

Validade e reprodutibilidade de inquérito telefônico de atividade física no Brasil

Validity and reliability of a telephone survey of physical activity in Brazil

Alexandra Dias Moreira¹, Rafael Moreira Claro¹,
Mariana Santos Felisbino-Mendes¹, Gustavo Velasquez-Melendez¹

RESUMO: *Introdução:* A atividade física no Brasil é avaliada pelo Sistema de Vigilância de Fatores de Risco e Proteção para Doenças Crônicas por Inquérito Telefônico (VIGITEL). *Objetivo:* Avaliar a reprodutibilidade e a validade dos indicadores de atividade física do VIGITEL. *Métodos:* Para o estudo de reprodutibilidade, 305 indivíduos em Belo Horizonte responderam a entrevistas do VIGITEL repetidas com intervalos de 7 a 15 dias, em 2013. Os indicadores avaliados foram “suficientemente ativos no lazer”, “ativos no deslocamento”, “inativos em quatro domínios da atividade física” (lazer, trabalho, transporte e atividades domésticas) e “assistir TV por longos períodos”. O Coeficiente Kappa (k) foi utilizado para medir concordância entre as entrevistas. Para a validade, os participantes responderam também o Questionário Global de Atividade Física (GPAQ), método de referência para a comparação dos indicadores VIGITEL. A comparação foi avaliada por sensibilidade, especificidade, valores preditivos positivos (VPP) e negativos (VPN). *Resultados:* A reprodutibilidade mostrou concordância substancial para os indivíduos ativos no lazer (k = 0,70) e inativos (k = 0,64). A concordância do hábito de assistir TV foi moderada (k = 0,56) e o deslocamento mostrou concordância regular (k = 0,35). No estudo de validade, a sensibilidade variou de 54,8 a 67,7 na frequência de inativos e ativos no lazer, respectivamente. O deslocamento foi representado por 11,9 de sensibilidade. A especificidade variou de 72,0 a 91,2 nos quatro domínios da atividade física e inativos. *Conclusão:* O questionário de atividade física utilizado pelo Sistema de Vigilância parece ser confiável em todos os domínios, exceto nas questões de deslocamento. O VIGITEL foi comparável ao GPAQ na maioria dos aspectos da atividade física.

Palavras-chave: Atividade motora. Reprodutibilidade dos testes. Validade dos testes. Questionários. Eficiência. Epidemiologia.

¹Escola de Enfermagem do Departamento Materno Infantil e Saúde Pública da Universidade Federal de Minas Gerais – Belo Horizonte (MG), Brasil.

²Escola de Enfermagem do Departamento de Nutrição da Universidade Federal de Minas Gerais – Belo Horizonte (MG), Brasil.
Autor correspondente: Gustavo Velasquez-Melendez. Escola de Enfermagem do Departamento de Enfermagem Materno-Infantil e Saúde Pública da Universidade Federal de Minas Gerais. Avenida Alfredo Balena, 190, Santa Efigênia. CEP: 30130-100. Belo Horizonte, MG, Brasil. E-mail: jguveme@ufmg.br

Conflito de interesses: nada a declarar – **Fonte de financiamento:** Ministério da Saúde (Convênio 188).

ABSTRACT: Introduction: Physical activity in Brazil is assessed by the Telephone-based Risk Factor Surveillance System for Chronic Diseases (VIGITEL). **Objectives:** To evaluate the reliability and concurrent validity of VIGITEL physical activity indicators. **Methods:** For the reliability study, 305 individuals responded to VIGITEL interviews, which were repeated with intervals of 7 – 15 days, in Belo Horizonte, 2013. The evaluated indicators included “sufficiently active on leisure time,” “active in transportation,” “inactive in four domains of physical activity (leisure, work, transportation, and housework),” and “watching TV for long periods.” Kappa coefficient (k) was used to measure agreement between both interviews. For concurrent validity assessment, the same subjects also responded to the Global Physical Activity Questionnaire (GPAQ), used as reference method for comparison of VIGITEL indicators. Comparison was assessed by measures of sensitivity, specificity, positive (PPV), and negative predictive values (NPV). **Results:** Reliability study showed substantial agreement for active individuals in leisure time ($k = 0.70$) and inactive individuals ($k = 0.64$). The agreement of watching TV for long periods was moderate ($k = 0.56$) and the activeness in transportation showed fair agreement ($k = 0.35$). In comparison to the reference method, sensitivity ranged from 54.8 to 67.7 in frequency of inactive, active in leisure time and work. Transportation domain was represented by 11.9 of sensitivity. Specificity ranged from 72.0 to 91.2 among four domains of physical activity and inactive. **Conclusion:** Physical activity questionnaire used by the surveillance system seems to be reliable in all domains, except in transportation questions. VIGITEL was considered comparable to GPAQ in most aspects of physical activity evaluation.

Keywords: Motor activity. Reproducibility of results. Validity of tests. Questionnaires. Efficiency. Epidemiology.

INTRODUÇÃO

A inatividade física destaca-se como um dos principais fatores de risco modificáveis para ocorrência de doenças e agravos não transmissíveis (DANT)¹. No entanto, frequências elevadas de prática insuficiente de atividade física são verificadas tanto nos países desenvolvidos quanto naqueles em desenvolvimento². No Brasil, sabe-se que aproximadamente 1 a cada 2 adultos (49,4%) não atinge o volume recomendado de prática de atividade física (AF)².

A Organização Mundial de Saúde (OMS) define como saudável a prática semanal de, ao menos, 150 minutos de atividade de intensidade moderada ou equivalente (75 minutos de atividades intensas ou, ainda, combinação equivalente de ambas)³. A aferição do volume de AF praticado por indivíduos ou populações é realizada, frequentemente, por meio do uso de questionários com respostas autorreferidas^{4,5}. Nesse sentido, a OMS desenvolveu, em 2002, o Questionário Global de Atividade Física (GPAQ), com objetivo de permitir uma ampla aferição da AF (em diversos domínios), a ser utilizada internacionalmente^{6,7}. Seu desenvolvimento levou em consideração diferenças culturais, assim como subpopulações, considerando, por exemplo, pessoas com baixo nível socioeconômico e indivíduos analfabetos. Esse instrumento foi traduzido para diversos idiomas e apresentou boa confiabilidade em diversos estudos realizados até o presente momento^{5,7}. Esse instrumento encontra-se validado para população adulta brasileira⁵ e seu uso faz-se presente em diversos estudos em

nível internacional⁶, frequentemente aplicado com mediação presencial de um entrevistador, ou, ainda, por meio de entrevistas por telefone⁸.

No Brasil, dados oficiais e sistemáticos de vigilância da prática de AF estão disponíveis a partir do ano de 2006, coletados pelo Sistema de Vigilância de Fatores de Risco e Proteção para Doenças Crônicas por Inquérito Telefônico (VIGITEL), baseado em entrevistas não presenciais para aferição de dados autorreferidos. Entretanto, a natureza do VIGITEL — aplicação por telefone e com questionário envolvendo todos os principais fatores de risco e proteção para DANT — impede a utilização integral de instrumentos como o GPAQ, dada a necessidade do uso de questões simples e curtas com respostas objetivas e diretas⁹. Desse modo, a aferição da prática de AF no VIGITEL é feita por meio de um conjunto de questões adaptadas de inquéritos já existentes há época de sua concepção (como o *Behavioral Risk Factor Surveillance System* do *Center for Disease Control* – BRFSS/CDC) e questionários mundialmente utilizados na área (como o GPAQ e o Questionário Internacional de Atividade Física – IPAQ). Isso criou a necessidade de estudos de validação desse questionário e análises de sua reprodutibilidade. Em estudo realizado na cidade de São Paulo, no ano de 2005, a validade e reprodutibilidade dos indicadores de AF do VIGITEL foram investigadas tendo como padrão de referência as informações obtidas em três recordatórios de 24 horas (R24H)¹⁰. Nessa ocasião, resultados satisfatórios foram obtidos tanto para validade quanto para reprodutibilidade dos indicadores. No entanto, posteriormente, foi verificado que são necessários pelo menos quatro dias de aplicação do R24H para estimar uma semana de atividades físicas (três dias durante a semana e um dia no final de semana)¹¹. Além disso, após esses achados, pouco se avançou nesse sentido, de modo que questões importantes permanecem desconhecidas, como a comparabilidade do VIGITEL com questionários de uso internacional.

Portanto, pretende-se, com este estudo, verificar a reprodutibilidade e validade dos indicadores de AF do VIGITEL (com métodos mais avançados e de forma mais acurada).

MÉTODOS

POPULAÇÃO DE ESTUDO E COLETA DOS DADOS

Trata-se de um estudo epidemiológico transversal com abordagem diagnóstica, realizado na cidade de Belo Horizonte, Minas Gerais. Para o cálculo amostral, o valor de kappa esperado para o estudo foi de 0,9, 50% de proporção positiva esperada para primeira e segunda observação, precisão absoluta de 5% e nível de confiança de 95%. Dessa forma, com a ajuda do *software* Stata 9.1, foi definida uma amostra de 300 indivíduos. Inicialmente, uma subamostra aleatória dos indivíduos estudados pelo VIGITEL, no ano de 2013, na cidade de Belo Horizonte, entre março e abril (réplicas amostrais 1, 2, 3 e 4 — maiores detalhes sobre o procedimento de amostragem do VIGITEL encontram-se publicados no relatório anual do sistema⁹) foi definida e estabelecida como contato inicial (linha de base) junto aos

participantes no estudo. Assim, foram incluídos 418 indivíduos de forma a permitir que a amostra mínima pretendida fosse atingida até mesmo com recusas totais próximo de 30%.

Um número total de 305 pessoas aceitou participar do estudo e foram incluídas nas etapas subsequentes de coleta de dados, para o estudo de reprodutibilidade e validade. Para isso, os participantes eram contatados novamente, por telefone, de 7 a 15 dias após terem respondido a entrevista original do VIGITEL, sendo solicitados a responder novamente ao bloco de questões sobre AF deste inquérito (estudo de reprodutibilidade) e ao questionário GPAQ, versão 2, método considerado de referência para o estudo de validade. As entrevistas para o presente estudo foram garantidas de tal forma que o segundo entrevistador fosse sempre diferente daquele que havia conduzido a entrevista original.

O GPAQ foi aplicado conforme recomendações do *Global Physical Activity Questionnaire Analyses Guide*, proposto pela OMS¹². Simulações de entrevistas foram realizadas, a fim de que possíveis erros fossem identificados antes do início do estudo. Além disso, durante o período de coleta de dados, foi feito um controle periódico de qualidade, por meio de gravações de entrevistas escolhidas aleatoriamente.

ORGANIZAÇÃO DOS DADOS E CONFEÇÃO DOS INDICADORES DE INTERESSE

O questionário do VIGITEL inclui informações sobre atividades físicas realizadas em quatro domínios: no tempo livre, no trabalho, no deslocamento e nas atividades domésticas. Para o estudo de reprodutibilidade, todos os indicadores do VIGITEL foram utilizados para a comparação entre a entrevista repetida e original do sistema⁹.

Foram considerados suficientemente ativos no tempo livre os indivíduos que praticavam pelo menos 150 minutos semanais de AF de intensidade moderada ou pelo menos 75 minutos semanais de AF intensa. Foram classificadas como atividades de intensidade moderada: caminhada, caminhada em esteira, musculação, hidroginástica, ginástica em geral, natação, artes marciais, ciclismo e voleibol. As AFs consideradas intensas foram: corrida, corrida em esteira, ginástica aeróbica, futebol, basquetebol e tênis⁹. Vale ressaltar que no VIGITEL não existem perguntas sobre a percepção individual da intensidade da atividade física, de acordo com as modalidades consideradas⁹.

Foram considerados ativos no deslocamento os indivíduos que geralmente se deslocavam para o trabalho ou curso/escola de bicicleta ou caminhando e que despendessem 30 minutos ou mais diários no percurso de ida e volta⁹.

Foram considerados fisicamente inativos os indivíduos que não praticaram nenhuma atividade física no tempo livre nos últimos três meses e que não realizaram esforços físicos intensos no trabalho, não se deslocaram para o trabalho ou curso/escola caminhando ou de bicicleta, perfazendo um mínimo de dez minutos por trajeto por dia e não foram responsáveis pela limpeza pesada de suas casas⁹.

Adicionalmente, foi apresentada a frequência de adultos que têm o hábito de assistir televisão por três horas ou mais por dia. Esse indicador vincula o hábito de ver televisão

ao comportamento sedentário, que, por sua vez, está relacionado ao risco de desenvolvimento de patologias crônicas, como obesidade, diabetes tipo II, doenças cardiovasculares e síndrome metabólica, independente da prática recomendada de atividade física⁹.

O GPAQ inclui 16 perguntas com informações sobre atividades físicas realizadas em 3 oportunidades ou domínios: no tempo livre (prática de esporte, exercício recreativo), no deslocamento (caminhada ou uso de bicicleta) e no trabalho (trabalho remunerado ou voluntário, atividades domésticas), além de comportamento sedentário (tempo sentado lendo, assistindo TV, conversando com amigos, viajando). As AF são avaliadas segundo frequência e duração na última semana e classificadas como moderadas ou vigorosas. Essas questões permitiram elaborar três dos quatro indicadores de AF utilizados pelo sistema VIGITEL, considerando a duração, a frequência e a intensidade das atividades: “atividade física no tempo livre” (suficientemente ativos são indivíduos que praticam pelo menos 150 minutos semanais de AF de intensidade moderada ou pelo menos 75 minutos semanais de AF intensa), “atividade física no deslocamento” (suficientemente ativos são indivíduos que se deslocam de um lugar para o outro de bicicleta ou caminhando e que despendem 30 minutos ou mais diários no percurso de ida e volta) e “inatividade” (inativos são indivíduos que não praticaram nenhuma atividade física no tempo livre e que não realizaram esforços físicos no trabalho, não se deslocaram de um lugar para outro caminhando ou de bicicleta). O indicador “hábito de assistir à televisão” não foi construído, pois a pergunta relativa a comportamento sedentário no GPAQ está relacionada a vários momentos em que o indivíduo permanece sentado durante a semana, não somente ao hábito de assistir televisão.

ANÁLISE DOS DADOS

Para o estudo de reprodutibilidade, as diferenças entre as proporções de ativos no lazer, deslocamento, inativos e indivíduos com hábito de assistir TV por 3 horas ou mais nas entrevistas original e repetida foram testadas por meio do teste McNemar, considerando significância < 5%.

A concordância entre os resultados da entrevista inicial (VIGITEL) e da entrevista repetida foi investigada por meio de teste Kappa. Foram considerados como valores de referência: 0,80 como concordância quase perfeita; entre 0,61 e 0,80 concordância substancial; entre 0,41 e 0,60 concordância moderada; entre 0,21 e 0,40 concordância regular; e abaixo de 0,21 concordância leve¹³. Por meio desse procedimento, analisou-se a concordância entre todos os quatro indicadores de AF presentes no VIGITEL.

Em seguida, o método *Bland and Altman* foi utilizado para investigar a relação entre os erros de medida e os valores reais. Como o valor real não é conhecido, a média entre as aferições foi assumida como a melhor estimativa da realidade¹⁴. Foi considerado, nessa análise, o total de minutos gastos em uma semana em AF no tempo livre e o gráfico foi construído a partir da diferença entre entrevista repetida e original em minutos.

O GPAQ foi utilizado como método de referência para a validação concorrente do bloco de perguntas sobre AF do sistema VIGITEL. Os indicadores de AF no tempo livre,

deslocamento e inatividade, elaborados a partir da segunda entrevista do VIGITEL, foram comparados àqueles indicadores obtidos por meio da entrevista com o GPAQ. As análises de comparação entre os questionários, utilizando como referência o GPAQ, foram realizadas por meio de cálculo de sensibilidade (S), especificidade (E), valor preditivo positivo (VPP) e valor preditivo negativo (VPN), além da área sob a curva ROC (*Receiver Operating Characteristic*).

Os cálculos obtidos por meio do GPAQ estiveram de acordo com o guia de análises proposto pela OMS, o *Global Physical Activity Questionnaire Analyses Guide*¹², sendo excluídos 3 indivíduos que apresentaram um período maior que 16 horas por dia de atividade física. A amostra final para o estudo de validade foi, portanto, de 302 indivíduos.

O presente estudo foi apreciado e aprovado pelo Comitê de Ética em Pesquisa da Universidade Federal de Minas Gerais, parecer nº 325.242 (CAAE: 15689713.3.00005149) e não há conflito de interesse. O consentimento livre e esclarecido foi obtido na ocasião do contato telefônico, assim como é feito nas entrevistas originais do VIGITEL. Foi garantido sigilo aos participantes sobre as informações coletadas e oferecida liberdade de desistência em qualquer etapa da pesquisa.

RESULTADOS

A amostra obtida foi de 305 indivíduos para o estudo de reprodutibilidade. Dos participantes, 58% eram mulheres e 42% homens. A média de idade para ambos os estudos foi de 49,7 anos (DP = 18,2) e cerca da metade dos entrevistados referiram ser casados legalmente. Para o estudo de validade, a amostra final foi constituída de 302 indivíduos, com características similares àqueles participantes do estudo de reprodutibilidade.

Em relação à comparação entre entrevista original e repetida do VIGITEL, para o indicador “AF no tempo livre” se obteve frequência de ativos de 37,0 e 32,8%, respectivamente. O coeficiente Kappa (k) encontrado foi de 0,70, indicando concordância substancial. Com relação ao indicador “AF no deslocamento”, foram encontradas frequências de ativos de 11,5 e 10,5% para entrevista original e repetida, respectivamente, com um k de 0,35, indicando concordância regular. Além disso, 14,7% de participantes da amostra foram considerados inativos fisicamente na entrevista original e 16,4% na entrevista repetida, com concordância substancial para o indicador de inatividade, segundo k de 0,64. Com relação ao “hábito de assistir à televisão”, encontrou-se proporção de pessoas com esse hábito de 30,2 e 35,1%, respectivamente para entrevista original e repetida, com concordância moderada. Nenhuma diferença estatisticamente significativa foi encontrada na comparação das proporções entre as entrevistas após teste de McNemar (Tabela 1).

As diferenças entre o tempo total para AF no tempo livre, em minutos, nas duas entrevistas (entrevista repetida – entrevista original) em relação à média das mesmas apresentaram-se, de forma geral, próximas da média (Gráfico 1). Não foram verificadas tendências de sub ou superestimação das variáveis autorreferidas para esse indicador.

Em relação às análises referentes à validade (Tabela 2), não se observam no indicador “AF no tempo livre” diferenças significativas nas frequências obtidas por meio do VIGITEL

Tabela 1. Reprodutibilidade dos indicadores de atividade física e sedentarismo em adultos obtidos por entrevista telefônica. Belo Horizonte, MG, 2013. Amostra de 305 indivíduos.

Indicador	Entrevista telefônica original	Entrevista telefônica repetida	Valor p*	k (IC95%)
	n (%)	n (%)		
Ativos no tempo livre				
Homem	58 (45,3)	53 (41,1)	0,424	0,60 (0,46 – 0,74)
Mulher	55 (31,1)	47 (26,5)	0,076	0,78 (0,67 – 0,88)
Total	113 (37,0)	100 (32,8)	0,059	0,70 (0,62 – 0,78)
Ativos no deslocamento				
Homem	8 (6,2)	9 (7,0)	1,000	0,31 (0,01 – 0,61)
Mulher	27 (15,2)	23 (12,3)	0,571	0,35 (0,16 – 0,54)
Total	35 (11,5)	32 (10,5)	0,749	0,35 (0,19 – 0,51)
Inativos				
Homem	21 (16,4)	20 (15,6)	1,000	0,56 (0,37 – 0,76)
Mulher	24 (13,6)	30 (16,9)	0,179	0,69 (0,55 – 0,84)
Total	45 (14,7)	50 (16,4)	0,458	0,64 (0,52 – 0,76)
Hábito de assistir à televisão				
Homem	95 (25,8)	39 (30,5)	0,263	0,61 (0,46 – 0,76)
Mulher	59 (33,3)	68 (38,4)	0,199	0,52 (0,39 – 0,65)
Total	154 (30,2)	107 (35,1)	0,067	0,56 (0,46 – 0,66)

*Valores p para teste McNemar; k: Coeficiente Kappa; IC95%: intervalo de confiança de 95%.

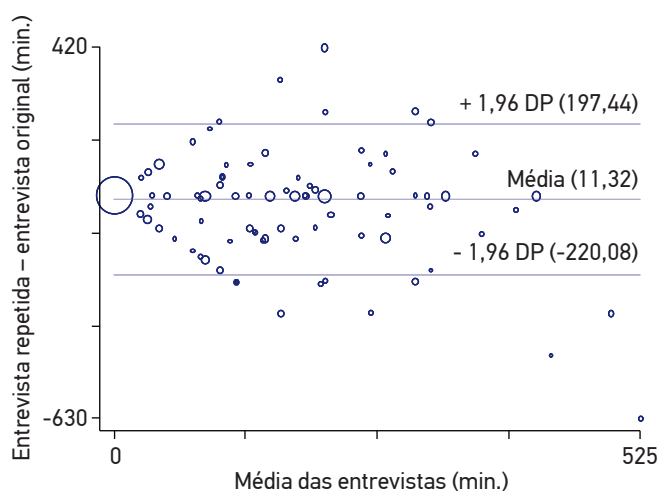


Gráfico 1. Diferenças do tempo de atividade física no tempo livre obtidas na entrevista repetida e original¹. Belo Horizonte, MG, 2013.

(32,8%) em comparação àquelas do GPAQ (30,8%) ($p = 0,538$). Obtiveram-se valores de sensibilidade e especificidade satisfatórios — 67,7% e 82,8% respectivamente —, área sob curva *receiver operator characteristic* (ROC) de 0,75 e um VPP de 63,3% e VPN de 85,2%.

Considerando o indicador AF no deslocamento, houve diferença significativa entre as frequências obtidas pelo questionário VIGITEL e pelo GPAQ ($p < 0,001$). Apenas 9,9% da população estudada foi considerada ativa nesse indicador pelo sistema VIGITEL, enquanto 36,1% foi considerada ativa pelo GPAQ, evidenciando-se grande subestimação dessa medida pelo primeiro questionário, para ambos os sexos. Conseqüentemente, o valor da sensibilidade foi de apenas 11,9%. O valor da

Tabela 2. Frequência, sensibilidade, especificidade, valor preditivo positivo e negativo dos indicadores de atividade física no lazer, atividade física no deslocamento e inatividade. Belo Horizonte, MG, 2013. Amostra de 302 indivíduos.

Indicador	Segunda entrevista VIGITEL		GPAQ		S	E	AUC (IC95%)	VPP	VPN
	n (%)	IC95%	n (%)	IC95%					
Atividade física no tempo livre (lazer)									
Homens	53 (42,1)	33,6 – 51,1	51 (40,5)	31,3 – 49,0	78,0	81,3	0,80 (0,72 – 0,87)	73,6	84,7
Mulheres	46 (26,1)	19,5 – 32,5	42 (23,9)	17,9 – 30,7	55,8	83,6	0,70 (0,62 – 0,78)	52,2	85,5
Total	99 (32,8)	27,4 – 38,1	93 (30,8)	25,5 – 36,0	67,7	82,8	0,75 (0,70 – 0,81)	63,6	85,2
Atividade física no deslocamento*									
Homens	8 (6,4)	2,04 – 10,7	46 (36,5)	28,2 – 45,3	8,7	94,9	0,52 (0,47 – 0,57)	50,0	64,1
Mulheres	20 (12,5)	7,52 – 17,3	63 (35,8)	28,4 – 43,7	14,3	88,6	0,51 (0,46 – 0,58)	40,9	65,2
Total	30 (9,9)	6,54 – 13,3	109 (36,1)	30,6 – 41,5	11,9	91,2	0,52 (0,48 – 0,55)	43,3	64,7
Inatividade*									
Homens	20 (16,0)	9,5 – 22,5	13 (10,3)	4,97 – 15,8	30,8	85,7	0,58 (0,45 – 0,72)	20,0	91,4
Mulheres	30 (17,0)	11,4 – 22,5	18 (10,2)	4,97 – 15,8	72,2	89,3	0,81 (0,70 – 0,92)	43,3	96,6
Total	50 (16,5)	12,3 – 20,8	31 (10,3)	5,67 – 14,7	54,8	87,8	0,71 (0,62 – 0,80)	34,0	94,4

*Valor p para teste McNemar ($< 0,05$); n: número amostral; IC95%: intervalo de confiança de 95%; GPAQ: Questionário Global de Atividade Física; S: sensibilidade; E: especificidade; AUC: área sob curva ROC; VPP: valor preditivo positivo; VPN: valor preditivo negativo.

especificidade foi satisfatório, atingindo 91,2%, com VPP de 43,3% e VPN de 64,7%. A área sob a curva ROC foi de 0,52.

Finalmente, para o indicador de inativos nos quatro domínios de AF, houve superestimação do sistema VIGITEL com relação ao GPAQ ($p = 0,007$), tendo sido obtidas frequências de 16,6% e 10,3%, respectivamente, para a amostra total do estudo. Isso fez com que a sensibilidade ficasse um pouco menor (54,8%) em relação ao domínio de AF no lazer. A especificidade apresentou um valor alto (87,8%) e a área sob a curva ROC foi de 0,71 com VPP de 34,0% e VPN de 94,4%.

DISCUSSÃO

A consolidação do VIGITEL como instrumento central do monitoramento de fatores de risco e proteção para doenças crônicas não transmissíveis no país¹⁵ oportunizou a realização de estudos a respeito da reprodutibilidade e validade dos dados obtidos pelo sistema. Assim, os resultados do presente estudo expõem a qualidade das informações a respeito da prática de AF aferidas pelo VIGITEL, comparando-o ao GPAQ, um questionário validado e recomendado pela OMS. Em relação à reprodutibilidade dos indicadores, os resultados foram aceitáveis tanto na comparação das frequências populacionais (frequências próximas entre os indicadores) quanto na análise intraindividual (k variando de concordância moderada a substancial) para os indicadores investigados — “inatividade”, “hábito de ver televisão” e “AF no tempo livre”. De forma semelhante, a comparação dos dados do VIGITEL com aqueles do método de referência GPAQ apresentou resultados aceitáveis para AF no tempo livre e inatividade, porém para deslocamento observou-se pior desempenho no VIGITEL. O indicador que obteve melhor desempenho na reprodutibilidade entre as entrevistas foi o de AF no tempo livre. Trata-se de um achado importante, uma vez que esse é o domínio de prática mais importante para prevenção de DCNT e redução da mortalidade em geral¹⁶⁻¹⁸. Resultados semelhantes foram encontrados em outras pesquisas sobre o tema^{5,10,19}. Os indicadores “hábito de ver televisão” e “inatividade” tiveram concordâncias aproximadas a de outros estudos de reprodutibilidade^{5,10,20}. Os presentes achados expandem o entendimento sobre o tema identificado nesses estudos, uma vez que se apoia em metodologia inédita, abordando não apenas o indicador de prática de AF, mas também suas dimensões (como a duração da prática).

A reprodutibilidade apresentada pela maior parte dos indicadores sugere padronização adequada das entrevistas e que os participantes compreendem bem as questões, fornecendo respostas similares em diferentes momentos de aferição. Esses achados favorecem a implantação de políticas públicas baseadas em dados populacionais fornecidos pelo questionário de AF do VIGITEL, assim como a realização de outras pesquisas na área de AF a partir do uso desse instrumento, com ressalva ao indicador “AF no deslocamento”. Tal situação pode ser explicada, em parte, pela complexidade dos deslocamentos nas capitais ou até mesmo pela fragilidade das questões utilizadas pelo VIGITEL para a aferição desse indicador, o que

pode ser observado em estudos com resultados semelhantes a este^{5,19,21}. Sabe-se que a seção do questionário do VIGITEL destinada a aferir a AF no deslocamento foi alterada em 2009, visando o melhor desempenho desse indicador (sem que sua validade fosse testada após esse fato). Ainda que não seja possível comparar nossos resultados àqueles obtidos antes dessa alteração, fica explícita a necessidade de nova revisão do questionário.

Quando comparado ao GPAQ, o VIGITEL mostrou desempenho aceitável na maioria das análises. Observaram-se especificidades e valores preditivos negativos elevados para os indicadores de AF no tempo livre e no deslocamento, demonstrando eficácia do sistema em detectar indivíduos que não praticam AF nesses domínios. Além disso, o valor de sensibilidade deste estudo para AF no tempo livre foi satisfatório, similar ou superior ao de outras pesquisas de validação^{10,22}, demonstrando que o VIGITEL tem um desempenho adequado quando comparado a outros instrumentos.

Em outro extremo, para o domínio de AF no deslocamento, obteve-se um valor de sensibilidade baixo e uma grande diferença entre as frequências, verificando-se subestimação dessa medida pelo questionário VIGITEL quando comparado ao GPAQ. Esse resultado insatisfatório pode ser explicado pelas diferenças existentes na composição das perguntas para dimensionar o indicador de deslocamento. Enquanto o sistema VIGITEL considera apenas o deslocamento para curso/escola e trabalho, o GPAQ considera todos os deslocamentos realizados no dia. Portanto, dada a subestimação da real prevalência de ativos no deslocamento na população, o monitoramento pode ser ineficiente nesse aspecto. Além disso, foi verificada também neste estudo baixa reprodutibilidade relacionada a esse indicador, reforçando as fragilidades das perguntas do VIGITEL sobre esse domínio.

Avaliam-se como principal limitação do estudo de validade as diferenças existentes nas perguntas relacionadas aos domínios de AF entre o VIGITEL e o GPAQ, o que dificultou em certa forma a comparabilidade entre os mesmos. Apesar disso, a maioria dos testes de desempenho foi satisfatório e similar àqueles encontrados em muitos estudos de validação existentes na literatura. Ademais, a validação de um questionário que utiliza como referência outro também baseado em autorrelato pode conduzir a erros comuns em ambos os inquéritos, pois tratam-se de duas medidas subjetivas. Entretanto, sabe-se que o GPAQ é um questionário internacionalmente validado para uso em países em desenvolvimento utilizado por muitos pesquisadores para avaliar AF e, por isso, é referência neste estudo. Medidas objetivas poderiam ter sido utilizadas como padrão-ouro — como instrumentos denominados acelerômetros; no entanto, o uso dessa tecnologia também apresenta algumas limitações de aplicabilidade e alto custo.

CONCLUSÃO

Os indicadores de AF e sedentarismo aferidos por entrevistas telefônicas do sistema VIGITEL são adequados para monitorar tendências de indicadores de AF no tempo livre e

de sedentarismo. Porém, foram obtidos resultados insatisfatórios para as questões relativas ao deslocamento. Constatou-se, ainda, a necessidade de outras avaliações com medidas de referência de maior precisão para a confirmação desses resultados.

REFERÊNCIAS

1. Institute of Health Metrics and Evaluation (IHME). GBD Profile: Brazil, 2010. p. 1-4. Disponível em: https://www.healthdata.org/sites/default/files/files/country_profiles/GBD/ihme_gbd_country_report_brazil.pdf (Acessado em 06 de março de 2015).
2. World Health Organization (WHO). Prevalence of insufficient physical activity. Geneva: WHO, 2008.
3. World Health Organization (WHO). Global recommendations on physical activity for health. Geneva, Switzerland: WHO, 2010.
4. Matsudo S, Araújo T, Matsudo V, Andrade D, Andrade E, Oliveira LC, et al. Questionário internacional de atividade física (IPAQ): estudo de validade e reprodutibilidade no Brasil. *Rev Bras Ativ Fís Saúde* 2001; 6(2): 05-18.
5. Bull FC, Maslin TS, Armstrong T. Global physical activity questionnaire (GPAQ): nine country reliability and validity study. *J Phys Act Health* 2009; 6(6): 790-804.
6. Dumith SC. Physical activity in Brazil: a systematic review. *Cad Saúde Pública*. 2009; 25(Suppl 3): S415-26.
7. Armstrong T, Bull F. Development of the World Health Organization Global Physical Activity Questionnaire (GPAQ). *J Public Health* 2006; 14: 66-70.
8. Macías R, Garrido-Muñoz M, Tejero-González CM, Lucia A, López-Adán E, Rodríguez-Romo G. Prevalence of leisure-time sedentary behaviour and sociodemographic correlates: a cross-sectional study in Spanish adults. *BMC Public Health* 2014; 14: 972.
9. Brasil. Ministério da Saúde. Secretaria de Vigilância em Saúde. Secretaria de Gestão Estratégica e Participativa. VIGITEL Brasil 2011. Vigilância de fatores de risco e proteção para doenças crônicas por inquérito telefônico. Brasília: Ministério da Saúde, 2012.
10. Monteiro CA, Florindo AA, Claro RM, Moura EC. Validade de indicadores de atividade física e sedentarismo obtidos por inquérito telefônico. *Rev Saúde Pública* 2008; 42(4): 575-81.
11. Osti RFI, Garcia LMT, Florindo AA. Validation of the 24-hour physical activity recall in elderly adults. *Rev Bras Cineantropom Desempenho Hum* 2014; 16(1): 15-26.
12. World Health Organization (WHO). Global Physical Activity Questionnaire (GPAQ). Analysis Guide. Disponível em: http://www.who.int/chp/steps/resources/GPAQ_Analysis_Guide.pdf (Acessado em 20 de maio de 2015).
13. Lands JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977; 33(1): 159-74.
14. Bland JM, Altman DG. Measuring agreement in method comparison studies. *Stat Methods Med Res* 1999; 8(2): 135-60.
15. Brasil. Ministério da Saúde. Secretaria de Vigilância em Saúde. Departamento de Análise de Situação de Saúde. Plano de ações estratégicas para o enfrentamento das doenças crônicas não transmissíveis (DCNT) no Brasil 2011-2022. Brasília: Ministério da Saúde, 2011.
16. Wagnacker DS, Pitanga FJG. Atividade física comparada a atividade física no tempo livre como fator de proteção à hipertensão arterial sistêmica. *Rev Bras Ci e Mov* 2009; 17(1): 1-19.
17. Perez LG, Pratt M, Simoes EJ, de Moura L, Malta DC. Association between leisure-time physical activity and self-reported hypertension among Brazilian adults, 2008. *Prev Chronic Dis* 2013; 10: E172.
18. Zhao G, Li C, Ford ES, Fulton JE, Carlson SA, Okoro CA, et al. Leisure-time aerobic physical activity, muscle-strengthening activity and mortality risks among US adults: the NHANES linked mortality study. *Br J Sports Med* 2014; 48(3): 244-9.
19. Hallal PC, Simoes E, Reichert FF, Azevedo MR, Ramos LR, Pratt M, et al. Validity and reliability of the telephone-administered international physical activity questionnaire in Brazil. *J Phys Act Health* 2010; 7(3): 402-9.
20. Pettee KK, Ham SA, Macera CA, Ainsworth BE. The reliability of a survey question on television viewing and associations with health risk factors in US adults. *Obesity* 2009; 17(3): 487-93.
21. Matthews CE, Shu XO, Yang G, Jin F, Ainsworth BE, Liu D, et al. Reproducibility and validity of the Shanghai Women's Health Study Physical Activity Questionnaire. *Am J Epidemiol* 2003; 158(11): 1114-22.
22. Medina C, Barquera S, Janssen I. Validity and reliability of the International Physical Activity Questionnaire among adults in Mexico. *Rev Panam Salud Publica* 2013; 34(1): 21-8.

Recebido em: 11/09/2015

Versão final apresentada em: 04/01/2016

Aprovado em: 09/08/2016