

Sistema de Vigilância de Fatores de Risco e Proteção para Doenças Crônicas por Inquérito Telefônico (Vigitel): mudança na metodologia de ponderação

doi: 10.5123/S1679-49742017000400003

Surveillance System for Risk and Protective Factors for Chronic Diseases by Telephone Survey (Vigitel): changes in weighting methodology

Sistema de Vigilancia para Factores de riesgo y Protección de Enfermedades Crónicas por Encuesta Telefónica (Vigitel): cambio en la metodología de ponderación

Regina Tomie Ivata Bernal¹
Betine Pinto Moehlecke Iser²
Deborah Carvalho Malta³
Rafael Moreira Claro⁴

¹Universidade de São Paulo, Núcleo de Pesquisas Epidemiológicas em Nutrição e Saúde, São Paulo-SP, Brasil

²Universidade do Sul de Santa Catarina, Programa de Pós-Graduação em Ciências da Saúde, Tubarão- SC, Brasil

³Universidade Federal de Minas Gerais, Departamento Materno-Infantil e de Saúde Pública, Belo Horizonte-MG, Brasil

⁴Universidade Federal de Minas Gerais, Departamento de Nutrição, Belo Horizonte-MG, Brasil

Resumo

Objetivo: apresentar a metodologia utilizada na construção dos pesos de pós-estratificação do Sistema de Vigilância de Fatores de Risco e Proteção para Doenças Crônicas por Inquérito Telefônico (Vigitel) 2012 e comparar a tendência de indicadores estimados pelos métodos por célula e *rake*. **Métodos:** neste painel de estudos transversais, foram estimadas as prevalências de fumantes, excesso de peso, consumo de frutas, verduras e legumes no período 2006-2012, pelos métodos por célula e *rake*. **Resultados:** não houve diferenças nas tendências temporais dos indicadores estimados por ambos métodos; porém, as prevalências de fumantes estimadas pelo método *rake* foram inferiores às aquelas estimadas pelo método por célula, enquanto as prevalências do consumo de frutas, verduras e legumes foram superiores; para o excesso de peso, não houve diferença entre os métodos. **Conclusão:** o método *rake* forneceu maior precisão das estimativas quando comparado ao método por célula, revelando-se mais vantajoso, embora apresente vício de cobertura de cadastro.

Palavras-chave: Inquéritos Epidemiológicos; Telefone; Viés (Epidemiologia); Vigilância; Estudos Transversais Seriados.

Endereço para correspondência:

Regina Tomie Ivata Bernal – Universidade de São Paulo, Núcleo de Pesquisas Epidemiológicas em Nutrição e Saúde, Av. Dr. Arnaldo, nº 715, São Paulo-SP, Brasil. CEP: 01246-904
E-mail: rbernal@usp.br

Introdução

O Sistema de Vigilância de Fatores de Risco e Proteção para Doenças Crônicas por Inquérito Telefônico (Vigitel), implantado em 2006, está consolidado como sistema de vigilância em saúde e gestão.¹⁻⁶ A população-alvo do Vigitel é composta por adultos com 18 anos ou mais de idade, residentes em domicílios com e sem telefone fixo, nas 26 capitais dos estados brasileiros e no Distrito Federal (DF). Suas informações são utilizadas para o planejamento de políticas públicas, diretrizes e ações específicas de promoção de saúde e prevenção de riscos e doenças.⁷

No processo de amostragem, o Vigitel utiliza os cadastros de telefones residenciais para fins de sorteio de amostras de domicílios. Dados da Pesquisa de Orçamentos Familiares (POF), realizada pela Fundação Instituto Brasileiro de Geografia e Estatística (IBGE) entre 2008 e 2009, indicam que 66,7% das famílias residentes nas capitais dos estados brasileiros e no DF tinham despesas com telefonia fixa nos 30 dias antecedentes ao estudo, e que essa cobertura se manteve estável nos últimos seis anos (2003 a 2008).⁸ Dados do Censo Demográfico 2010⁸ mostram que 61,0% dos domicílios particulares localizados nas 26 capitais e no DF possuíam pelo menos um telefone fixo. Entretanto, a cobertura de telefone fixo não era uniforme no país: as regiões Norte e Nordeste apresentavam 38,0% e 44,0% de cobertura, respectivamente, enquanto o Centro-Oeste, o Sul e o Sudeste referiam coberturas de 56,0%, 70,0% e 74,0%, respectivamente.

Quando parte da população de estudo é excluída da amostra, em consequência da baixa cobertura do cadastro utilizado para fins de sorteio, introduzem-se vícios não amostrais.⁹⁻¹¹ Tal situação é experimentada pelo inquérito Vigitel, sendo especialmente relevante nas capitais do Norte e Nordeste, cuja cobertura de telefonia fixa era inferior a 50% em 2010. Bernal *et al.* mostram que esse vício é desprezível em regiões com cobertura acima de 70%.¹²

A aplicação de pesos de pós-estratificação para compensar a baixa cobertura é comumente utilizada para minimizar esses vícios potenciais.¹³⁻¹⁶ Tal procedimento utiliza variáveis disponíveis na amostra e na população de referência, obtida de fontes externas, para ajustar a distribuição da amostra com telefone àquela verificada para o conjunto completo da população de referência. A escolha das variáveis utilizadas na construção dos pesos leva em consideração as características da população excluída, de forma a minimizar o vício.

Desde a primeira edição do Vigitel, era adotado o método de ponderação por célula¹⁷ na construção dos pesos de pós-estratificação, para cada capital, utilizando-se como base dados da população derivados do Censo Demográfico 2000.⁸ As combinações das categorias das variáveis associadas à posse de telefone fixo – por faixa etária (18 a 24; 25 a 34; 35 a 44; 45 a 54; 55 a 64; 65 ou mais anos), sexo (masculino; feminino) e escolaridade (0 a 8; 9 a 11; 12 anos ou mais de estudo) – compõem as 36 células utilizadas na construção dos pesos, mediante razão da frequência relativa, de cada célula, da população do Censo Demográfico 2000 e da amostra ponderada pelos pesos básicos de delineamento.⁸

Optou-se por rever a metodologia utilizada para construção dos pesos de pós-estratificação do Vigitel, buscando métodos alternativos de ponderação que minimizassem o efeito dos pesos na precisão das estimativas e permitissem considerar as transições sociodemográficas.

Contudo, nas edições recentes do Vigitel, nota-se alta variabilidade nos pesos de pós-estratificação em razão da presença de células com poucas informações na amostra. As células que compõem os estratos da população jovem (18 a 24 anos de idade) e com baixa escolaridade (0 a 8 anos de estudo) são as que apresentam maiores pesos, resultado da baixa frequência na amostra em decorrência das mudanças no perfil sociodemográfico ocorridas no período de 2000 a 2010.

Frente a esses resultados e com a divulgação e disponibilização dos dados da população pelo Censo Demográfico 2010, optou-se por rever a metodologia utilizada para construção dos pesos de pós-estratificação do Vigitel, buscando métodos alternativos de ponderação que minimizassem o efeito dos pesos na precisão das estimativas e permitissem considerar as transições sociodemográficas vivenciadas pela população brasileira na década compreendida entre os Censos Demográficos 2000 e 2010.

O método *rake*^{11,18} foi selecionado para o cálculo dos pesos de pós-estratificação do Vigitel. Esse método utiliza a distribuição de frequência simples de cada variável,

como faixa etária, sexo e escolaridade da população para cada capital, e ainda permite a utilização de diferentes fontes externas no período intercensitário, para construção desses pesos. A principal vantagem desse método é a utilização de distribuições univariadas de diferentes fontes externas, em diferentes períodos de tempo. Como desvantagem do método *rake*, está a necessidade de um algoritmo para o cálculo dos pesos. De sua parte, o método por célula tem como vantagem a facilidade no cálculo dos pesos, e como desvantagem, a necessidade das bases de dados para gerar a tabela de contingência utilizada no cálculo dos pesos.¹¹

O objetivo deste estudo foi apresentar a metodologia utilizada na construção dos pesos de pós-estratificação do Vigitel 2012 e comparar a tendência de indicadores estimados pelo método por célula e pelo método *rake*.

Métodos

Neste painel de estudos transversais, foram estimadas as prevalências de fumantes, excesso de peso, e consumo de frutas, legumes e verduras regularmente em cinco ou mais dias da semana, a partir de dados dos inquéritos Vigitel realizados de 2006 a 2012, pelos métodos por célula e *rake*.

O método *rake*^{11,18} foi utilizado como método alternativo no cálculo dos pesos de pós-estratificação para expansão da amostra do inquérito Vigitel. Esse método utiliza a distribuição de frequência anual das variáveis 'faixa etária', 'sexo' e 'escolaridade' da população. Devido à ausência da projeção anual do nível de instrução da população, optou-se por estimar as projeções do nível de instrução, sexo e faixa etária a partir dos dados dos Censos 2000 e 2010.

A seguir, são apresentados os procedimentos utilizados no cálculo das projeções e a construção dos pesos pelo método *rake* para cada capital. Inicialmente, foram considerados os microdados dos Censos Demográficos de 2000¹⁸ e 2010¹⁹ para cada capital (i), no cálculo da taxa média geométrica ($rg_{(i)}$) do crescimento anual da população adulta utilizada para projeção da população no período intercensitário – Expressão 1:

$$rg_{(i)} = \left(\sqrt[10]{\frac{\text{População}_{(i)}^{2010}}{\text{População}_{(i)}^{2000}}} \right) - 1$$

Sobre cada capital, calculou-se a taxa de crescimento da população, estratificada por sexo, para cada uma das seis categorias de idade (18 a 24; 25 a 34; 35 a 44;

45 a 54; 55 a 64; e 65 anos ou mais), e para quatro níveis de instrução (sem instrução ou com Ensino Fundamental incompleto; com Ensino Fundamental completo ou Ensino Médio incompleto; com Ensino Médio completo ou Ensino Superior incompleto; com Ensino Superior completo). Devido a mudanças nas perguntas entre os Censos 2000 e 2010, o IBGE disponibilizou as combinações das variáveis classificatórias do nível de instrução em cada censo. O uso de técnica de imputação de dados fez-se necessário na presença da categoria 'Não determinada' da variável 'nível de instrução', adotando-se o critério do valor da categoria mais frequente dentro de cada faixa etária – estratificada por sexo – para os casos faltantes.

As estimativas da população adulta de cada capital, no período intercensitário (t), foram obtidas a partir da taxa média geométrica de crescimento anual estimada para cada um dos 14 grupos da população (seis faixas de idade para o conjunto total da população e quatro faixas de instrução para cada um dos sexos). Adicionalmente, a estimativa do total da população do Vigitel por idade foi ajustada pela população total do IBGE, divulgada pelo Departamento de Informática do Sistema Único de Saúde (Datasus)¹⁹ a cada ano do inquérito. Esse ajuste é resultante da multiplicação da frequência total relativa da idade, estratificada por sexo, pelo número total de adultos com 18 anos ou mais de idade na população. O mesmo procedimento foi usado para o nível de instrução, estratificado por sexo. Em ambos casos, o total populacional estimado a partir da amostra de adultos foi ajustado ao tamanho total da população.

Para cada capital, foram calculadas as estimativas populacionais por faixa etária (seis categorias) e por nível de instrução estratificado por sexo (oito categorias), no período de 2006 a 2011 – exceto 2010. Essas informações foram utilizadas na construção dos pesos de pós-estratificação do Vigitel pelo método *rake*.^{11,18} Este método trabalha com uma variável de cada vez, equiparando a distribuição do total da variável na amostra ponderada pelos pesos amostrais e na população, mediante procedimentos de iteração. Logo, esse processo é repetido em cada uma das variáveis usadas na construção dos pesos, fazendo com que a distribuição da amostra fique idêntica à da população para essas variáveis. Os novos pesos de pós-estratificação do Vigitel, para cada capital, foram calculados sobre o aplicativo estatístico SAS (utilizou-se a função macro *Rakinge.sas*²¹).

Os pesos amostrais do Vigitel, em cada capital, são calculados pela razão entre o número de adultos e o número de telefones fixos, em razão do processo de amostragem que envolve sorteio da amostra em dois estágios: no primeiro, são utilizados os cadastros de telefones residenciais para fins de sorteio de amostras de domicílios; e no segundo estágio, é realizado o sorteio de um adulto residente no domicílio sorteado.

As análises dos dados do Vigitel requerem consideração do plano complexo de amostragem (PCA) do inquérito. A medida usada para avaliar o impacto dos pesos na precisão da prevalência é o efeito de delineamento (*deff*),⁹ que expressa a razão entre as variâncias das estimativas calculadas com pesos e a amostra casual simples. Um valor *deff* igual a 1 indica que os pesos não alteraram a precisão da estimativa, enquanto valores *deff* maiores que 1 indicam perda de precisão decorrente do uso dos pesos de pós-estratificação. Dessa maneira, foi calculado o *deff* para cada indicador usando-se os pesos de pós-estratificação pelos métodos por célula e *rake*. Espera-se que o *deff* obtido pelo método *rake* seja menor que o *deff* obtido pelo método por célula.

Com o propósito de apresentar a metodologia *rake* aplicada no inquérito Vigitel a partir de 2012, serão ilustradas as seguintes etapas:

- a) estimação das taxas médias geométricas anuais de crescimento da população da cidade de Belém-PA segundo faixa etária e nível de instrução, de 2000 a 2010;
- b) comparação das estimativas da população por grupo etário, fornecidas pelo IBGE e disponibilizadas no sítio eletrônico do Datatus, com as estimativas produzidas pelo Vigitel;
- c) detalhamento metodológico da ponderação *rake*;
- d) estimação do efeito de delineamento – *deff* – dos pesos do Vigitel na prevalência fumantes, por capital, por ano – entre 2006 e 2011 – e por método de ponderação – *rake* e por célula –; e
- e) comparação das diferenças entre séries temporais de indicadores selecionados (prevalência de fumantes; excesso de peso; frequência de consumo de frutas, verduras e legumes), no período de 2006 a 2012, segundo o emprego dos distintos métodos de ponderação – método por célula e método *rake*.

O estudo utilizou apenas dados secundários, sem identificação dos sujeitos. Foi realizado em conformidade com os princípios éticos previstos na Resolução do Conselho Nacional de Saúde (CNS) nº 510, de 7 de abril de 2016.

Resultados

No período de 2000 a 2010, a população adulta residente na capital Belém-PA aumentou, em média, 1,2% ao ano, embora na estratificação por idade, a taxa de variação do tamanho da população no grupo etário de 18 a 24 anos tenha se mostrado negativa. Observou-se maior crescimento no grupo de 55 a 64 anos. A população sem instrução ou com o Ensino Fundamental incompleto diminuiu 1,7% a cada ano, ao passo que a população com Ensino Médio completo ou Ensino Superior incompleto aumentou, em média, 5,5% ao ano. Já a população com Ensino Superior completo aumentou quase 7% a cada ano, com crescimento superior entre as mulheres: 7,6% (Tabela 1).

Em 2011, a saída do programa Raking.sas executado para Belém-PA (Figura 1) mostra, como exemplo, a saída da primeira e última iteração para o cálculo dos pesos de pós-estratificação. A primeira variável na construção dos pesos é a faixa etária, apresentada na coluna (1), com seis categorias (18 a 24; 25 a 34; 35 a 44; 45 a 54; 55 a 64; e 65 anos ou mais de idade). Na primeira iteração da faixa etária, observa-se que o número de adultos ponderado pelo peso básico da amostra é igual a 6.116,75 (2); e na população, igual a 999.061 adultos (3). Quando comparada a distribuição relativa da amostra (4) com a da população (5), nota-se que o Vigitel tem menos adultos entre 18 e 44 anos, e mais adultos com 45 anos ou mais de idade (6). A segunda variável é o nível de instrução estratificado por sexo (7), com oito categorias:

- i) homens sem instrução ou com Ensino Fundamental incompleto;
- ii) homens com Ensino Fundamental completo ou Ensino Médio incompleto;
- iii) homens com Ensino Médio completo ou Ensino Superior incompleto;
- iv) homens com Ensino Superior completo;
- v) mulheres sem instrução ou com Ensino Fundamental incompleto;
- vi) mulheres com Ensino Fundamental completo ou Ensino Médio incompleto;
- vii) mulheres com Ensino Médio completo ou Ensino Superior incompleto; e
- viii) mulheres com Ensino Superior completo.

Nessa etapa, o total da amostra expandida do Vigitel (8) já é igual ao total da população,

Tabela 1 – Taxa média geométrica anual de variação proporcional (%) do tamanho da população segundo faixa etária, nível de instrução e sexo, no município de Belém, Pará, 2000-2010

Variáveis	Sexo		Total
	Masculino	Feminino	
Faixa etária (em anos)			
18-24	-0,797	-0,983	-0,895
25-34	1,520	1,525	1,522
35-44	2,083	2,173	2,131
45-54	3,474	3,492	3,483
55-64	4,193	4,143	4,165
≥65	3,942	3,816	3,864
Nível de instrução			
Sem instrução/Ensino Fundamental incompleto	-1,477	-1,952	-1,727
Ensino Fundamental completo/Ensino Médio incompleto	0,396	0,476	0,438
Ensino Médio completo/Ensino Superior incompleto	5,645	5,319	5,461
Ensino Superior completo	6,275	7,565	6,980
Total	1,193	1,200	1,197

segundo o IBGE (9); porém, quando comparada a distribuição relativa do nível de instrução na amostra e na população, observa-se no Vigitel a diferença de -8,8% de homens e -5,9% de mulheres, ambos sem instrução ou com Ensino Fundamental incompleto. Por sua vez, verifica-se mais adultos com Ensino Superior completo. Na segunda iteração, observa-se que a amostra ponderada da variável 'faixa etária' é a mesma da população; contudo, a diferença entre as distribuições relativas da população e da amostra é diferente de zero. Esse procedimento é repetido até que a diferença relativa entre o Vigitel e a população seja igual a zero. Nesse exemplo, para se atender ao critério de convergência, foram necessárias sete iterações até que as distribuições da faixa etária (10) e do nível de instrução estratificado por sexo (11) fossem iguais a zero. No final do processo, é gerado o arquivo com os pesos de pós-estratificação.

A avaliação do impacto dos pesos na precisão da prevalência de fumantes, expressa pela medida de efeito de delineamento (*deff*), mostrou valores $deff > 1$; isto significa que há perda de precisão no uso de pesos de pós-estratificação. Notou-se valores *deff* produzidos pelo método por célula superiores àqueles produzidos pelo método *rake*, em todas as capitais e anos. Quanto maior o valor *deff*, menor a precisão da

estimativa (Tabela 2). A perda de precisão decorre da alteração no tamanho efetivo da amostra, expressa pela seguinte fórmula:

$$n_{deff} = n/deff$$

Como exemplo, em 2006, a capital Palmas-TO apresenta *deff* de 5,12 e 2,54 – o que corresponde ao tamanho efetivo da amostra de 391 (2000/5,12) e 787 (2000/2,54), respectivamente. Ou seja, o uso de pesos faz com que o tamanho da amostra igual a 2000 corresponda ao tamanho da amostra efetivo, equivalente a 391 e 787 entrevistas respectivamente (Tabela 2).

Comparando-se as séries históricas da prevalência de adultos fumantes, para o conjunto das capitais e o DF, estimada pelos métodos por célula e *rake*, observou-se tendência de declínio pelos dois métodos; porém, no método *rake*, a tendência mostrou-se mais acentuada, com prevalências inferiores àquelas estimadas pelo método por célula. Para a estimativa da prevalência de excesso de peso, não se observou diferença entre os dois métodos. A série da estimativa da prevalência de consumo de frutas, legumes e verduras regularmente em cinco ou mais dias da semana apresentou-se crescente, e as prevalências estimadas pelo método *rake* foram superiores àquelas pelo método por célula (Figura 2).

Raking by FET, iteration - 1						
(1)	(2)	(3)		(4)	(5)	(6)
FET	Calculated margin	Marginal control total	Difference	Calculated %	Marginal Control %	Difference in %
1	1047.67	181627	180579.33	17.128	18.180	-1.052
2	1333.50	262280	260946.50	21.801	26.253	-4.452
3	1037.75	208865	207827.25	16.966	20.906	-3.940
4	1043.67	156756	155712.33	17.062	15.690	1.372
5	807.00	99599	98792.00	13.193	9.969	3.224
6	847.17	89934	89086.83	13.850	9.002	4.848
	=====	=====		=====	=====	
	6116.75	999061		100.00	100.00	
Belem - 2011 16:37 Sunday, February 1, 2009 106						
Raking by SEXOFXESC, iteration - 1						
(7)						
SEXOFXESC	Calculated margin	Marginal control total	Difference	Calculated %	Marginal Control %	Difference in %
1	55018.68	142750	87731.32	5.507	14.288	-8.781
2	64278.60	90902	26623.40	6.434	9.099	-2.665
3	199603.57	172392	-27211.57	19.979	17.255	2.724
4	87414.01	53609	-33805.01	8.750	5.366	3.384
5	93738.04	153077	59338.96	9.383	15.322	-5.939
6	85415.87	98583	13167.13	8.550	9.868	-1.318
7	286341.14	217849	-68492.14	28.661	21.805	6.856
	=====	=====		=====	=====	
(8)	999061.00	999061	(9)	100.00	100.00	
	-	-	-	-	-	-
	-	-	-	-	-	-
	-	-	-	-	-	-
Raking by FET, iteration - 7						
FET	Calculated margin	Marginal control total	Difference	Calculated %	Marginal Control %	Difference in %
1	181623.37	181627	3.62626	18.179	18.180	-0.000
2	262277.63	262280	2.37050	26.252	26.253	-0.000
3	208864.48	208865	0.52154	20.906	20.906	-0.000
4	156757.15	156756	-1.15096	15.690	15.690	0.000
5	99600.72	99599	-1.71775	9.969	9.969	0.000
6	89937.65	89934	-3.64959	9.002	9.002	0.000
	=====	=====		=====	=====	
	999061.00	999061		100.00	100.00	
Belem - 2011 16:37 Sunday, February 1, 2009 118						
Raking by SEXOFXESC, iteration - 7						
FET	Calculated margin	Marginal control total	Difference	Calculated %	Marginal Control %	Difference in %
1	142748.58	142750	1.41543	14.288	14.288	-0.000
2	90902.30	90902	-0.30409	9.099	9.099	0.000
3	172393.06	172392	-1.05909	17.256	17.255	0.000
4	53609.17	53609	-0.16651	5.366	5.366	0.000
5	153075.42	153077	1.58073	15.322	15.322	-0.000
6	98583.07	98583	-0.07237	9.868	9.868	0.000
7	217850.19	217849	-1.19216	21.805	21.805	0.000
8	69899.20	69899	-0.20193	6.996	6.996	0.000
	=====	=====		=====	=====	
	999061.00	999061		100.00	100.00	

Figura 1 – Saída do programa Rakinge.sas para cálculo do peso de pós-estratificação no município de Belém, Pará, 2011

Tabela 2 – Distribuição do efeito de delineamento dos pesos do Sistema de Vigilância de Fatores de Risco e Proteção para Doenças Crônicas por Inquérito Telefônico (Vigitel) produzidos pelos métodos por célula e por *rake*, na prevalência de fumantes, por capital e ano, nas 26 capitais dos estados brasileiros e no Distrito Federal, 2006-2011

Capital	2006		2007		2008		2009		2010		2011	
	Por célula ^a	Por <i>rake</i> ^b	Por célula ^a	Por <i>rake</i> ^b	Por célula ^a	Por <i>rake</i> ^b	Por célula ^a	Por <i>rake</i> ^b	Por célula ^a	Por <i>rake</i> ^b	Por célula ^a	Por <i>rake</i> ^b
Aracaju	2,84	2,22	5,16	2,57	3,69	2,43	4,04	2,07	3,84	2,54	3,87	2,11
Belém	1,87	1,66	2,81	2,10	3,43	2,47	3,49	2,23	5,48	3,09	5,25	2,52
Belo Horizonte	2,02	1,53	2,31	1,61	4,84	1,62	2,71	1,46	2,96	1,57	2,80	1,50
Boa Vista	3,00	2,22	3,86	2,31	5,66	3,64	6,18	2,87	8,80	2,79	7,54	3,16
Campo Grande	2,51	1,79	3,48	2,06	3,40	2,02	4,14	1,78	3,11	1,62	3,74	1,53
Cuiabá	3,47	2,02	2,40	1,78	3,05	2,08	2,40	1,56	3,73	1,82	4,26	1,93
Curitiba	1,88	1,45	1,78	1,42	2,22	1,44	1,99	1,49	2,40	1,52	3,05	1,45
Florianópolis	2,19	1,55	2,54	1,77	2,40	1,47	3,25	1,63	2,72	1,66	2,45	1,53
Fortaleza	2,82	2,02	3,74	2,29	2,86	1,94	4,17	2,34	5,19	2,10	4,29	1,95
Goiânia	2,22	1,53	2,47	1,64	2,80	1,47	5,25	1,73	3,47	1,48	2,84	1,39
João Pessoa	2,87	2,01	3,41	2,28	10,40	2,53	4,30	2,45	7,37	3,30	3,11	2,28
Macapá	3,27	2,46	4,17	2,66	4,43	2,29	5,51	3,32	5,45	2,94	4,21	2,37
Maceió	3,03	2,01	2,76	2,29	3,14	1,98	5,72	3,15	6,36	2,66	4,08	2,29
Manaus	3,03	1,91	2,49	1,69	3,43	2,29	4,23	2,54	3,90	2,22	3,86	2,06
Natal	2,67	1,93	2,36	2,02	2,91	2,13	4,05	2,14	5,15	2,12	7,53	2,67
Palmas	5,12	2,54	3,79	2,42	7,68	2,47	6,22	2,46	7,86	2,44	6,30	2,42
Porto Alegre	3,05	1,75	2,62	1,79	2,07	1,52	2,37	1,76	2,62	1,98	2,20	1,49
Porto Velho	3,10	2,25	3,42	2,45	3,37	2,60	3,78	2,34	3,67	2,37	4,34	2,61
Recife	2,06	1,64	4,03	2,07	2,23	1,71	3,24	1,75	3,57	1,81	3,66	1,69
Rio Branco	3,02	2,07	3,92	2,77	6,91	2,73	5,38	3,27	7,13	2,77	4,61	2,23
Rio de Janeiro	1,53	1,38	2,66	1,53	2,81	1,62	2,77	1,56	2,41	1,45	3,06	1,52
Salvador	2,01	1,41	3,01	2,00	2,50	1,59	5,18	2,04	2,14	1,68	4,70	1,94
São Luis	2,45	1,95	3,06	2,15	2,91	2,64	3,38	2,32	3,80	2,06	6,40	2,64
São Paulo	2,19	1,43	2,12	1,48	4,92	1,56	3,76	1,61	2,29	1,56	2,63	1,48
Teresina	3,75	2,07	3,24	2,23	3,92	2,33	5,18	2,57	4,82	2,53	6,73	2,43
Vitória	2,37	1,52	2,92	1,43	4,42	1,50	4,35	1,72	3,37	1,76	3,39	1,48
Distrito Federal	4,25	2,68	3,23	1,68	3,64	1,73	8,45	4,35	11,22	6,08	5,16	1,71

Notas:

a) população do Censo 2000 e método de ponderação por célula.

b) estimativa da população no ano da pesquisa e método de ponderação *rake*.

Discussão

Este é o primeiro estudo a comparar os dois métodos de pós-estratificação do Vigitel: por célula e *rake*. O método *rake* passou a ser utilizado no Vigitel em 2012 e apresenta vantagens em relação ao método por célula, ao utilizar dados da população anual, os quais refletem as mudanças ocorridas na população brasileira, enquanto o método por célula utiliza o

mesmo dado da população para diferentes anos da pesquisa. O método *rake*, apesar de não eliminar o vício de cobertura de cadastro, produz estimativas que expressam, com maior precisão, as tendências dos indicadores de fatores de risco na população, na comparação com o método por célula. O novo método de ponderação já foi utilizado para produção de novas estimativas com dados do Vigitel, no período de 2006 a 2011.

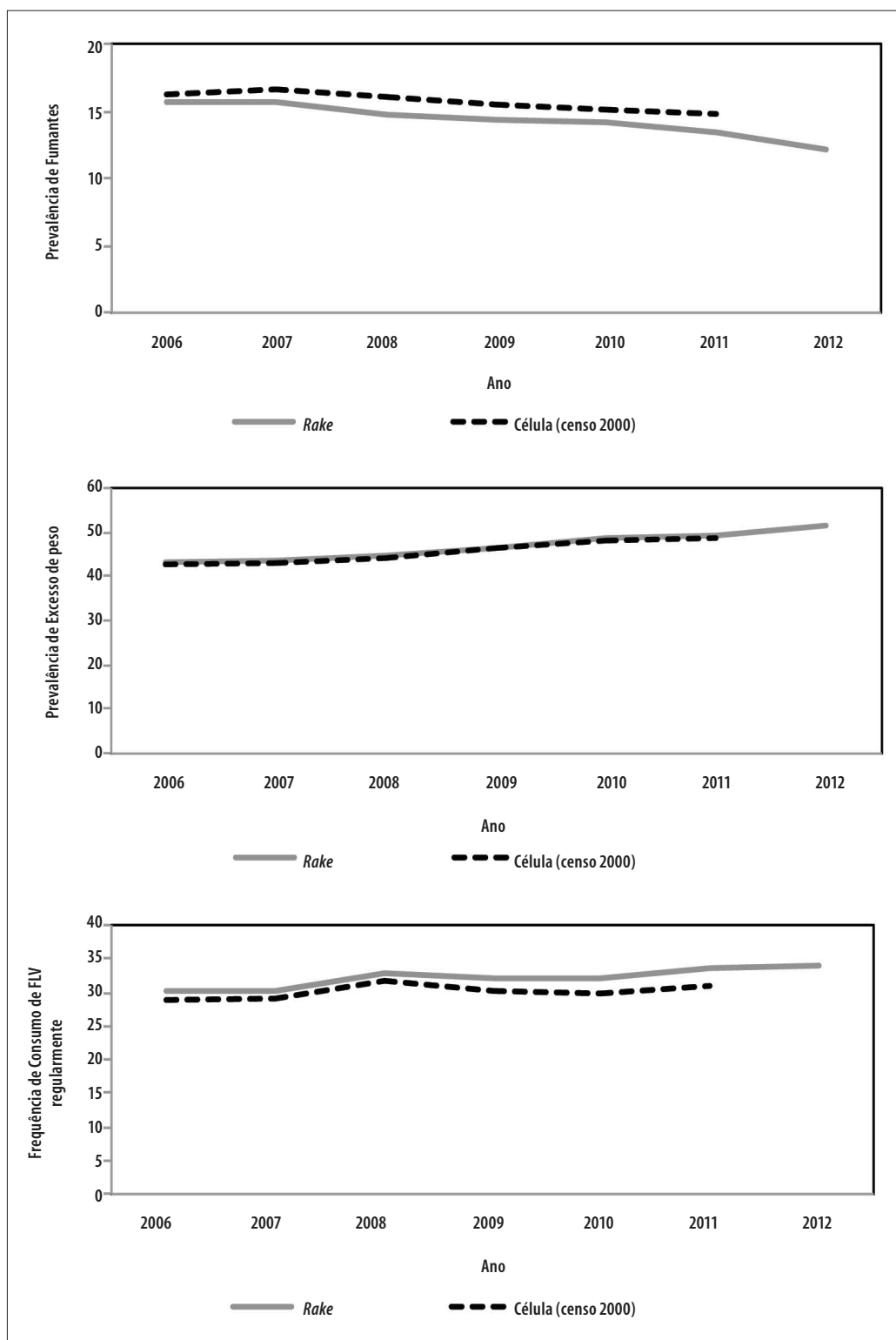


Figura 2 – Série temporal das prevalências de fumantes, excesso de peso e consumo de frutas, verduras e legumes regularmente em cinco ou mais dias da semana, obtidas com dados do Vigitel, segundo o método de ponderação, nas 26 capitais dos estados brasileiros e no Distrito Federal, 2006-2012

Há vários métodos para construir pesos de pós-estratificação. Entre eles, o método de ponderação por célula é amplamente utilizado devido à simplicidade nos cálculos dos pesos, que utilizam as células das tabelas de contingência da população e da amostra.¹¹ O fator limitante desse método reside na necessidade de se ter os dados disponíveis da população para gerar a tabela de contingência das variáveis, visto que essas informações são disponibilizadas tão somente a cada dez anos (pelo Censo Demográfico do IBGE). A população brasileira passa por um período de transição demográfica e socioeconômica,²⁰ e esse longo período de uma década de espera constitui uma limitação ao uso desse método. Por sua vez, o método *rake* utiliza apenas as frequências simples de cada variável da população e da amostra, o que permite o uso de informações populacionais de diferentes fontes externas, como também de interpolações das variáveis da população no período intercensitário, refletindo as mudanças da população ocorridas no período. O fator limitador do método *rake* encontra-se na maior complexidade do processo de cálculo dos pesos, porque necessita de um algoritmo disponível no pacote estatístico para o cálculo dos pesos de pós-estratificação.

Na comparação entre as estimativas de ambas metodologias de ponderação, as tendências dos três indicadores foram similares; porém, os valores das prevalências foram distintos entre o método *rake* e o método por célula. Notaram-se diferenças nas séries das prevalências de fumantes e de consumo de frutas, legumes e verduras, enquanto para a prevalência de excesso de peso não houve diferença. As diferenças observadas podem ser explicadas pela associação entre os indicadores, com referência à posse de telefone fixo.^{21,22}

No caso dos indicadores que estiverem positivamente associados à escolaridade, estima-se que o aumento da escolarização de população na década intercensitária, bem como seu envelhecimento, levariam a maior redução nas tendências desses indicadores, conforme demonstrado por este estudo. Outros estudos apontam que a prevalência do fumo está associada a baixa escolaridade.²³ O aumento da escolarização dos brasileiros no período intercensitário justifica a redução da prevalência de fumantes, captada de forma mais consistente pela metodologia de ponderação *rake*, haja vista esse método atualizar as estimativas populacionais de sexo, idade e escolaridade. Assim, o método *rake* conseguiu

revelar prevalências menores; e uma tendência de declínio mais acentuado do tabagismo na população brasileira, confirmada pela Pesquisa Nacional de Saúde (PNS), realizada pelo Ministério da Saúde em parceria com o IBGE, que apontou redução de 20% na prevalência do tabagismo na população brasileira entre 2008 e 2013.²⁴ Portanto, o uso do método *rake* possibilitou obter estimativas mais precisas, a partir dos dados do Vigitel.

De forma semelhante, nos Estados Unidos da América (EUA), Battaglia *et al.*¹³ avaliaram os pesos de pós-estratificação dos dados de 2003 disponibilizados pelo *Behavioral Risk Factor Surveillance System* (BRFSS), cuja proporção de resposta foi inferior a 50%. Foram comparadas as estimativas obtidas pelo método de ponderação por célula e pelo método *rake*, ambos utilizando os dados populacionais do *Current Population Survey* (CPS) para construção dos pesos de pós-estratificação, com vistas à redução do vício de não resposta. Os resultados do estudo norte-americano mostram que o BRFSS tendeu a subestimar os indicadores de fatores de risco da população devido à baixa proporção de resposta, inferior a 50%, sobretudo entre a população de baixa renda – a qual, por sua vez, apresentava maior prevalência dos fatores de risco. Já as estimativas das prevalências obtidas pelo método *rake* foram maiores, quando comparadas às fornecidas pelo método de ponderação por célula, dada a associação entre as variáveis utilizadas na construção dos pesos e os fatores de risco.

Em 2011, os Centers for Disease Control and Prevention (CDC) dos EUA implantaram várias mudanças no BRFSS.²⁵ O CDC/EUA incluiu na amostra de seu inquérito a população que usa apenas o telefone móvel (somada à população atendida pela telefonia fixa) e mudou a metodologia de ponderação. O sistema passou a utilizar o método *rake* na construção dos pesos de pós-estratificação, tendo em vista o cadastro duplo (telefone fixo e móvel), a redução do vício de não resposta e a melhoria na precisão das estimativas. O processo de mudança na metodologia adotada pelo BRFSS foi recomendado por especialistas da área, a partir da divulgação de estudos que avaliaram diferentes variáveis utilizadas na construção dos pesos de pós-estratificação e métodos de ponderação, com vistas à redução de vício decorrente da baixa taxa de resposta.^{26,27}

Quanto ao Vigitel, destaca-se que os métodos de pós-estratificação utilizados visam reduzir possíveis distorções

nas estimativas inerentes ao método de seleção dos participantes, por meio de telefones fixos, o que gera viés de seleção. Para eliminar o vício, o Vigitel deveria introduzir em sua pesquisa a população sem telefone fixo.

Enquanto não forem incorporadas modificações no processo de amostragem do Vigitel, recomenda-se a utilização do método *rake*. Este possibilita a obtenção de estimativas de prevalências mais precisas – na comparação com os resultados obtidos pelo método por célula –, tornando-as mais próximas da realidade apontada no Censo 2010. Com isso, é possível melhor apoiar o monitoramento dos fatores de risco e proteção para doenças crônicas não transmissíveis, ademais

de subsidiar, de forma mais adequada, as políticas públicas de promoção da saúde.

Contribuição dos autores

Malta DC contribuiu no delineamento do estudo e revisão do artigo. Bernal RTI contribuiu no delineamento do estudo, realizou a análise estatística e elaborou o texto do artigo. Iser BPM e Claro RM deram contribuições substanciais à análise dos dados, redação e revisão do artigo. Todos os autores aprovaram a versão final e declaram-se responsáveis pelo conteúdo integral do artigo, garantindo sua precisão e integridade.

Referências

1. Ministério da Saúde (BR). Secretaria de Vigilância em Saúde. Secretaria de Gestão Estratégica e Participativa. Vigitel Brasil 2006: vigilância de fatores de risco e proteção para doenças crônicas por inquérito telefônico [Internet]. Brasília: Ministério da Saúde, 2007 [citado 2017 ago 23]. 297 p. Disponível em: http://bvsms.saude.gov.br/bvs/publicacoes/vigitel_brasil_2006.pdf
2. Ministério da Saúde (BR). Secretaria de Vigilância em Saúde. Secretaria de Gestão Estratégica e Participativa. Vigitel Brasil 2007: vigilância de fatores de risco e proteção para doenças crônicas por inquérito telefônico [Internet]. Brasília: Ministério da Saúde, 2008 [citado 2017 ago 23]. 136 p. Disponível em: http://bvsms.saude.gov.br/bvs/publicacoes/vigitel_brasil_2007.pdf
3. Ministério da Saúde (BR). Secretaria de Vigilância em Saúde. Secretaria de Gestão Estratégica e Participativa. Vigitel Brasil 2008: vigilância de fatores de risco e proteção para doenças crônicas por inquérito telefônico [Internet]. Brasília: Ministério da Saúde, 2009 [citado 2017 ago 23]. 112 p. Disponível em: http://bvsms.saude.gov.br/bvs/publicacoes/vigitel_brasil_2008.pdf
4. Ministério da Saúde (BR). Secretaria de Vigilância em Saúde. Secretaria de Gestão Estratégica e Participativa. Vigitel Brasil 2009: vigilância de fatores de risco e proteção para doenças crônicas por inquérito telefônico [Internet]. Brasília: Ministério da Saúde, 2010 [citado 2017 ago 23]. 150 p. Disponível em: <http://www.abeso.org.br/uploads/downloads/75/553a24c01eb80.pdf>
5. Ministério da Saúde (BR). Secretaria de Vigilância em Saúde. Secretaria de Gestão Estratégica e Participativa. Vigitel Brasil 2010: vigilância de fatores de risco e proteção para doenças crônicas por inquérito telefônico [Internet]. Brasília: Ministério da Saúde, 2011 [citado 2017 ago 23]. 152 p. Disponível em: http://bvsms.saude.gov.br/bvs/publicacoes/vigitel_2010.pdf
6. Ministério da Saúde (BR). 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 [Internet]. Brasília: Ministério da Saúde, 2012 [citado 2017 ago 23]. 132 p. Disponível em: http://bvsms.saude.gov.br/bvs/publicacoes/vigitel_brasil_2011_fatores_risco_doencas_cronicas.pdf
7. Ministério da Saúde (BR). 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 [Internet]. Brasília: Ministério da Saúde, 2011 [citado 2017 ago 23]. 148 p. Disponível em: http://bvsms.saude.gov.br/bvs/publicacoes/plano_acoes_enfrent_dcnt_2011.pdf
8. Instituto Brasileiro de Geografia e Estatística. Microdados da pesquisa de orçamentos familiares [Internet]. 2014 [citado 2013 maio 14]. Disponível em: http://downloads.ibge.gov.br/downloads_estatisticas.htm
9. Kish L. Survey sampling. New York: John Wiley & Sons; 1965.
10. Cochran WG. Sampling techniques. 3rd ed. New York: John Wiley & Sons; 1977.

11. Kalton G. Compensating for missing survey data. Michigan: Institute for Social Research, The University of Michigan; 1983.
12. Bernal R, Silva NN. Cobertura de linhas telefônicas residenciais e vícios potenciais em estudos epidemiológicos. *Rev Saúde Pública*. 2009;43(3):421-6.
13. Battaglia MP, Frankel MR, Link MW. Improving standard poststratification techniques for random-digit-dialing telephone surveys. *Surv Res Methods* [Internet]. 2008 [cited 2013 Nov 24];2(1):11-9. Available from: <https://ojs.ub.uni-konstanz.de/srm/article/view/597>.
14. Lepkowski JM, Tucker NC, Brick JM, Leeuw E, Japac L, Lavraskas PJ, et al. *Advances in telephone survey methodology*. Hoboken (NJ): John Wiley & Sons; 2008.
15. Flores-Cervantes I, Brick JM, Jones ME, Westat J, DiSogra C, Yen W. Weighting for nontelephone household in the 2001 California Health Interview Survey [Internet]. In: *Proceedings of the Joint Statistical Meetings Section on Survey Research Methods*; 2002 Aug 11-15; New York. p.1002-7 [cited 2013 May 14]. Available from: <http://www.amstat.org/sections/srms/Proceedings/y2002/Files/JSM2002-000661.pdf>.
16. Bernal RTI, Malta DC, Araújo TS, Silva NN. Inquérito por telefone: pesos de pós-estratificação para corrigir vícios de baixa cobertura em Rio Branco, AC. *Rev Saúde Pública*. 2013 abr;47(2):316-25.
17. Moura EC, Moraes Neto OL, Malta DC, Moura L, Silva NN, Bernal R, et al. Vigilância de fatores de risco para doenças crônicas por inquérito telefônico nas capitais dos 26 Estados brasileiros e no Distrito Federal (2006). *Rev Bras Epidemiol*. 2008;11 (supl 1):20-37.
18. Izrael D, Hoaglin DC, Battaglia MP. A SAS macro for balancing a weighted sample [Internet]. In: *Proceedings of the Twenty-Fifth Annual SAS Users Group International Conference*, Paper 275, 2000 [cited 2010 Dec 12]. Available from: <http://www2.sas.com/proceedings/sugi29/207-29.pdf>.
19. Ministério da Saúde (BR). Portal da saúde. Datasus. Demográficas e socioeconômicas [Internet]. 2012 [citado 2013 dez 12]. Disponível em: <http://www2.datasus.gov.br/DATASUS/index.php?area=0206>.
20. Alves JED, Cavenaghi S. Tendências demográficas, dos domicílios e das famílias no Brasil [Internet]. *Aparte: inclusão social em debate*. 2012 [citado 2013 out 11]. Disponível em: http://www.ie.ufrj.br/aparte/pdfs/tendencias_demograficas_e_de_familia_24ago12.pdf.
21. Bernal RTI. Inquéritos por telefone: inferências válidas em regiões com baixa taxa de cobertura de linhas residenciais [tese]. São Paulo (SP): Universidade de São Paulo, 2011.
22. Malta DC, Iser BPM, Sá NNB, Yokota RTC, Moura L, Claro RM, et al. Tendências temporais no consumo de tabaco nas capitais brasileiras, segundo dados do Vigitel, 2006 a 2011. *Cad Saúde Pública* [Internet]. 2013 abr [citado 2013 set 01];29(4):812-22. Disponível em: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0102-311X2013000400018&lng=en&nrm=iso
23. Giovino GA, Mirza SA, Samet JM, Gupta PC, Jarvis MJ, Bhalra N, et al. Tobacco use in 3 billion individuals from 16 countries: an analysis of nationally representative cross-sectional household surveys. *Lancet*. 2012 Aug;380(9842):668-79.
24. Malta DC, Vieira ML, Szwarcwald CL, Caixeta R, Brito SMF, Reis AAC. Tendência de fumantes na população brasileira segundo a pesquisa nacional de amostra de domicílios 2008 e a pesquisa nacional de saúde 2013. *Rev Bras Epidemiol*. 2015 dez;18(suppl 2):45-56.
25. Center for Disease Control and Prevention. Behavioral risk factor surveillance system: BRFSS 2011 survey data and documentation [Internet]. Atlanta: Center for Disease Control and Prevention; 2011 [cited 2017 May 17]. Available from: http://www.cdc.gov/brfss/annual_data/annual_2011.htm
26. Brennan N, Kenney G, Rajan S, Scheuren F, Wang K, Zuckerman S, et al. 1997 NASF technical papers: report no 16: assessing the new federalism an urban institute program to assess changing social policies. Washington: Urban Institute; 2001.
27. Link MW et al. Reaching the U.S. cell phone generation: comparison of cell phone survey results with an ongoing landline telephone survey. *Public Opin Q* [Internet]. 2007 [cited 2013 Nov 24];71(5):814-39. Available from: <http://www.bebr.ufl.edu/files/Reaching%20the%20US%20Cell%20Phone%20Generation.pdf>

Abstract

Objective: to introduce the methodology used to calculate post-stratification weights of the 2012 Surveillance System for Risk and Protective Factors for Chronic Diseases by Telephone Survey (Vigitel) and to compare the trends of indicators estimated by cell-by-cell weighting and raking methods. **Methods:** in this panel of cross-sectional studies, the prevalences of smokers, overweight, and intake of fruits and vegetables from 2006 to 2012 were estimated using the cell-by-cell weighting and raking methods. **Results:** there were no differences in time trends of the indicators estimated by both methods, but the prevalence of smokers estimated by the raking method was lower than the estimated by cell-by-cell weighting, whilst the prevalence of fruit and vegetable intake was higher; for overweight, there was no difference between the methods. **Conclusion:** raking method presented higher accuracy of the estimates when compared to cell-by-cell weighting method, proving to be most convenient, although it presents register coverage bias.

Keywords: Health Surveys; Telephone; Bias; Surveillance; Serial Cross-Sectional Studies.

Resumen

Objetivo: presentar la metodología utilizada para la construcción de pesos de post-estratificación del Sistema de Vigilancia de Factores de Riesgo y Protección para enfermedades Crónicas por Encuesta Telefónica (Vigitel) 2012 y comparar una tendencia de indicadores estimados por métodos de rake y célula. **Métodos:** en este panel de estudios transversales, estimamos las prevalencias de fumadores, exceso de peso, consumo de frutas, verduras y legumbres, de 2006 a 2012, por métodos célula y rake. **Resultados:** no hubo diferencias en las tendencias temporales de los indicadores; sin embargo las prevalencias de fumadores estimadas por el método rake fueron inferiores a las estimadas por célula, mientras que las prevalencias de consumo de frutas y legumbres fueron superiores; para el exceso de peso, no hubo diferencia entre los métodos. **Conclusión:** método rake tuvo mayor precisión en las estimaciones cuando comparado con célula, mostrando más ventajas, aunque presenta vicio de cobertura de registro.

Palabras-clave: Encuestas Epidemiológicas; Teléfono; Sesgo (Epidemiología); Vigilancia; Estudios Transversales Seriales.

Recebido em 30/11/2016
Aprovado em 24/04/2017