

ConVid – Pesquisa de Comportamentos pela Internet durante a pandemia de COVID-19 no Brasil: concepção e metodologia de aplicação

ConVid – Behavior Survey by the Internet during the COVID-19 pandemic in Brazil: conception and application methodology

ConVid – Encuesta de Comportamientos por Internet durante la pandemia de COVID-19 en Brasil: concepción y metodología de aplicación

Celia Landmann Szwarcwald ¹
Paulo Roberto Borges de Souza Júnior ¹
Giseli Nogueira Damascena ¹
Deborah Carvalho Malta ²
Marilisa Berti de Azevedo Barros ³
Dalia Elena Romero ¹
Wanessa da Silva de Almeida ¹
Luiz Otávio Azevedo ¹
Ísis Eloah Machado ⁴
Margareth Guimarães Lima ³
André Oliveira Werneck ⁵
Danilo Rodrigues Pereira da Silva ⁶
Crizian Saar Gomes ⁷
Arthur Pate de Souza Ferreira ¹
Renata Gracie ¹
Maria de Fátima de Pina ¹

doi: 10.1590/0102-311X00268320

Resumo

A ConVid – Pesquisa de Comportamentos foi realizada no Brasil de 24 de abril a 24 de maio de 2020, com o objetivo de investigar as mudanças nos estilos de vida e nas condições de saúde durante a pandemia de COVID-19. Neste artigo, apresentamos a concepção e metodologia da pesquisa. Estudo de corte transversal com a utilização de um questionário pela Internet, com questões validadas em inquéritos de saúde anteriores. O método de amostragem foi o “bola de neve virtual” e foram usados os procedimentos de pós-estratificação. Os resultados relativos às doenças crônicas não transmissíveis e estilos de vida pré-pandemia foram comparados às estimativas da Pesquisa Nacional de Saúde de 2013 e da Vigilância de Fatores de Risco e Proteção para Doenças Crônicas por Inquérito Telefônico de 2019. A amostra total foi de 45.161 pessoas. Após a ponderação dos dados, as distribuições amostrais das variáveis demográficas foram semelhantes às populacionais. Apenas as pessoas de baixo nível de instrução foram sub-representadas. A comparação com os resultados anteriores mostrou similaridade na maioria das estimativas: consumo recomendado de frutas e legumes (22,1%), atividade física recomendada (35,2%), fumo de cigarros (12,3%), consumo frequente e abusivo de álcool (6,7%), obesidade (21,2%), prevalências autorreferidas de hipertensão (18,6%), diabetes (7,1%) e doença do coração (4,4%). O inquérito online possibilitou conhecer as condições de saúde da população durante a pandemia. A similaridade dos indicadores com os obtidos em pesquisas tradicionais permitiu validar as estimativas médias. Estudos são necessários para investigar como os efeitos endógenos das redes sociais virtuais podem ser levados em consideração na estimação da variância.

COVID-19; Internet; Inquéritos Epidemiológicos

Correspondência

C. L. Szwarcwald
Departamento de Informações para a Saúde, Instituto de Comunicação e Informação Científica e Tecnológica em Saúde, Fundação Oswaldo Cruz,
Av. Brasil 4365, Biblioteca de Manguinhos, sala 225,
Rio de Janeiro, RJ 21040-360, Brasil.
celia.szwarcwald@icict.fiocruz.br

¹ Instituto de Comunicação e Informação Científica e Tecnológica em Saúde, Fundação Oswaldo Cruz, Rio de Janeiro, Brasil.

² Escola de Enfermagem, Universidade Federal de Minas Gerais, Belo Horizonte, Brasil.

³ Faculdade de Ciências Médicas, Universidade Estadual de Campinas, Campinas, Brasil.

⁴ Escola de Medicina, Universidade Federal de Ouro Preto, Ouro Preto, Brasil.

⁵ Faculdade de Saúde Pública, Universidade de São Paulo, São Paulo, Brasil.

⁶ Centro de Ciências Biológicas e da Saúde, Universidade Federal de Sergipe, São Cristóvão, Brasil.

⁷ Faculdade de Medicina, Universidade Federal de Minas Gerais, Belo Horizonte, Brasil.



Introdução

A grande expansão da Internet, a partir da década de 1990, tornou a comunicação à distância muito mais abrangente e acessível, aproximando pessoas de grupos e culturas distintos. No Brasil, a Internet surgiu no final da década de 1980 como meio de divulgação de informações na área de ensino e pesquisa, mas, rapidamente, atingiu a população geral. Em 2018, estimou-se que 79,1% dos domicílios particulares permanentes tinham acesso à Internet ¹.

Nos últimos 20 anos, as tecnologias de informação e comunicação têm crescido intensamente, promovendo profundas modificações na maneira como as pessoas se comunicam, captam e compartilham informações, e emitem opiniões ^{2,3,4}. Ocorre uma grande proliferação de *sites*, *chats*, redes sociais, tornando o mundo uma enorme rede de pessoas conectadas, independentemente de suas localizações geográficas ⁵.

Tendo em vista que as abordagens tradicionais de coleta de informações de pesquisas de base populacional nem sempre são viáveis devido a restrições orçamentárias, limites de tempo para a obtenção das informações, ou ausência de uma lista ou sistema de referências, a elaboração de pesquisas em ambiente virtual é uma tendência atual para a coleta de dados ⁶, e tem estimulado pesquisadores a utilizarem questionários *online* como um método alternativo para a obtenção de respostas em pesquisas científicas, incluindo inquéritos na área de saúde ^{7,8}. As informações podem ser coletadas por uma plataforma única para a coleta de dados ⁹ (como Facebook ou um sítio institucional) ou utilizando um procedimento de amostragem em cadeia, como a metodologia conhecida como “bola de neve” adaptada às redes sociais virtuais ¹⁰.

Norteadas pela teoria dos seis graus de separação ¹¹, as técnicas de amostragem em cadeia têm sido usadas, frequentemente, para facilitar o acesso a populações consideradas de difícil acesso pelo fato de adotarem comportamentos estigmatizados socialmente ou praticarem atividades ilegais ¹². O método “bola de neve” é um procedimento de amostragem não probabilístico, que funciona a partir da indicação de um grupo inicial de pessoas que fazem parte da população-alvo (denominadas de sementes), que indicam pares do mesmo grupo populacional, e assim sucessivamente, semelhante à formação de uma bola de neve ¹³.

De maneira semelhante, o método de levantamento de dados por “bola de neve virtual” é iniciado pelo envio/apresentação do *link* de acesso a um questionário eletrônico, por meio de *e-mail* ou de alguma rede social virtual. No corpo da mensagem, além da apresentação da pesquisa, há um pedido para que o convite de participação seja compartilhado com a rede de contatos da pessoa. Essa técnica não probabilística de coleta de dados permite a definição de uma amostra por meio de referências feitas por pessoas, que indicam outras que satisfaçam os critérios de inclusão na pesquisa, e assim sucessivamente ¹⁰.

Com a chegada da pandemia de COVID-19, a adoção das medidas de restrição de contato físico em vários países do mundo limitou a realização de pesquisas de saúde com entrevistas presenciais ¹⁴. Por outro lado, a necessidade de aquisição de maior conhecimento sobre a doença e de outros problemas de saúde relacionados direta ou indiretamente à pandemia estimulou o uso da Internet como meio para a obtenção das informações de saúde de forma ágil ^{15,16}. Além da aderência às medidas de proteção e de restrição de contatos físicos ¹⁷, os estudos abordaram outras questões relacionadas ao período de pandemia como sintomas ¹⁸, transtornos psicológicos ¹⁹, dificuldades de acesso aos serviços de saúde ²⁰ e a adoção de comportamentos não saudáveis ^{21,22}.

No Brasil, a partir de março de 2020, medidas rígidas de restrição de contatos entre as pessoas foram impostas pelos estados e municípios. O distanciamento de familiares e amigos, a incerteza sobre a doença, as mudanças substanciais no contexto socioeconômico e a falta de controle sobre a própria vida têm provocado danos à saúde física e mental ^{23,24}.

Com a finalidade de investigar as mudanças nos estilos de vida e nas condições de saúde da população brasileira durante a pandemia de COVID-19, foi realizada a *ConVid – Pesquisa de Comportamentos* em ambiente virtual. No presente trabalho, apresentamos a concepção, os objetivos e a metodologia, e analisa-se a representação da população na amostra alcançada.

Metodologia

Concepção e desenho do estudo

A *ConVid – Pesquisa de Comportamentos* (<https://convid.fiocruz.br/>) foi realizada, em âmbito nacional, pela Fundação Oswaldo Cruz (Fiocruz), em parceria com a Universidade Federal de Minas Gerais e a Universidade Estadual de Campinas. Trata-se de um estudo de corte transversal com a utilização de um questionário virtual, autopreenchido por meio de celular ou computador com acesso à Internet. A coleta de dados ocorreu entre 24 de abril e 24 de maio de 2020.

A pesquisa teve os seguintes objetivos: descrever a intensidade de aderência da população brasileira às medidas de restrição social; investigar as mudanças na situação de trabalho e rendimento; avaliar as dificuldades na realização das atividades de rotina; analisar as condições de saúde e descrever as mudanças de comportamentos adotados durante a pandemia de COVID-19. Foram utilizadas questões validadas em inquéritos de saúde aplicados, previamente, no Brasil. As perguntas sobre o estado de ânimo foram adaptadas da *Pesquisa Mundial de Saúde* ²⁵, e as referentes ao diagnóstico de doenças crônicas não transmissíveis e aos estilos de vida (hábitos de fumo, alimentação, consumo de bebidas alcoólicas, prática de atividade física e sedentarismo) foram fundamentadas das questões da *Pesquisa Nacional de Saúde* de 2013 (PNS 2013) ²⁶ e monitoradas pela *Vigilância de Fatores de Risco e Proteção para Doenças Crônicas por Inquérito Telefônico* de 2019 (Vigitel 2019) ²⁷. Quanto aos problemas no sono, foi usada uma questão única de autoavaliação do sono, adaptada do conjunto de questões sobre o sono do *Inquérito de Saúde de Base Populacional em Campinas, 2014-2015 (ISA-CAMP)* ²⁸. Para avaliar as dificuldades dos idosos (60 anos ou mais) em realizar as atividades da vida diária utilizou-se apenas uma pergunta, síntese de um conjunto de questões da PNS 2013 ²⁹. Em relação à situação de trabalho e perdas de rendimento, as questões foram baseadas nas usadas na pesquisa *Qualidade de Vida em Pacientes em Terapia Antirretroviral no Brasil*, realizada em 2008 ³⁰.

Para grande parte dos temas abordados foram feitas perguntas sobre a situação anterior e posterior à chegada da pandemia no Brasil. Foram avaliadas as mudanças na situação de trabalho, nas dificuldades de realização das atividades de rotina, nas condições de saúde e nos estilos de vida. O questionário completo está disponibilizado no *site* da ConVid (<https://convid.fiocruz.br/>).

Para a elaboração do questionário foi utilizado o aplicativo RedCap (Research Electronic Data Capture. <https://www.project-redcap.org/>). As informações foram coletadas diretamente pela Internet e armazenadas no servidor do Instituto de Comunicação e Informação Científica e Tecnológica em Saúde (ICICT/Fiocruz). Os critérios de inclusão para a participação na pesquisa foram: ter 18 anos ou mais de idade e residir no território brasileiro durante a pandemia de COVID-19. Todas as respostas foram anônimas e sem qualquer tipo de identificação dos participantes. Maiores detalhes da *ConVid – Pesquisa de Comportamentos* estão no *site* da pesquisa (<https://convid.fiocruz.br/>).

O projeto foi aprovado pela Comissão Nacional de Ética em Pesquisa (CONEP) em 19 de abril de 2020 (parecer nº 3.980.277).

Amostragem

O método de amostragem foi o “bola de neve virtual”, iniciado a partir do envio de convites com o *link* de acesso ao questionário eletrônico por meio da rede social virtual WhatsApp ou por *e-mail*. A pesquisa foi denominada de “ConVid” para sugerir que o participante deveria convidar outras pessoas de suas redes sociais: “*Se você achou importante, ConVide mais pessoas para participar!*”. Ao final do questionário, havia, igualmente, uma solicitação para que o convite de participação na pesquisa fosse compartilhado com a rede de contatos da pessoa: “*Faça parte da Rede ConVid e compartilhe esta pesquisa com três ou mais convidados da sua rede social*”.

Na primeira etapa do recrutamento de participantes, para obter abrangência nacional, 15 pesquisadores que compunham a equipe do estudo escolheram, aproximadamente, 200 outros pesquisadores de diferentes estados do Brasil. Adicionalmente, cada pesquisador da equipe do estudo selecionou 20 pessoas da sua rede social, totalizando cerca de 500 pessoas que foram chamadas de sementes por terem desencadeado a rede de participantes. Após responder ao questionário, as sementes constituíram a primeira onda da cadeia de recrutamento. Com vistas a obter diversidade das características

sociodemográficas na amostra ³¹, as sementes enviaram o *link* da pesquisa para pelo menos três pessoas das suas redes sociais em cada um dos 12 estratos formados por sexo, faixa de idade (18-39; 40-59; 60+ anos) e grau de escolaridade (Ensino Médio incompleto ou menos; Ensino Médio completo ou mais). As pessoas convidadas pelas sementes compuseram a segunda onda da cadeia de recrutamento. A cada pessoa da segunda onda foi solicitado que ela convidasse pelo menos outras três pessoas de suas redes sociais e assim por diante. Ao final do período de coleta das informações (24 de abril a 24 de maio de 2020), o tamanho total de amostra atingido foi de 45.161 pessoas.

Procedimento de pós-estratificação

Uma vez que as probabilidades de seleção são desconhecidas, a amostragem por redes não é probabilística, não sendo possível calcular os pesos naturais do desenho de amostragem.

Para obter uma amostra representativa da população de acordo com a localização geográfica e com as características sociodemográficas, foram realizadas ponderações calculadas por procedimentos de pós-estratificação ³² por: Unidade da Federação (UF), capital/restante da UF, sexo, faixa etária (18-29; 30-39; 40-49; 50-59; 60+ anos), grau de escolaridade (Superior incompleto; Superior completo) e raça/cor da pele com base nas estimativas populacionais da *Pesquisa Nacional de Amostra por Domicílios* de 2019 (PNAD 2019) ³³.

Matematicamente, seja Nh a estimativa populacional da PNAD 2019 ³³ em cada estrato h , composto por UF, capital/restante da UF, sexo, faixa etária, grau de instrução e raça; e N o número total de observações. Considerando nh o número de observações por estrato h na amostra da ConVid e n o total da amostra, o fator de ponderação (Wh) em cada estrato foi calculado por:

$$Wh = (Nh/nh).(n/N)$$

Antes do início da realização dos procedimentos de ponderação, os dados foram analisados quanto à presença de duplicidades (todas as respostas exatamente iguais) e de dados faltantes nas variáveis usadas para a pós-estratificação.

Representatividade

Para verificar a representação da população brasileira na amostra obtida na ConVid, as distribuições amostrais das características sociodemográficas utilizadas e não utilizadas na pós-estratificação (número de moradores no domicílio e condição na força de trabalho) foram comparadas com as da PNAD 2019 ³³.

Considerando-se as questões da ConVid sobre a situação anterior à chegada da pandemia no Brasil, os indicadores relativos aos comportamentos saudáveis foram comparados aos da PNS 2013 ^{29,34}, considerada o padrão-ouro dos inquéritos nacionais de saúde. Além disso, para ter uma comparação com dados mais atuais, foram utilizados, igualmente, indicadores da Vigitel 2019 ²⁷.

Foram considerados os seguintes indicadores de estilos de vida: proporção de pessoas com 18 anos ou mais com o consumo recomendado de frutas e hortaliças (FLV); proporção de pessoas com 18 anos ou mais com o consumo de feijão por 5 dias ou mais por semana (feijão 5 dias ou mais); proporção de pessoas com 18 anos ou mais com o consumo abusivo e frequente de bebida alcoólica, definido como o consumo de 8 ou mais doses de bebida alcoólica por semana para as mulheres e 15 ou mais doses para os homens (beber pesado); proporção de pessoas com 18 anos ou mais que são fumantes atuais de cigarros (fumantes de cigarros); proporção de pessoas com 18 anos ou mais que praticam o nível recomendado de atividade física no lazer (atividade física ≥ 150 minutos); proporção de pessoas com 18 anos ou mais que assistem à TV ou *tablet* no tempo livre por 3 horas ou mais (TV/*tablet* 3 horas ou mais).

Quanto à prática de atividade física no tempo recomendado (150 minutos de atividade física moderada ou 75 minutos de atividade física intensa por semana), uma vez que as questões relativas ao número de dias (1-2; 3-4; 5 ou mais dias) e ao tempo de prática de atividade física (< 30; 30-45; 46-59; 60 ou mais minutos) são categóricas na ConVid e não houve perguntas sobre o tipo de atividade física praticada, a estimativa correspondeu à média entre a proporção mínima (indivíduos que praticam

3-4 dias por semana por 46 minutos ou mais e 5 ou mais dias por semana e 30 ou mais minutos) e a máxima (indivíduos que praticam 1-2 dias por semana e 60 ou mais minutos ou 3-4 dias por semana por 30 minutos ou mais e 5 ou mais dias por semana).

Entre os idosos (60+ anos), foi considerado o indicador de limitação funcional que consiste em precisar de ajuda para a realização das atividades da vida diárias (limitação AVD), tais como comer, tomar banho, ir ao banheiro, se vestir, se locomover pela casa, ou se deitar.

Adicionalmente, os indicadores de morbidade referida foram também comparados aos obtidos na PNS 2013^{29,34} e no Vigitel 2019²⁷. Foram considerados os seguintes indicadores: proporções de pessoas com 18 anos ou mais de idade que referem diagnóstico de hipertensão arterial (hipertensão), de diabetes, de doença do coração e de depressão; proporção de pessoas com 18 anos ou mais de idade que referem diagnóstico de doenças crônicas não-transmissíveis de risco de agravamento de COVID-19 (DCNT risco COVID-19); proporção de pessoas com 18 anos ou mais com excesso de peso, definido pelo índice de massa corporal (IMC) maior ou igual a 25kg/m² (excesso de peso); proporção de pessoas com 18 anos ou mais com obesidade, definida pelo IMC maior ou igual a 30kg/m² (obesidade).

A presença de DCNT risco COVID-19 foi calculada com base no diagnóstico referido de pelo menos uma das seguintes doenças: diabetes, hipertensão, asma/enfisema/doença respiratória crônica ou outra doença do pulmão, doença do coração ou câncer.

Para cada um dos indicadores de estilos de vida e prevalências de DCNT autorreferidas, foram consideradas três diferentes estimativas: amostra total sem ponderação; amostra total com ponderação; e, para reduzir o viés de seleção³⁵ do procedimento de recrutamento por bola de neve, foi considerada a amostra sem as duas primeiras ondas, isto é, excluindo os primeiros 18.500 participantes (500 sementes que convidaram três pessoas em cada um dos 12 estratos).

Uma vez que não foram coletadas informações sobre a rede de contatos, não foi possível levar em consideração a dependência das observações na estimação da variância. No presente estudo, adotou-se a proposta de Goel & Salganik³⁵, e as estimativas dos indicadores de estilos de vida e das prevalências de DCNT autorreferidas foram calculadas com base na amostra da ConVid, sem considerar as duas primeiras cadeias de recrutamento. Os indicadores de cada pesquisa foram descritos utilizando-se estimativas de proporções e intervalos de 95% de confiança (IC95%).

Para verificar a diversidade da amostra da ConVid em termos geográficos, foi elaborado um mapa com todos os municípios do Brasil alcançados pela pesquisa.

Resultados

Ao todo, 47.184 pessoas aceitaram participar da pesquisa. Após a exclusão dos questionários com dados faltantes nas variáveis usadas para pós-estratificação (4,3%), um total de 45.161 questionários foi considerado para a análise de dados.

Na Tabela 1, comparam-se as distribuições das variáveis usadas para a ponderação da base de dados obtidas com os dados da PNAD 2019 e da ConVid. Percebem-se poucas diferenças entre as distribuições das variáveis utilizadas para ponderação. Em relação às variáveis não usadas para a pós-estratificação, tanto para o número de moradores no domicílio como na condição na força de trabalho as distribuições foram semelhantes.

Três diferentes estimativas dos indicadores de estilos de vida e prevalências de diagnóstico de DCNT estão apresentadas na Tabela 2. A primeira estimativa realizada com a amostra total sem ponderação em comparação à amostra ponderada mostra a influência do maior nível de instrução entre os participantes da ConVid: para o consumo regular de frutas e vegetais, o percentual foi maior em 7 pontos, e o percentual de fumantes foi menor em 1,3 ponto. Quanto às DCNT, a maior diferença ocorreu para a depressão. Já comparando as estimativas correspondentes à amostra com a exclusão das duas primeiras ondas com as obtidas na amostra total ponderada, percebem-se pequenas diferenças e mostram que as estimativas são significativas, mesmo após a exclusão dos participantes das duas primeiras cadeias de recrutamento.

Na Tabela 3, as estimativas de indicadores de estilos de vida realizadas por meio da ConVid são comparadas com as obtidas na PNS 2013^{29,34} e Vigitel 2019²⁷. Quanto ao indicador de consumo

Tabela 1

Comparação das distribuições das variáveis utilizadas e não utilizadas para a pós-estratificação dos dados da *ConVid - Pesquisa de Comportamentos* com as distribuições das variáveis na *Pesquisa Nacional por Amostra de Domicílios* de 2019 (PNAD 2019) ³³.

Variáveis/Categorias	Distribuição (%)	
	ConVid	PNAD 2019
Variáveis utilizadas na pós-estratificação		
Região		
Norte	7,7	7,9
Nordeste	25,0	26,3
Sudeste	45,3	43,4
Sul	15,1	14,7
Centro-oeste	6,9	7,6
Capital/Restante da UF		
Capital	25,1	24,7
Restante da UF	74,9	75,3
Sexo		
Masculino	47,0	47,3
Feminino	53,0	52,7
Faixa etária (anos)		
18-29	24,7	23,6
30-39	21,0	20,7
40-49	18,1	18,4
50-59	15,9	16,2
60+	20,3	21,0
Grau de escolaridade		
Médio completo	83,6	84,3
Superior completo	16,4	15,7
Raça/Cor da pele		
Branca	44,7	43,8
Preta	9,6	9,7
Parda	44,9	45,4
Outra	0,8	1,1
Variáveis não utilizadas na pós-estratificação		
Condição em relação à força de trabalho *		
Trabalha	71,6	70,7
Não trabalha	28,4	29,3
Número de moradores no domicílio		
1	9,5	7,2
2	26,4	24,1
3	27,3	27,8
4	22,6	23,0
5	9,5	10,4
6+	4,7	7,4

UF: Unidade da Federação.

* Entre os indivíduos de 18-69 anos.

Tabela 2

Estimativas dos indicadores utilizando a amostra sem ponderação, a amostra com ponderação e a amostra não considerando as duas primeiras ondas.

Indicador	Estimativa 1 *	Estimativa 2 **	Estimativa 3 ***
	%	%	%
Estilos de vida			
FLV	29,1	22,1	22,1
Feijão 5 dias ou mais	37,7	43,2	42,4
Atividade física \geq 150 minutos	34,2	35,51	35,2
TV/Tablet 3 horas ou mais	66,0	60,6	61,6
Beber pesado	7,1	6,5	6,7
Fumantes de cigarros	10,7	12,0	12,3
Limitação AVD	5,7	6,1	6,2
Prevalências de doenças crônicas não transmissíveis (dados autorreferidos)			
Hipertensão	18,5	19,3	18,6
Diabetes	7,0	7,2	7,1
Doença do coração	3,8	4,3	4,4
Depressão	16,6	14,9	14,9
DCNT risco COVID-19	33,3	33,9	33,3
Excesso de peso	54,0	56,3	55,7
Obesidade	19,8	21,4	21,2

Atividade física \geq 150 minutos: pratica de atividade física no lazer no nível recomendado; Beber pesado: consumo abusivo e frequente de bebida alcoólica, definido como o consumo de 8 ou mais doses de bebida alcoólica por semana para as mulheres e 15 ou mais doses para os homens; DCNT risco COVID-19: doenças crônicas não-transmissíveis de risco de agravamento de COVID-19; Feijão 5 dias ou mais: consumo de feijão por 5 dias ou mais por semana; FLV: consumo recomendado de frutas e hortaliças; Fumantes de cigarros: que são fumantes atuais de cigarros; Limitação AVD: limitação funcional que consiste em precisar de ajuda para a realização das atividades da vida diárias entre idosos; TV/Tablet 3 horas ou mais: assiste à TV ou tablet no tempo livre por 3 horas ou mais.

* Sem ponderação;

** Com ponderação;

*** Com ponderação, não considerando as duas primeiras ondas.

recomendado de frutas e legumes, os resultados das 3 pesquisas são muito próximos, variando de 22,1 a 22,9%. Já o percentual de consumo de feijão por 5 dias ou mais da ConVid (42,4%) foi bem menor que o da PNS 2013 (71,9%) e do Vigitel 2019 (59,7%).

Na comparação da atividade física no lazer no tempo recomendado por semana (\geq 150 minutos), o percentual da ConVid (35,2%) está no mesmo patamar que o do Vigitel 2019 (39%). A pequena diferença de 3,8 pontos percentuais pode ser justificada pelo método de cálculo do indicador com os dados da ConVid, correspondente à média entre a proporção mínima (30%) e a máxima (40,5%), muito próxima à do Vigitel 2019. Em relação ao tempo de TV/tablet por 3 horas ou mais, a proporção encontrada na ConVid, de 61,6%, foi similar à do Vigitel, 62,7% (Tabela 2).

Em relação ao hábito de beber, a prevalência do consumo abusivo e frequente de álcool (beber pesado) calculada pelos dados da ConVid, de 6,7%, foi similar à estimada com os dados da PNS 2013, de 6,1%.

No tocante à proporção de fumantes de cigarros, a estimativa da ConVid (12,3%) ficou entre a proporção encontrada no Vigitel 2019 (9,8%) e a verificada na PNS 2013 (14,7%).

A análise comparativa do indicador de limitação funcional entre as pessoas com 60 anos ou mais que necessitam de ajuda para a realização das atividades da vida diária (Tabela 2) mostrou percentuais muito semelhantes na ConVid (6,2%) e na PNS 2013, (6,1%).

Na Tabela 4, comparam-se as estimativas de indicadores de morbidade referida. As proporções de pessoas que referem diagnóstico de hipertensão são próximas nas três pesquisas, com valores de 18,6% (ConVid), 21,4% (PNS 2013) e 24,5% (Vigitel 2019). Quanto às proporções que referem diagnóstico de diabetes, os percentuais foram de 6,2% (PNS 2013), 7,1% (ConVid) e 7,4% (Vigitel 2019), consis-

Tabela 3

Comparação de indicadores de estilos de vida obtidos com os dados da *ConVid – Pesquisa de Comportamentos, Pesquisa Nacional de Saúde* de 2013 (PNS 2013) ²⁶ e *Vigilância de Fatores de Risco e Proteção para Doenças Crônicas por Inquérito Telefônico* de 2019 (Vigitel 2019) ²⁷.

Prevalências	ConVid		PNS 2013		Vigitel 2019	
	%	IC95%	%	IC95%	%	IC95%
FLV	22,1	21,6-22,6	22,8	22,1-23,5	22,9	22,1-23,7
Feijão 5 dias ou mais	42,4	41,8-43,0	71,9	71,2-71,6	59,7	58,8-60,6
Atividade física ≥ 150 minutos *	35,21	34,6-35,8	22,5	21,8-23,1	39,0	38,0-39,9
TV/Tablet 3 horas ou mais	61,6	61,0-62,2	-	-	62,7	61,8-63,6
Beber pesado	6,7	6,4-7,0	6,1	5,8-6,4	-	-
Fumantes de cigarros	12,3	11,9-12,7	14,5	14,5-15,0	9,8	9,2-10,5
Limitação AVD	6,2	5,9-6,5	6,1	5,4-6,8	-	-

Atividade física ≥ 150 minutos: prática de atividade física no lazer no nível recomendado; Beber pesado: consumo abusivo e frequente de bebida alcoólica, definido como o consumo de 8 ou mais doses de bebida alcoólica por semana para as mulheres e 15 ou mais doses para os homens; Feijão 5 dias ou mais: consumo de feijão por 5 dias ou mais por semana; FLV: consumo recomendado de frutas e hortaliças; Fumantes de cigarros: que são fumantes atuais de cigarros; IC95%: intervalo de 95% de confiança; Limitação AVD: limitação funcional que consiste em precisar de ajuda para a realização das atividades da vida diárias entre idosos; TV/Tablet 3 horas ou mais: assiste à TV ou *tablet* no tempo livre por 3 horas ou mais.

* Na ConVid o número de dias e o tempo praticado em atividade física foram variáveis categóricas e a estimativa foi feita pela média entre os percentuais mínimo (30,1%) e máximo (40,9%).

Tabela 4

Comparação das estimativas das prevalências de doenças crônicas não transmissíveis obtidas com os dados autorreferidos da *ConVid – Pesquisa de Comportamentos, Pesquisa Nacional de Saúde* de 2013 (PNS 2013) ²⁶ e *Vigilância de Fatores de Risco e Proteção para Doenças Crônicas por Inquérito Telefônico* de 2019 (Vigitel 2019) ²⁷.

Prevalências	ConVid		PNS 2013		Vigitel 2019	
	%	IC95%	%	IC95%	%	IC95%
Hipertensão	18,6	18,1-19,1	21,4	20,8-22,0	24,5	23,8-25,3
Diabetes	7,1	6,8-7,4	6,2	5,9-6,6	7,4	7,0-7,9
Doença do coração	4,4	4,2-4,6	4,2	3,9-4,5	-	-
Depressão	14,9	14,5-15,3	7,6	7,2-8,1	-	-
DCNT risco COVID-19	33,3	32,7-33,9	34,0	33,2-34,7	-	-
Excesso de peso	55,7	55,1-56,3	56,4	55,6-57,2	55,4	54,4-56,3
Obesidade	21,2	20,7-21,7	20,8	20,2-21,4	20,3	19,5-21,0

DCNT risco COVID-19: doenças crônicas não-transmissíveis de risco de agravamento de COVID-19; IC95%: intervalo de 95% de confiança.

tentes com as estimativas obtidas com os resultados de exames laboratoriais de hemoglobina glicada da PNS (2014-2015), que variaram de 6,6 a 9,4% utilizando diferentes critérios para a presença da doença ³⁶. A comparação das prevalências de diagnóstico de doença do coração da ConVid com a PNS 2013 mostra valores similares, de 4,4% e 4,2%, respectivamente. Em relação ao percentual de pessoas com doenças crônicas de risco de agravamento de COVID-19 (33,3%), a estimativa foi também muito próxima à calculada com os dados da PNS 2013 (34%) ³⁹. No tocante ao excesso de peso e obesidade, as prevalências das três pesquisas são semelhantes. A única diferença considerável em relação à PNS foi quanto ao relato de depressão com uma proporção superior na ConVid (14,9% vs. 7,6%).

O mapa apresentado na Figura 1 mostra a grande abrangência geográfica da ConVid. Em um mês de pesquisa, foram alcançadas todas as UF e 1.696 municípios com pelo menos um participante.

Discussão

A realização da *ConVid – Pesquisa de Comportamentos* possibilitou descrever as mudanças nos comportamentos dos brasileiros, nas condições de saúde e no acesso aos serviços de saúde, avaliar os transtornos psicológicos e os problemas socioeconômicos relacionados às medidas rígidas de restrição social impostas pelos governos estaduais e municipais após a intensa disseminação da COVID-19 no Brasil.

O trabalho usou um questionário virtual com os participantes recrutados por amostragem em cadeia, do tipo bola de neve virtual, o que permitiu obter uma amostra grande, de mais de 45 mil pessoas em um mês, sem necessidade de recursos externos. Os resultados foram também divulgados pelo *site* da pesquisa (<https://convid.fiocruz.br/>) e pela mídia, de maneira ágil, o que possibilitou orientar a população a manter os comportamentos saudáveis e a procurar os atendimentos de telemedicina para atenuar os transtornos psicológicos.

A comparação dos indicadores de comportamentos saudáveis e de morbidade referida com os resultados de pesquisas tradicionais, como a PNS 2013 e o *Vigitel* 2019, mostrou grande similaridade na maioria das estimativas, possibilitando dizer que há representação da população brasileira na amostra da *ConVid*. No conjunto dos indicadores de estilos de vida, as estimativas da *ConVid* referentes ao fumo e à prática regular de atividade física foram mais próximas às do *Vigitel* 2019 e indicam os avanços alcançados nestes comportamentos ao longo dos anos^{27,38}. Apenas o percentual de consumo regular de feijão mostrou valor dissonante, associado, provavelmente, ao maior nível de escolaridade

Figura 1

Municípios com pelo menos um respondente da *ConVid – Pesquisa de Comportamentos*. Brasil, 24 de abril a 24 de maio de 2020.



dos participantes da ConVid, sendo compatível com o percentual encontrado entre os indivíduos com curso superior no Vigitel 2019 ²⁷. No conjunto dos indicadores de morbidade referida, o único indicador discrepante foi a prevalência autorreferida de depressão quando comparada à estimativa da PNS 2013. Tendo em vista o grande percentual de pessoas que procurou atendimento por problema de saúde mental durante a pandemia, a proporção de pessoas com diagnóstico de depressão pode, de fato, ter aumentado.

Embora as pesquisas pela Internet tenham uma série de vantagens, problemas de diferentes ordens podem ocorrer na amostragem em redes. No que se refere à ConVid, as pessoas que não têm acesso à Internet têm probabilidade nula de seleção e foi utilizado o procedimento de bola de neve por redes virtuais, que é uma técnica não probabilística de amostragem em que os participantes são voluntários, e as probabilidades de seleção e a taxa de não-resposta não podem ser estimadas. Além disso, como a lógica de recrutamento não é gerenciada pelo pesquisador, as conexões entre os participantes não são conhecidas, não sendo possível considerar a dependência das observações na estimação da variância.

Para a obtenção de uma amostra representativa da população, foram realizadas ponderações calculadas por procedimentos de pós-estratificação com base na PNAD 2019. O problema da pós-estratificação é que podem existir fatores relacionados aos indicadores de interesse, mas não considerados no conjunto de variáveis usadas na estratificação devido às limitações no tamanho da amostra para a estratificação por várias variáveis ou à falta de alcance da amostra de determinados estratos. Em artigos recentes ^{39,40}, foram discutidas alternativas para a estimação das probabilidades de seleção em amostras não probabilísticas com base em uma amostra probabilística de referência ou censo demográfico. No caso do Brasil, onde a PNAD é realizada anualmente, a aplicação dessas metodologias é viável e de muito interesse por propiciar a realização de inferência estatística em pesquisas com amostras não probabilísticas realizadas pela Internet.

Na ConVid, para as variáveis demográficas (sexo, faixa etária e raça) foi obtida a diversidade necessária para ponderar os dados, de modo a obter a representatividade da população brasileira. Quanto ao alcance da distribuição geográfica, o trabalho alcançou todas as UF e cerca de 1.700 municípios, permitindo representar as macrorregiões e a localização do município de residência (capital/restante da UF). Devido ao fato de os participantes terem a tendência de convidar pares para participar da pesquisa com características semelhantes às suas, denominado de efeito de homofilia ⁴¹, a estratificação das sementes, como realizada na ConVid, foi um fator importante para alcançar a diversidade na amostra ³¹.

Entretanto, como o percentual de pessoas com o ensino médio incompleto na amostra (11,1%) foi muito menor que o da PNAD 2019 (48,7%), foram utilizadas apenas duas categorias do grau de escolaridade na pós-estratificação (nível superior ou não), para não atribuir ponderações muito altas a um pequeno grupo de indivíduos. Segundo os dados da PNAD Contínua 2018 ¹, entre as pessoas que não acessaram a Internet em 2018, 41,6% alegaram que não sabiam usar. As dificuldades em participar de pesquisas pela Internet afetam também a velocidade de desenvolvimento das redes, que se desenvolvem mais rapidamente entre os indivíduos que têm maior facilidade no preenchimento de questionário, e acabam sendo predominantes na amostra.

Em dados coletados por amostragem em cadeia, o efeito de homofilia contribui, igualmente, para a dependência das observações, sendo necessário usar procedimentos que levem em consideração as conexões entre os participantes para a estimação da variância e do efeito de desenho ⁴². Quando há o registro dos pares recrutador/recrutado, o procedimento de *bootstrap* tem sido recomendado para a estimação da variância dos indicadores de interesse por meio de várias simulações de amostras geradas com o mesmo processo que originou a amostra total ^{43,44}, mas não foi possível a sua utilização na ConVid porque não foram coletadas informações sobre a rede de contatos. Outra proposta de estimação da variância em amostras por redes consiste em desagregar os efeitos endógenos e contextuais por meio da comparação das covariâncias dos resultados dos pares com as covariâncias dos resultados dos pares dos pares ⁴⁵.

Estudos apontam para a maior diversidade das pessoas nas redes sociais, no que diz respeito à localização geográfica e às características socioeconômicas, do que nos conglomerados clássicos utilizados nas pesquisas tradicionais por amostragem de domicílio ^{3,10}. Sob a suposição de que o efeito de homofilia decresce na medida em que aumenta a “distância” do participante à semente ⁴⁵, a expansão da rede em várias ondas de recrutamento convergiria para a composição de uma amostra

grande e abrangente das características populacionais. No caso da ConVid, as estimativas realizadas com a amostra total sem ponderação mostram a influência do grau de escolaridade nas estimativas dos indicadores de comportamento. Prováveis explicações são a exclusão social de acesso à Internet e o desenvolvimento mais rápido da rede entre os indivíduos de maior escolaridade.

Embora o uso das pesquisas *online* seja cada vez mais frequente para estimar a prevalência de eventos de interesse para a saúde pública, raramente são realizadas inferências sobre as propriedades estruturais das redes sociais ⁴⁶. Para corrigir o viés de seleção dos processos de recrutamento em cadeia, Goel & Salganick ³⁵ propuseram excluir os participantes das primeiras ondas da rede. As estimativas dos indicadores de estilos de vida e de DCNT, sem considerar as duas primeiras ondas, mostraram-se significativas e muito próximas às calculadas com a amostra total ponderada, de forma que esta pode ser uma opção mais adequada para a utilização em pesquisas em redes virtuais e foi adotada no presente trabalho.

As pesquisas *online* são muito promissoras para a área da saúde, principalmente no que se refere à aquisição de conhecimento de maneira ágil e imediata, refletindo a situação das condições de saúde em tempo real. Para imprimir mais qualidade às pesquisas realizadas pela Internet foi proposta uma lista de recomendações que vão desde a identificação de pessoas que preenchem o questionário eletrônico mais de uma vez, a elaboração do questionário, até o cálculo da taxa de participação ⁴⁷. Entretanto, existe ainda a preocupação em saber se o grupo de pessoas alcançadas na amostra por redes representa a população geral. No tocante à ConVid, a comparação da distribuição por grau de escolaridade na amostra com a obtida na PNAD 2019 mostrou a divisão social no acesso à Internet no Brasil, considerada como um dos marcadores mais importantes das desigualdades socioeconômicas ⁴⁸. Apesar da lacuna de pessoas de baixa escolaridade, as estimativas de vários indicadores de saúde foram compatíveis, na sua maioria, aos valores encontrados na PNS 2013 e no Vigitel 2019. No que concerne à estimação da variância, são necessários estudos para investigar como os efeitos endógenos das redes sociais virtuais podem ser levados em consideração, para que a inferência estatística possa ser usada de maneira adequada e com mais fidedignidade.

Colaboradores

C. L. Szwarcwald, P. R. B. Souza Júnior e G. N. Damacena participaram da concepção e redação do artigo, análise e interpretação dos dados, aprovação da versão final e são responsáveis por todos os aspectos do trabalho na garantia da exatidão e integridade de qualquer parte da obra. D. C. Malta, M. B. A. Barros, D. E. Romero, L. O. Azevedo colaboraram na concepção e redação do artigo, aprovação da versão final e são responsáveis por todos os aspectos do trabalho na garantia da exatidão e integridade de qualquer parte da obra. W. S. Almeida contribuiu na análise e interpretação dos dados, aprovação da versão final do artigo e é responsável por todos os aspectos do trabalho na garantia da exatidão e integridade de qualquer parte da obra. I. E. Machado, M. G. Lima, A. O. Werneck, D. R. P. Silva, C. S. Gomes, A. P. S. Ferreira, R. Gracie e M. F. Pina colaboraram na revisão crítica relevante do conteúdo intelectual, aprovação da versão final do artigo e são responsáveis por todos os aspectos do trabalho na garantia da exatidão e integridade de qualquer parte da obra.

Informações adicionais

ORCID: Celia Landmann Szwarcwald (0000-0002-7798-2095); Paulo Roberto Borges de Souza Júnior (0000-0002-8142-4790); Giseli Nogueira Damacena (0000-0002-7059-3353); Deborah Carvalho Malta (0000-0002-8214-5734); Marilisa Berti de Azevedo Barros (0000-0003-3974-195X); Dalia Elena Romero (0000-0002-2643-9797); Wanessa da Silva de Almeida (0000-0002-5164-8603); Luiz Otávio Azevedo (0000-0002-4876-5948); Ísis Eloah Machado (0000-0002-4678-2074); Margareth Guimarães Lima (0000-0001-6996-0745); André Oliveira Werneck (0000-0002-9166-4376); Danilo Rodrigues Pereira da Silva (0000-0003-3995-4795); Crizian Saar Gomes (0000-0001-6586-4561); Arthur Pate de Souza Ferreira (0000-0002-6686-0105); Renata Gracie (0000-0003-0225-3696); Maria de Fátima de Pina (0000-0002-1521-7865).

Agradecimentos

Agradecemos aos pesquisadores Ana Maria de Brito, Angélica Espinosa Miranda, Paulo Germano de Frias e Rita Suely Bacuri de Queiroz, pelo apoio na divulgação da pesquisa e disseminação da rede.

Referências

1. Instituto Brasileiro de Geografia e Estatística. Pesquisa Nacional por Amostra de Domicílios Contínua: acesso à Internet e à televisão e posse de telefone móvel celular para uso pessoal 2018. PNAD Contínua. https://agenciadendoticias.ibge.gov.br/media/com_mediaibge/arquivos/10d5c0576ff8d726467f1d4571dd8e62.pdf (acessado em 01/Set/2020).
2. Evans RS. Electronic health records: then, now, and in the future. *Yearb Med Inform* 2016; (Suppl 1):S48-61.
3. Fenner Y, Garland SM, Moore EE, Jayasinghe Y, Fletcher A, Tabrizi SN, et al. Web-based recruiting for health research using a social networking site: an exploratory study. *J Med Internet Res* 2012; 14:e20.
4. De Cocker K, Chastin SFM, De Bourdeaudhuij I, Imbo I, Stragier J, Cardon G. Citizen science to communicate about public health messages: the reach of a playful online survey on sitting time and physical activity. *Health Commun* 2019; 34:720-5.
5. Milgram S. The small-world problem. *Psychol Today* 1967; 2:60-7.
6. Faleiros F, K appler C, Pontes FAR, Silva SSC, Goes FSN, Cucik CD. Uso de question ario online e divulga o virtual como estrat gia de coleta de dados em estudos cient ficos. *Texto Contexto Enferm* 2016; 25:e3880014.
7. Ekman A, Litton JE. New times, new needs; e-epidemiology. *Eur J Epidemiol* 2007; 22:285-92.
8. van Gelder MM, Bretveld RW, Roeleveld N. Web-based questionnaires: the future in epidemiology? *Am J Epidemiol* 2010; 172:1292-8.
9. Nolte MT, Shauver MJ, Chung KC. Analysis of four recruitment methods for obtaining normative data through a web-based questionnaire: a pilot study. *Hand (N Y)* 2015; 10:529-34.
10. Costa BRL. Bola de neve virtual: o uso das redes sociais virtuais no processo de coleta de dados de uma pesquisa cient fica. *Revista Interdisciplinar de Gest o Social* 2018; 7:15-37.
11. Aglipay M, Wylie JL, Jolly AM. Health research among hard-to-reach people: six degrees of sampling. *CMAJ* 2015; 187:1145-9.
12. Magnani R, Sabin K, Saidel T, Heckathorn D. Review of sampling hard-to-reach and hidden populations for HIV surveillance. *AIDS* 2005; 19 Suppl 2:S67-72.
13. Faugier J, Sargeant M. Sampling hard to reach populations. *J Adv Nurs* 1997; 26:790-7.
14. Ali SH, Foreman J, Capasso A, Jones AM, Tozan Y, Di Clemente RJ. Social media as a recruitment platform for a nationwide online survey of COVID-19 knowledge, beliefs, and practices in the United States: methodology and feasibility analysis. *BMC Med Res Methodol* 2020; 20:116.

15. Gharpure R, Hunter CM, Schnall AH, Barrett CE, Kirby AE, Kunz J, et al. Knowledge and practices regarding safe household cleaning and disinfection for COVID-19 prevention – United States, May 2020. *Morb Mortal Wkly Rep* 2020; 69:705-9.
16. Reuben RC, Danladi MMA, Saleh DA, Ejembi PE. Knowledge, attitudes and practices towards COVID-19: an epidemiological survey in north-central Nigeria. *J Community Health* 2020; [Epub ahead of print].
17. Machida M, Nakamura I, Saito R, Nakaya T, Hanibuchi T, Takamiya T, et al. Adoption of personal protective measures by ordinary citizens during the COVID-19 outbreak in Japan. *Int J Infect Dis* 2020; 94:139-44.
18. Denis F, Galmiche S, Dinh A, Fontanet A, Scherpereel A, Benezit F, et al. Epidemiological observations on the association between anosmia and COVID-19 infection: analysis of data from a self-assessment web application. *J Med Internet Res* 2020; 22:e19855.
19. Guo Y, Cheng C, Zeng Y, Li Y, Zhu M, Yang W, et al. Mental health disorders and associated risk factors in quarantined adults during the COVID-19 outbreak in China: a cross-sectional study. *J Med Internet Res* 2020; 22:e20328.
20. Vordenberg SE, Zikmund-Fisher BJ. Older adults' strategies for obtaining medication refills in hypothetical scenarios in the face of COVID-19 risk. *J Am Pharm Assoc* (2003) 2020; 60:915-22.
21. Steele EM, Rauber F, Costa CS, Leite MA, Gabe KT, Louzada MLC, et al. Dietary changes in the NutriNet Brasil cohort during the COVID-19 pandemic. *Rev Saúde Pública* 2020; 54:91.
22. Bommele J, Hopman P, Walters BH, Geboers C, Croes E, Fong GT, et al. The double-edged relationship between COVID-19 stress and smoking: implications for smoking cessation. *Tob Induc Dis* 2020; 18:63.
23. Mukhtar S. Psychological health during the coronavirus disease 2019 pandemic outbreak. *Int J Soc Psychiatry* 2020; 66:512-6.
24. Qiu Y, Chen X, Shi W. Impacts of social and economic factors on the transmission of coronavirus disease 2019 (COVID-19) in China. *J Popul Econ* 2020; [Online ahead of print].
25. Szwarcwald CL, Viacava F. *Pesquisa Mundial de Saúde no Brasil*, 2003. *Cad Saúde Pública* 2005; 21 Suppl 1:S4-5.
26. Szwarcwald CL, Malta DC, Pereira CA, Vieira ML, Conde WL, Souza Júnior PR, et al. *Pesquisa Nacional de Saúde no Brasil: concepção e metodologia de aplicação*. *Ciênc Saúde Colet* 2014; 19:333-42.
27. Departamento de Análise em Saúde e Vigilância de Doenças Não Transmissíveis, Secretaria de Vigilância em Saúde, Ministério da Saúde. *Vigitel Brasil 2019. Vigilância de fatores de risco e proteção para doenças crônicas por inquérito telefônico: estimativas sobre frequência e distribuição sociodemográfica de fatores de risco e proteção para doenças crônicas nas capitais dos 26 estados brasileiros e no Distrito Federal em 2019*. Brasília: Ministério da Saúde; 2020.
28. Barros MBA, Lima MG, Ceolin MF, Zancanello E, Cardoso TAMO. Quality of sleep, health status and subjective well-being in a population-based study. *Rev Saúde Pública* 2019; 53:82.
29. Pereira CA, Malta DC, Szwarcwald CL. *Pesquisa Nacional de Saúde 2013: ciclos de vida*. Rio de Janeiro: Instituto Brasileiro de Geografia e Estatística; 2015.
30. Souza Júnior PR, Szwarcwald CL, Castilho EA. Self-rated health by HIV infected individuals undergoing antiretroviral therapy in Brazil. *Cad Saúde Pública* 2011; 27 Suppl 1:S56-66.
31. McKnight C, Des Jarlais D, Bramson H, Tower L, Abdul-Quader AS, Nemeth C, et al. Respondent-driven sampling in a study of drug users in New York City: notes from the field. *J Urban Health* 2006; 83 Suppl 6:i54-9.
32. Szwarcwald CL, Damacena GN. Amostras complexas em inquéritos: planejamento e implicações na análise estatística de dados. *Rev Bras Epidemiol* 2008; 11 Suppl 1:38-45.
33. Instituto Brasileiro de Geografia e Estatística. *Pesquisa Nacional por Amostra de Domicílios (PNAD) 2019. Notas técnicas. Versão 1.5. 4ª Ed.* Rio de Janeiro: Instituto Brasileiro de Geografia e Estatística; 2019.
34. Pereira CA, Malta DC, Szwarcwald CL. *Pesquisa Nacional de Saúde 2013: percepção do estado de saúde, estilos de vida e doenças crônicas*. Rio de Janeiro: Instituto Brasileiro de Geografia e Estatística; 2014.
35. Goel S, Salganik MJ. Respondent-driven sampling as Markov chain Monte Carlo. *Stat Med* 2009; 28:2202-29.
36. Malta DC, Duncan BB, Schmidt MI, Machado IE, Silva AGD, Bernal RTI, et al. Prevalence of diabetes mellitus as determined by glycated hemoglobin in the Brazilian adult population, National Health Survey. *Rev Bras Epidemiol* 2019; 22 Suppl 2:E190006.
37. Rezende LFM, Thome B, Schweitzer MC, Souza Júnior PRB, Szwarcwald CL. Adults at high-risk of severe coronavirus disease-2019 (COVID-19) in Brazil. *Rev Saúde Pública* 2020; 54:50.

38. Pinto M, Bardach A, Palacios A, Biz A, Alcaraz A, Rodriguez B, et al. Burden of smoking in Brazil and potential benefit of increasing taxes on cigarettes for the economy and for reducing morbidity and mortality. *Cad Saúde Pública* 2019; 35:e00129118.
39. Elliott MR, Valliant R. Inference for nonprobability samples. *Stat Sci* 2017; 32:249-64.
40. Valliant R. Comparing alternatives for estimation from nonprobability samples. *J Surv Stat Methodol* 2020; 8:231-63.
41. Salganik MJ, Heckathorn DD. Sampling and estimation in hidden populations using respondent-driven sampling. *Sociol Methodol* 2004; 34:193-240.
42. Szwarcwald CL, Souza Júnior PR, Damacena GN, Barbosa Junior A, Kendall C. Analysis of data collected by RDS among sex workers in 10 Brazilian cities, 2009: estimation of the prevalence of HIV, variance, and design effect. *J Acquir Immune Defic Syndr* 2011; 57 Suppl 3:S129-35.
43. Spiller MW, Gile KJ, Handcock MS, Mar CM, Wejnert C. Evaluating variance estimators for respondent-driven sampling. *J Surv Stat Methodol* 2017; 2017:smx018.
44. Baraff AJ, McCormick TH, Raftery AE. Estimating uncertainty in respondent-driven sampling using a tree bootstrap method. *Proc Natl Acad Sci U S A* 2016; 113:14668-73.
45. Rose CD. Identification of peer effects through social networks using variance restrictions. *Econom J* 2017; 20:S47-60.
46. Verdery AM, Fisher JC, Siripong N, Abdesselem K, Bauldry S. New survey questions and estimators for network clustering with respondent-driven sampling data. *Sociol Methodol* 2017; 47:274-306.
47. Eysenbach G. Improving the quality of web surveys: the checklist for reporting results of internet e-surveys. *J Med Internet Res* 2004; 6:e34.
48. Feehan DM, Cobb C. Using an online sample to estimate the size of an offline population. *Demography* 2019; 56:2377-92.

Abstract

The ConVid – Behavior Survey was conducted in Brazil from April 24 to May 24, 2020, aiming to investigate changes in lifestyles and health conditions during the COVID-19 pandemic. In this article, we present the conception and methodology of the research. We used a cross-sectional study using an internet questionnaire, with questions validated in previous health surveys. The sampling method “virtual snowball” was used, as well as post-stratification procedures. The results related to chronic non-communicable diseases and pre-pandemic lifestyles were compared with estimates from the 2013 Brazilian National Health Survey and 2019 Surveillance of Risk and Protective Factors for Chronic Diseases by Telephone Survey. The total sample was 45,161 people. After data weighing, the sample distributions of demographic variables were similar to population variables. Only people with a low schooling level were underrepresented. The comparison with the previous results showed similarity in most estimates: recommended consumption of fruits and vegetables (22.1%), recommended physical activity (35.2%), tobacco smoking habit (12.3%), frequent and abusive alcohol consumption (6.7%), obesity (21.2%), self-reported prevalence of hypertension (18.6%), diabetes (7.1%), and heart disease (4.4%). The online survey made it possible to know the population’s health conditions during the pandemic. The similarity of the indicators with those obtained in traditional research allowed the validation of the mean estimates. Studies are needed to investigate how the endogenous effects of virtual social networks can be considered when estimating variance.

COVID-19; Internet; Health Surveys

Resumen

La ConVid – Encuesta de Comportamientos se realizó en Brasil del 24 de abril al 24 de mayo de 2020, con el objetivo de investigar los cambios en los estilos de vida y en las condiciones de salud durante la pandemia de COVID-19. En este artículo, se presenta la concepción y metodología de la encuesta. Estudio de corte transversal, utilizando un cuestionario por Internet, con cuestiones validadas en encuestas de salud anteriores. El método de muestra fue el de “bola de nieve virtual” y se usaron procedimientos de posestratificación. Los resultados relativos a las enfermedades crónicas no transmisibles y estilos de vida prepandemia se compararon con las estimaciones de la Encuesta Nacional de Salud de 2013 y de la Vigilancia de Factores de Riesgo y Protección para Enfermedades Crónicas por Encuesta Telefónica de 2019. La muestra total fue de 45.161 personas. Tras la ponderación de los datos, las distribuciones de la muestra de las variables demográficas fueron semejantes a las poblacionales. Solamente las personas con un bajo nivel de formación estuvieron sub-representadas. La comparación con los resultados anteriores mostró similitud en la mayoría de las estimaciones: consumo recomendado de frutas y legumbres (22,1%), actividad física recomendada (35,2%), tabaco de liar para cigarrillos (12,3%), consumo frecuente y abusivo de alcohol (6,7%), obesidad (21,2%), prevalencias autoinformadas de hipertensión (18,6%), diabetes (7,1%) y enfermedad del corazón (4,4%). La encuesta online permitió conocer las condiciones de salud de la población durante la pandemia. La similitud de los indicadores con los obtenidos en investigaciones tradicionales permitió validar las estimaciones medias. Se necesitan estudios para investigar cómo los efectos endógenos de las redes sociales virtuales pueden ser tenidos en consideración en la estimación de la variancia.

COVID-19; Internet; Encuestas Epidemiológicas

Recebido em 04/Mar/2020

Versão final reapresentada em 18/Mai/2020

Aprovado em 19/Jun/2020