e sem nenhuma forma de identificação individual, este estudo não precisou ser submetido ao Comitê de Ética em Pesquisas com Seres Humanos.

Resultados

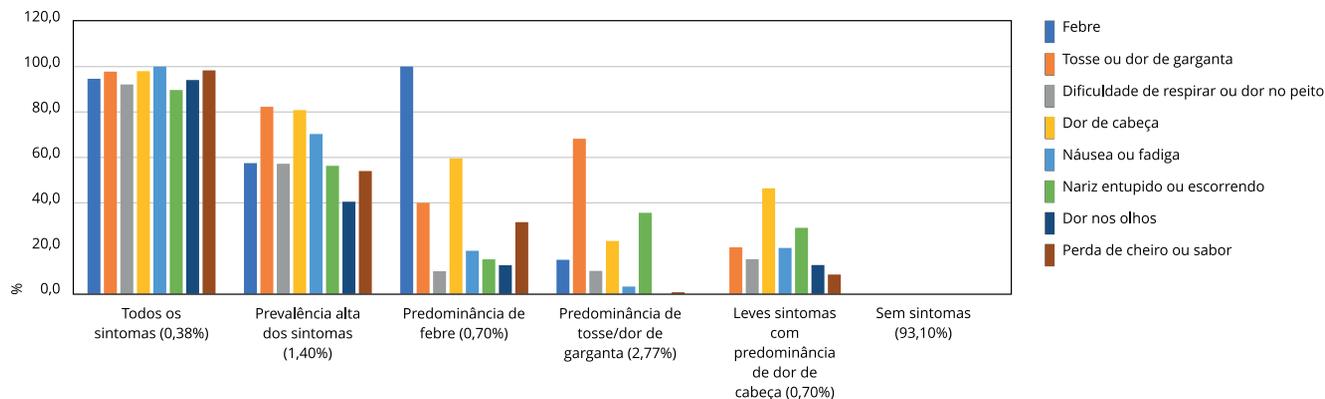
A Figura 2 ilustra a distribuição não condicional das classes latentes e a distribuição condicional das respostas positivas para cada um dos oito sintomas, segundo a classificação da ACL. O modelo que apresentou os melhores parâmetros de ajuste foi com seis classes. Elas foram classificadas como: (1) todos os sintomas; (2) prevalência alta dos sintomas; (3) predominância de febre; (4) predominância de tosse/dor de garganta; (5) leves sintomas com predominância de dor de cabeça; (6) ausência de sintomas.

A Tabela 1 apresenta os resultados do modelo múltiplo de regressão logística multinomial da ACL com covariáveis sociodemográficas. Pessoas do sexo feminino apresentaram mais chance de serem classificadas nas classes com maior número de sintomas. A chance foi 31% maior no sexo feminino de apresentar todos os sintomas (classe 1). Ser de cor parda apresentou uma chance 34% maior de apresentar todos os sintomas, ao passo que ser de cor preta apresentou chance 55% maior de predominância de febre, em comparação com os indivíduos de cor branca. Amarelos e indígenas apresentaram chances 80% maiores de prevalência alta dos sintomas (classe 2).

Observou-se que as categorias de maior escolaridade apresentaram maiores chances de possuírem todos os sintomas (classe 1), prevalência alta dos sintomas (classe 2) e predominância de tosse ou dor de garganta (classe 4). Isso também ocorreu com os quartis de idade mais avançada em todas as classes, exceto na de sintomas leves (classe 5). Comparadas com a Região Sudeste, as regiões Norte e Nordeste apresentaram mais chance de maior número de sintomas (classes 1 e 2). A Região Sul apresentou chance 14% maior de estar na classe com predominância de tosse ou dor de garganta (classe 4).

Figura 2

Distribuição de probabilidades não condicionais das classes latentes e condicionais dos oito sintomas relacionados à COVID-19 segundo as seis classes latentes. Brasil, 2020.



A Tabela 2 apresenta o resultado da distribuição da frequência relativa das classes latentes segundo o perfil do uso dos serviços de saúde. Observou-se que a procura por esses serviços esteve associada à prevalência das classes latentes de todos os sintomas (classe 1), prevalência alta dos sintomas (classe 2) e predominância de febre (classe 3). Com exceção das classes 5 (leves sintomas com predominância de dor de cabeça) e 6 (ausência de sintomas), todos os indivíduos classificados nas demais classes tomaram a providência de ficar em casa. Contudo, os atos de ligar para profissional de saúde, fazer uso de medicação (prescrita ou na forma de automedicação) e receber visita de um profissional de saúde particular estiveram associados às classes de maior gravidade dos sintomas. Receber visita por profissional do SUS esteve associado com possuir tanto todos os sintomas quanto leves sintomas com predominância de dor de cabeça. O local de busca de atendimento entre os indivíduos com todos os sintomas foi a UPA do SUS. Para aqueles que apresentaram maior prevalência de tosse e dor de garganta, a UBS foi o local mais procurado. Os hospitais tanto públicos como privados não foram os locais prioritários para busca por atendimento, assim como os consultórios particulares. Os prontos-socorros particulares foram buscados na prevalência alta dos sintomas. Houve um excesso de indivíduos da classe com todos os sintomas (classe 1) que ficaram internados ou que tentaram internação, mas não conseguiram.

Com relação à distribuição espacial dos quintis do número de casos confirmados por 100 mil habitantes e das áreas circulares com maior risco relativo da doença, observamos uma sobreposição de grande parte das regiões/áreas com maiores proporções de indivíduos que foram classificados com presença de todos os sintomas (classe 1). A classe de indivíduos sem sintomas (classe 6) não foi ilustrada espacialmente, para facilitar a visualização da composição das demais classes cuja escala de grandeza é bem distinta (Figura 3).

Tabela 1

Modelo múltiplo de regressão logística multinomial com estimativas de *odds ratio* (OR) e intervalos de 95% de confiança (IC95%) da análise de classes latentes com covariáveis sociodemográficas. Brasil, 2020.

	Todos os sintomas		Prevalência alta dos sintomas		Predominância de febre		Predominância de tosse/dor de garganta		Leves sintomas com predominância de dor de cabeça	
	OR	IC95%	OR	IC95%	OR	IC95%	OR	IC95%	OR	IC95%
Sexo										
Masculino	1,00		1,00		1,00		1,00		1,00	
Feminino	1,31	1,13-1,53	1,10	0,97-1,25	2,13	1,80-2,51	2,11	1,91-2,35	0,75	0,61-0,93
Cor										
Branca	1,00		1,00		1,00		1,00		1,00	
Preta	1,19	0,92-1,53	1,10	0,91-1,34	1,55	1,24-1,93	1,15	0,98-1,34	1,06	0,80-1,39
Amarela e indígena	1,43	0,95-2,15	1,80	1,08-3,01	1,02	0,50-2,05	1,33	0,98-1,81	1,47	0,94-2,28
Parda	1,34	1,13-1,58	1,16	1,02-1,31	1,20	1,04-1,39	1,06	0,97-1,15	1,03	0,91-1,18
Escolaridade										
Sem instrução	1,00		1,00		1,00		1,00		1,00	
Ensino Fundamental incompleto	2,01	1,41-2,86	1,32	0,99-1,76	1,26	0,84-1,87	1,97	1,40-2,87	0,89	0,67-1,18
Ensinos Fundamental e Médio	3,22	2,28-4,54	1,84	1,33-2,55	1,36	0,66-2,79	2,14	1,45-3,15	0,84	0,60-1,18
Ensino Superior e pós-graduação	3,18	2,23-4,52	1,24	0,85-1,81	1,65	0,82-3,30	3,04	2,01-4,61	1,27	0,71-2,26
Idade										
Quartil 1 (até 18 anos)	1,00		1,00		1,00		1,00		1,00	
Quartil 2 (19-36 anos)	2,88	2,27-3,66	1,55	1,28-1,87	7,07	2,52-19,86	5,30	3,73-7,54	0,92	0,65-1,29
Quartil 3 (37-53 anos)	3,64	2,90-4,57	1,73	1,38-2,17	9,64	3,34-27,84	6,01	4,12-8,78	0,71	0,43-1,16
Quartil 4 (53 anos ou mais)	2,23	1,71-2,91	1,17	0,95-1,44	9,63	2,33-39,73	3,63	2,27-5,81	1,25	0,65-2,38
Região										
Sudeste	1,00		1,00		1,00		1,00		1,00	
Norte	8,35	5,21-13,39	9,25	5,79-14,78	1,27	0,69-2,36	1,04	0,86-1,25	0,54	0,27-1,06
Nordeste	2,30	1,69-3,14	5,51	3,68-8,25	0,83	0,45-1,53	1,09	0,95-1,24	0,63	0,43-0,92
Sul	0,27	0,13-0,57	0,29	0,13-0,64	1,11	0,73-1,69	1,14	1,01-1,29	0,90	0,73-1,11
Centro-oeste	0,34	0,15-0,76	0,73	0,48-1,11	0,69	0,54-0,87	0,82	0,69-0,97	0,51	0,38-0,68

Tabela 2

Distribuição percentual das classes latentes de sintomas da COVID-19 segundo perfil de uso dos serviços de saúde. Brasil, 2020.

Variáveis	Todos os sintomas (n = 1.315; 0,38%)	Prevalência alta dos sintomas (n = 4.204; 1,21%)	Predominância de febre (n = 2.391; 0,70%)	Predominância de tosse/dor de garganta (n = 9.598; 2,77%)	Leves sintomas com predominância de dor de cabeça (n = 6.361; 1,84%)	Ausência de sintomas (n = 322.310; 93,10%)	Total (N = 346.179; 100,00%)
Procurou por serviço de saúde *							
Sim	47,00 **	31,90 **	28,80 **	13,30	12,70	8,00	15,50
Não	53,00	68,10	71,20	86,70 **	87,30 **	92,00 **	84,50
Providência tomada							
Ficou em casa *							
Sim	94,30 **	91,00 **	92,00 **	84,80 **	83,10	80,20	83,90
Não	5,70	9,00	8,00	15,20	16,90	19,80 **	16,10
Ligou para profissional de saúde *							
Sim	23,00 **	16,30 **	11,30 **	5,10	5,50	3,10	6,10
Não	77,00	83,70	88,70	94,90 **	94,50	96,90 **	93,90
Automedicação *							
Sim	68,70 **	72,10 **	78,20 **	55,70	55,20	54,10	58,00
Não	31,30	27,90	21,80	44,30 **	44,80 **	45,90 **	42,00
Medicação prescrita por médico *							
Sim	27,10 **	19,60 **	11,40	12,00	18,00 **	13,70	14,70
Não	72,90	80,40	88,60 **	88,00 **	82,00	86,30 **	85,30
Recebeu visita de profissional do SUS *							
Sim	6,00 **	4,00	4,40	2,80	3,80 **	2,40	3,10
Não	94,00	96,00	95,60	97,20	96,20	97,60 **	96,90
Recebeu visita de profissional particular *							
Sim	4,10 **	2,00 **	1,90 **	0,80	0,60	0,60	0,90
Não	95,90	98,00	98,10	99,20	99,40 **	99,40 **	99,10
Local onde buscou atendimento							
UBS/Equipe de saúde da família *							
Sim	47,50	48,20	46,90	54,20 **	48,90	41,60	48,20
Não	52,50	51,80	53,10	45,80	51,10	58,40 **	51,80
UPA *							
Sim	34,60 **	28,30 **	26,00	18,50	15,90	18,00	22,90
Não	65,40	71,70	74,00	81,50 **	84,10 **	82,00 **	77,10
Hospital SUS *							
Sim	24,10	20,60	21,40	14,80	18,50	20,40	19,50
Não	75,90	79,40	78,60	85,20 **	81,50	79,60	80,50
Consultório particular							
Sim	8,30	7,90	7,80	9,20	10,50	12,40	9,40
Não	91,70	92,10	92,20	90,80	89,50	87,60	90,60

(continua)

Tabela 2 (continuação)

Variáveis	Todos os sintomas (n = 1.315; 0,38%)	Prevalência alta dos sintomas (n = 4.204; 1,21%)	Predominância de febre (n = 2.391; 0,70%)	Predominância de tosse/dor de garganta (n = 9.598; 2,77%)	Leves sintomas com predominância de dor de cabeça (n = 6.361; 1,84%)	Ausência de sintomas (n = 322.310; 93,10%)	Total (N = 346.179; 100,00%)
Pronto-socorro particular *							
Sim	4,50	6,10 **	4,10	1,90	2,20	3,00	3,70
Não	95,50	93,90	95,90	98,10 **	97,80 **	97,00	96,30
Hospital particular							
Sim	14,30	13,10	10,60	10,00	11,80	11,30	11,80
Não	85,70	86,90	89,40	90,00	88,20	88,70	88,20
Internado *							
Sim	23,30 **	13,20	8,20	7,80	3,80	9,40	10,90
Não	70,60	85,20	90,20	91,40 **	95,00 **	90,20	87,30
Não conseguiu	6,10 **	1,70	1,60	0,80	1,20	0,50	1,80

SUS: Sistema Único de Saúde; UBS: unidade básica de saúde; UPA: unidade de pronto-atendimento.

* Teste de Rao-Scott com valor de $p < 0,05$;

** Resíduo padronizado $> 1,96$.

Discussão

A ACL possibilitou a compreensão do modo como os sintomas relacionados à COVID-19 se agrupam entre os indivíduos com padrões prodrômicos semelhantes. Nesse sentido, mostra-se como uma alternativa superior à imposição de um ponto de corte, que apenas classifica se o indivíduo teve um ou mais sintomas. Muitas vezes, podemos ter indivíduos com o mesmo número de sintomas, mas com forte distinção na sua natureza. Considerando a epidemiologia e patogenia da doença, os sintomas mais comuns no início da doença de COVID-19 são febre, tosse e fadiga, ao passo que outros sintomas incluem produção de escarro, dor de cabeça, hemoptise, diarreia, dispneia e linfopenia²⁴. Dessa forma, outros sintomas como diarreia e hemoptise poderiam ser adicionados à investigação sobre a sintomatologia, de modo a acolher o maior número de casos suspeitos.

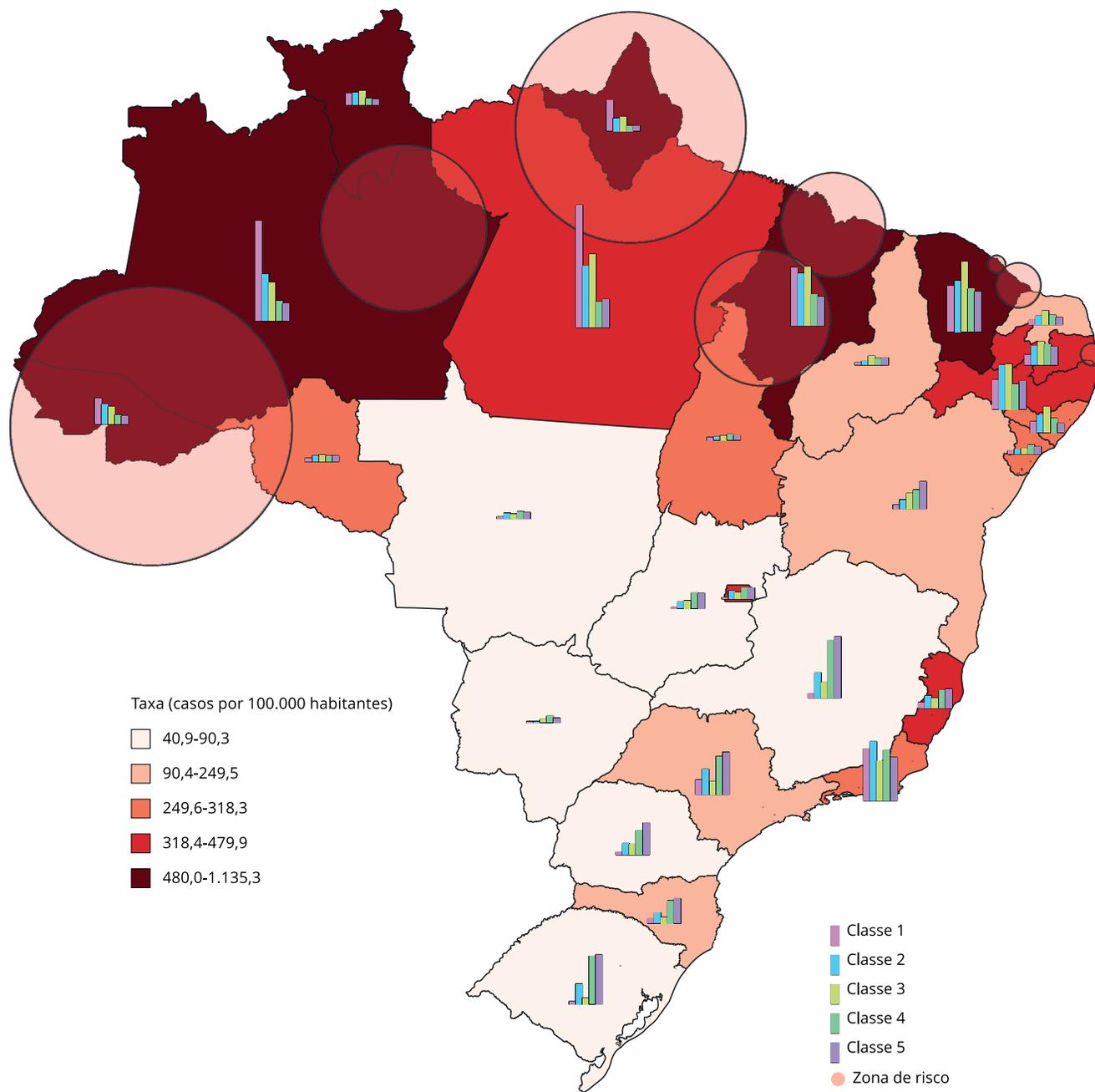
O perfil demográfico associado com a classe todos os sintomas (classe 1) foi formado principalmente por mulheres pardas, provenientes das regiões Norte e Nordeste, em todas as três faixas etárias mais velhas, em comparação com a faixa mais jovem. Nesse sentido, não se restringiu ao quartil mais velho (53 anos ou mais), apontando os grupos de adultos jovens (19-36 anos) e adultos (37-53 anos) com mais chances de pertencimento a todas as classes, exceto à classe com sintomas leves (classe 5). Ter algum nível de escolaridade esteve associado às classes 1, 2 e 4. Esse fato pode sugerir maior entendimento e capacidade de vocalização sobre a dimensão dos sintomas autorreferidos.

Observou-se que houve procura por serviços de saúde na presença de três classes de sintomas, desde todos os sintomas até predominância de febre. Por um lado, esse achado mostra a preocupação com a procura de assistência médica na presença desses sintomas em geral, mesmo que seja por uma febre. Isso pode ser reflexo das campanhas de divulgação dos sintomas mais comuns. Por outro lado, pode gerar maior probabilidade de infecção entre aqueles que não são confirmados, além de gerar maior sobrecarga nos serviços de saúde. Nesses casos, as iniciativas de teleconsulta surgem como possibilidade de acompanhamento dos pacientes classificados nas formas menos graves dos sintomas.

Apesar do crescimento nos últimos anos da telessaúde no Brasil, a epidemia acelerou a proliferação normativa e regulatória do campo, haja vista a potencialidade dessa estratégia no enfrentamento da COVID-19. As possibilidades de cuidado e tratamentos remotos, triagem, prevenção, monitoramento, detecção e vigilância configuram um cenário de fortalecimento da telessaúde na prática dos

Figura 3

Distribuição espacial das classes latentes de sintomas (barras), taxa de casos de COVID-19 (por 100 mil habitantes em quintis) e áreas de risco espacial na taxa de casos de COVID-19 (círculos). Brasil, 2020.



serviços no território. Suas aplicações vão desde a teleconsultoria, passando pelo telediagnóstico, telemonitoramento, telerregulação, teleducação, segunda opinião informativa e teleconsulta. Todavia, sua plena consolidação depende de investimentos em modelos de organização, sistemas, serviços, recursos humanos e infraestrutura, implicando maior destinação de recursos financeiros e tempo para sua constituição²⁵. Considerando a emblemática desigualdade social brasileira e o acelerado crescimento da pandemia, ainda estamos distantes de uma cobertura universal desses serviços.

O fato de terem sido as UPAs do SUS o local onde se buscou atendimento para aqueles que apresentaram maior prevalência de todos os sintomas indica a UPA como porta de entrada para a maioria dos indivíduos polissintomáticos. Esse achado sugere esse ponto da rede como prioritário para testagem, e não os hospitais ou UBS. Contudo, Medina et al.²⁶ defendem o protagonismo da atenção primária à saúde (APS) no enfrentamento da pandemia. Propõem a atuação da APS em quatro eixos: (1) vigilância em saúde nos territórios; (2) atenção aos usuários com COVID-19; (3) suporte social aos grupos vulneráveis e (4) continuidade das ações próprias da APS.

No que se refere ao eixo da vigilância em saúde, a capilaridade própria da APS, orientada pela Estratégia Saúde da Família (ESF), proporciona a detecção, notificação e acompanhamento dos casos. Entrementes, o papel de liderança comunitária e a proximidade territorial dos agentes comunitários de saúde (ACS) favorecem o engajamento nas medidas de isolamento social e empoderamento local, bem como a disseminação de informações verdadeiras a apoio nas atividades educativas. Na atenção aos usuários com COVID-19, é importante a organização de fluxos distintos dependentes da gravidade da doença, expressa pela sintomatologia e pelo histórico de contatos com pessoas infectadas. Nesse eixo, a triagem, a teleconsulta, o monitoramento dos sintomas e a articulação com os outros pontos da rede de cuidado são fundamentais para a continuidade da atenção à saúde. O suporte social aos grupos vulneráveis, especialmente idosos e portadores de comorbidades associadas, deve ser articulado em conjunto com os demais mecanismos de proteção social. Nesse eixo, destacam-se ações como abrigos em hotéis, escolas, apoio domiciliar e outras iniciativas de solidariedade comunitária. Finalmente, não podem ser negligenciadas as outras ações de rotina, próprias da APS, que buscam manter a população protegida, seja pela continuidade do calendário vacinal, seja pelos monitoramentos dos principais agravos e acompanhamentos dos indivíduos cujas morbidades são passíveis de intervenção e resolução nesse eixo²⁶. Nesse sentido, a ACL das informações sistematicamente coletadas pela PNAD-COVID19 se mostrou como uma potente ferramenta no apoio aos quatro eixos da APS, na medida em que permite a vigilância, a triagem e o monitoramento para a adequada articulação com os demais pontos da rede.

Entretanto, observou-se um excesso de indivíduos classificados com todos os sintomas (classe 1) que buscaram internação, mas não conseguiram. Esse fato chama a atenção para o colapso dos sistemas de saúde que não foram capazes, em alguns casos, de atender à demanda de internação dos pacientes mais graves.

Recente estudo longitudinal realizado em Londres, Inglaterra, por meio do aplicativo *COVID Symptom Study Smartphone*, detectou e validou seis grupos/clusters distintos de sintomas associados à COVID-19 que poderiam demandar diferentes níveis de cuidados médicos²⁷. Foram utilizadas 14 perguntas sobre os sintomas referidos pelos participantes da pesquisa. Os *clusters* 1 e 2 representaram formas de COVID-19 com 1,5% e 4,4%, respectivamente, exigindo suporte respiratório. Esses clusters mostraram predominantemente sintomas do trato respiratório superior e foram distinguidos entre si pela ausência de dor no *cluster* 2 em comparação com o *cluster* 1 e relatos ligeiramente aumentados de febre no *cluster* 2. O grupo 3 mostrou sintomas gastrointestinais mais fortes isoladamente (diarreia, refeições ignoradas) e uma redução relativa de necessidade de suporte respiratório (3,7%). Os *clusters* 4, 5 e 6 incluíram participantes que relataram COVID-19 mais grave, com 8,6%, 9,9% e 19,8% dos indivíduos que necessitavam de suporte respiratório, respectivamente. Esses três grupos representaram distintas apresentações, com o *cluster* 4, marcado pela presença precoce de fadiga severa, e o *cluster* 5 pela presença contínua de dor e tosse persistente. Por sua vez, os indivíduos do *cluster* 5 relataram confusão e fadiga severa. Finalmente, os indivíduos do *cluster* 6 relataram sintomas mais acentuados de dificuldade respiratória, incluindo início precoce de falta de ar, acompanhada de dor no peito.

No presente estudo, embora a maior parte dos indivíduos tenha sido agrupada na classe ausente de sintomas, o conhecimento específico das características associadas às demais classes latentes é de fundamental importância na identificação de grupos de maior risco. Entrementes, permite compreender

o padrão de uso dos serviços de saúde, no sentido de ofertar a melhor estratégia diagnóstica e assistiva no momento oportuno. Nesse sentido, buscando aplicar o pensamento epidemiológico sobre os critérios de diagnóstico de COVID-19, Tan ²⁸ aponta como o melhor método de detecção aquele que ocorre em três estágios: caso suspeito, caso clínico diagnosticado e caso definitivo diagnosticado. O Centro de Controle e Prevenção de Doenças (CDC), dos Estados Unidos, indica pacientes com COVID-19 como aqueles que apresentam tosse, falta de ar ou dificuldade de respirar, febre, calafrios, dor muscular, dor de garganta e perda de paladar ou olfato, além de considerar também os contatos próximos com pacientes com COVID-19 confirmados ²⁹.

Com efeito, a análise espacial mostrou a sobreposição das regiões de maior taxa de casos e das áreas circulares de maior risco com a maior frequência de indivíduos polissintomáticos (classe 1). Tais achados fornecem uma aproximação entre os sintomas e os casos confirmados, corroborando essa associação. Contudo, essa aproximação só pode ser feita em âmbito ecológico, considerando o efeito de agregado. Nesse sentido, não podemos atribuir aos indivíduos classificados com todos os sintomas (classe 1) o diagnóstico de COVID-19, evitando, assim, a falácia ecológica.

Em apertada síntese, considerando a hecatombe pandêmica e a ausência de uma vacina ou de um tratamento farmacológico cientificamente sustentado, a observação da tríade epidemiológica pessoa, espaço e tempo deve ser revisitada dentro do mesmo espírito investigativo de John Snow ³⁰, no enfrentamento da epidemia da cólera em Londres, ainda no século XIX. Sendo assim, o registro dos diferentes perfis sintomatológicos deve se concretizar como uma atividade contínua dos sistemas de vigilância epidemiológica. Os resultados deste estudo mostraram semelhanças entre os fatores associados aos casos de COVID-19 e às classes latentes de sintomas. A diferença está na facilidade operacional baseada no autorrelato dos sintomas registrados pela PNAD-COVID19 em relação ao cenário ainda distante da realização de testagens em massa.

Informação adicional

ORCID: Rafael da Silveira Moreira (0000-0003-0079-2901).

Agradecimentos

A André Luiz Sá de Oliveira, do Núcleo de Estatística e Geoprocessamento do Instituto Aggeu Magalhães, Fundação Oswaldo Cruz (Pernambuco, Brasil), pela excelente edição vetorial do mapa.

Referências

1. Zhu N, Zhang D, Wang W, Li X, Yang B, Song J, et al. A novel coronavirus from patients with pneumonia in China, 2019. *N Engl J Med* 2020; 382:727-33.
2. World Health Organization. Novel coronavirus – China. <http://www.who.int/csr/don/12-january-2020-novel-coronavirus-china/en/> (acessado em 05/Ago/2020).
3. Tsang KW, Ho PL, Ooi GC, Yee WK, Wang T, Chan-Yeung M, et al. A cluster of cases of severe acute respiratory syndrome in Hong Kong. *N Engl J Med* 2003; 348:1975-83.
4. Drosten C, Gunther S, Preiser W, van der Werf S, Brodt HR, Becker S, et al. Identification of a novel coronavirus in patients with severe acute respiratory syndrome. *N Engl J Med* 2003; 348:1967-76.
5. Yang X, Yu Y, Xu J, Shu H, Xia J, Liu H, et al. Clinical course and outcomes of critically ill patients with SARS-CoV-2 pneumonia in Wuhan, China: a single-centered, retrospective, observational study. *Lancet Respir Med* 2020; 8:475-81.

6. Jefferson T, Mar CBD, 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.
7. Rothe C, Schunk M, Sothmann P, Bretzel G, Froeschl G, Wallrauch C, et al. Transmission of 2019-nCoV infection from an asymptomatic contact in Germany. *N Engl J Med* 2020; 382:970-1.
8. Yu P, Zhu J, Zhang Z, Han Y. A familial cluster of infection associated with the 2019 novel coronavirus indicating possible person-to-person transmission during the incubation period. *J Infect Dis* 2020; 221:1757-61.
9. Li C, Ji F, Wang L, Wang L, Hao J, Dai M, et al. Asymptomatic and human-to-human transmission of SARS-CoV-2 in a 2-family cluster, Xuzhou, China. *Emerg Infect Dis* 2020; 26:1626-8.
10. Li X, Guan P, Wu X, Wang L, Zhou Y, Tong Y, et al. Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia. *N Engl J Med* 2020; 382:1199-207.
11. Wang W, Tang J, Wei F. Updated understanding of the outbreak of 2019 novel coronavirus (2019-nCoV) in Wuhan, China. *J Med Virol* 2020; 92:441-7.
12. To KKW, Tsang OT, Yip CC, Chan K, Wu T, Chan JM, et al. Consistent detection of 2019 novel coronavirus in saliva. *Clin Infect Dis* 2020; 71:841-3.
13. Kim YJ, Sung H, Ki CS, Hur M. COVID-19 Testing in South Korea: current status and the need for faster diagnostics. *Ann Lab Med* 2020; 40:349-50.
14. Hopman J, Allegranzi B, Mehtar S. Managing COVID-19 in low- and middle-income countries. *JAMA* 2020; 323:1549-50.
15. Lima NVT, Buss PM, Paes-Sousa R. A pandemia de COVID-19: uma crise sanitária e humanitária. *Cad Saúde Pública* 2020; 36:e00177020.
16. Instituto Brasileiro de Geografia e Estatística. PNAD-COVID19: plano amostral e ponderação. <https://biblioteca.ibge.gov.br/visualizacao/livros/liv101727.pdf> (acessado em 05/Ago/2020).
17. Moreira RS. COVID-19: unidades de terapia intensiva, ventiladores mecânicos e perfis latentes de mortalidade associados à letalidade no Brasil. *Cad Saúde Pública* 2020; 36:e00080020
18. Nylund KL, Asparouhov T, Muthén BO. Deciding on the number of classes in latent class analysis and growth mixture modeling: a Monte Carlo simulation study. *Struct Equ Modeling* 2007; 14:535-69.
19. Rao JNK, Scott AJ. On chi-squared tests for multiway contingency tables with cell proportions estimated from survey data. *Ann Stat* 1984; 12:46-60.
20. Pereira JCR. *Análise de dados qualitativos: estratégias metodológicas para as ciências da saúde, humanas e sociais*. São Paulo: Edusp; 2004.
21. Haberman SJ. *Analysis of qualitative data: introductory topics*. New York: Academic Press; 1978.
22. Wang J, Wang X. *Structural equation modeling: applications using Mplus*. Chichester: John Wiley & Sons; 2012.
23. Kulldorff MA. Spatial scan statistic. *Commun Stat Theory Methods* 1997; 26:1481-96.
24. Rothan HA, Byrareddy SN. The epidemiology and pathogenesis of coronavirus disease (COVID-19) outbreak. *J Autoimmun* 2020; 109:102433.
25. Caetano R, Silva AB, Guedes ACCM, Paiva CCN, Ribeiro GR, Santos DL, et al. Desafios e oportunidades para telessaúde em tempos da pandemia pela COVID-19: uma reflexão sobre os espaços e iniciativas no contexto brasileiro. *Cad Saúde Pública* 2020; 36:e00088920.
26. Medina MG, Giovanella L, Bousquat A, Mendonça MHM, Aquino R. Atenção primária à saúde em tempos de COVID-19: o que fazer? *Cad Saúde Pública* 2020; 36:e00149720.
27. Sudre C, Lee K, Lochlainn M, et al. Symptom clusters in covid19: a potential clinical prediction tool from the COVID Symptom study app. *medRxiv* 2020; 6 jun. <https://www.medrxiv.org/content/10.1101/2020.06.12.20129056v1>.
28. Tan HZ. An epidemiologic thinking on the diagnosis criteria of COVID-19. *Zhonghua Liu Xing Bing Xue Za Zhi* 2020; 41:998-9.
29. Centers for Disease Control and Prevention. Novel coronavirus, (2019). <https://www.cdc.gov/coronavirus/2019-ncov/about/transmission.html> (acessado em 05/Ago/2020).
30. Cameron D, Jones IG. John Snow, the broad street pump and modern epidemiology. *Int J Epidemiol* 1983; 12:393-6.

Abstract

The lack of mass testing for COVID-19 diagnosis creates the need to determine the magnitude of the disease based on its clinical symptoms. The study aimed to analyze the profile of COVID-19 symptoms and related aspects in Brazil. The author analyzed the sample of participants from the Brazilian National Household Sample Survey (PNAD-COVID19) conducted in May 2020. Latent class analysis (LCA) was performed with sociodemographic covariables and 11 symptoms reported by 346,181 individuals. Rao-Scott test and standardized residual analysis were used to measure the association with the pattern of health services use. Spatial scan analysis was performed to identify areas at risk of COVID-19 cases. LCA showed six classes of symptoms based on the pattern of answers by participants: (1) all the symptoms; (2) high prevalence of symptoms; (3) predominance of fever; (4) predominance of cough/sore throat; (5) mild symptoms with predominance of headache; and (6) absence of symptoms. Female sex, brown skin color, the North and Northeast regions of Brazil, and all three older age brackets showed stronger association with the class with all the symptoms (class 1). Most use of health services was also by this group of individuals, but with different profiles of use. Spatial analysis showed juxtaposition of this class with areas at greater risk of COVID-19. These findings underline the importance of investigating symptoms for the epidemiological identification of possible cases in a scenario with low population testing rates.

COVID-19; Latent Class Analysis; Pandemics; Epidemiological Monitoring

Resumen

La ausencia de pruebas masivas para el diagnóstico de la COVID-19 genera la necesidad de conocer la dimensión de la enfermedad mediante su sintomatología clínica. El objetivo fue investigar el perfil de síntomas relacionados con la COVID-19 y sus aspectos relacionados. Se analizó la muestra de participantes de la Encuesta Nacional por Muestra de Domicilios (PNAD-COVID 19), realizada en mayo de 2020. Se efectuó el análisis de clases latentes (ACL) con covariables sociodemográficas sobre 11 síntomas mencionados por 346.181 individuos. Se utilizaron pruebas de Rao-Scott y análisis de residuos estandarizados para medir la asociación con el patrón de utilización de los servicios de salud. Se efectuó un análisis espacial de barrido para identificar las áreas de riesgo con casos de COVID-19. La ACL mostró seis clases de sintomatología, según el patrón de respuestas de los individuos analizados: (1) todos los síntomas; (2) prevalencia alta de los síntomas; (3) predominancia de fiebre; (4) predominancia de tos/dolor de garganta; (5) leves síntomas con predominancia de dolor de cabeza y; (6) ausencia de síntomas. Personas del sexo femenino, mestizos/mulatos, provenientes de las regiones Norte y Nordeste y en las tres franjas etarias más viejas presentaron una mayor asociación con la clase de todos los síntomas (clase 1). La mayoría de búsquedas de servicios también se realizó por parte de este grupo de individuos, aunque con distintos perfiles de uso. El análisis espacial mostró superposición de esta clase con áreas de mayor riesgo de casos de COVID-19. Estos hallazgos sostienen la importancia de la investigación de los síntomas, sirviendo para la identificación epidemiológica de posibles casos en un escenario con baja tasa de pruebas entre la población.

COVID-19; Análisis de Clases Latentes; Pandemias; Monitoreo Epidemiológico

Recebido em 10/Ago/2020
Versão final reapresentada em 06/Out/2020
Aprovado em 09/Out/2020