
Notas técnicas

População de estudo

A população-alvo da pesquisa foi formada por escolares do 9º ano do ensino fundamental (antiga 8ª série) de escolas públicas ou privadas das capitais dos estados brasileiros e do Distrito Federal. O cadastro utilizado, para a seleção da amostra pesquisada, foi formado pelas escolas de ensino fundamental listadas pelo Censo Escolar 2007, realizado pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira, do Ministério da Educação - INEP/MEC, que informaram possuir turmas de 9º ano do ensino fundamental.

A escolha do 9º ano do ensino fundamental teve como justificativa o mínimo de escolarização necessária para responder questionário autoaplicável e também a proximidade da idade de referência preconizada pela Organização Mundial da Saúde - OMS, que é de 13 a 15 anos.

Foram excluídas as escolas com menos de 15 escolares na série escolhida, pois, embora representassem cerca de 10% das escolas, totalizavam menos de 1% do total de escolares. Considerou-se, portanto, que a seleção de uma ou mais escolas desse subconjunto representaria grande esforço de coleta para pouco retorno em termos de novos dados.

Dessa maneira, o cadastro de seleção da amostra foi constituído por 6 780 escolas que informaram possuir turmas do 9º ano do ensino fundamental.

Aspectos de amostragem

Uma decisão importante em qualquer planejamento amostral é a definição do tamanho da amostra a ser selecionada, por ser o momento em que os objetivos da pesquisa têm que ser compatibilizados com os recursos disponíveis para implementação da pesquisa. No presente inquérito optou-se por especificar os principais alvos de inferência e a margem de erro aceitável na estimação desta(s) quantidade(s), e então, partindo destes parâmetros e de algum conhecimento sobre o tamanho da população e a estrutura de variação na população das variáveis de interesse, determinar o tamanho mínimo de amostra que satisfaça aos requisitos de precisão estabelecidos. Os parâmetros usados foram definidos a priori pelo Grupo de Trabalho da Secretaria de Vigilância em Saúde, do Ministério da Saúde - SVS/MS.

No presente caso, o que se desejou foi uma amostra de escolares que estivessem cursando, em 2009, o 9º ano do ensino fundamental, em escolas públicas ou privadas, para que estes respondessem a um questionário sobre alguns fatores comportamentais de risco e de proteção à saúde.

Como o cadastro de todos os escolares matriculados no 9º ano do ensino fundamental no conjunto das escolas brasileiras disponível era referente ao ano de 2007, o que impossibilitou a seleção direta de uma amostra de escolares, o que se propôs para a pesquisa em questão foi um plano com amostragem de conglomerados em dois estágios, onde as Unidades Primárias de Amostragem - UPA foram as escolas e as Unidades Secundárias de Amostragem - USA foram as turmas do 9º ano do ensino fundamental das escolas selecionadas. A amostra de escolares foi formada, portanto, por todos os escolares das turmas USA selecionadas na amostra de escolas UPA.

As escolas UPA foram estratificadas, inicialmente, levando-se em conta sua localização geográfica e dependência administrativa, de tal modo que cada estrato geográfico correspondeu a um domínio de interesse para a divulgação de resultados da pesquisa. No caso, a estratificação geográfica correspondeu às capitais das Unidades da Federação e o Distrito Federal, totalizando 27 estratos. Em cada um destes, as escolas com classes do 9º ano do ensino fundamental foram agrupadas em escolas privadas ou públicas (federais, estaduais ou municipais).

O tamanho da amostra foi calculado para fornecer estimativas de proporções (ou prevalências) de algumas características de interesse, em cada um dos estratos geográficos, com um erro máximo de 0,03 em valor absoluto no nível de confiança de 95%. Para garantir que isto ocorra, foi dimensionada a amostra, considerando que a prevalência (proporção) é da ordem de 0,5 (ou 50%), pois para proporções desse valor, a variância dos estimadores amostrais é máxima. Os estratos formados pelo cruzamento dos estratos geográficos com a dependência administrativa das escolas foram utilizados apenas para alocação da amostra, de maneira a garantir a presença de escolas públicas e privadas na amostra, de forma proporcional a sua existência no cadastro de seleção.

Em cada estrato geográfico, a amostra foi obtida em dois estágios. Primeiro, foram selecionadas escolas, por meio de método de seleção com probabilidades proporcionais ao tamanho. A medida de tamanho considerada na seleção das escolas foi o número total de turmas do 9º ano do ensino fundamental de cada escola, conforme o cadastro de escolas do Censo Escolar 2007, realizado pelo INEP/MEC. Cada uma

das escolas selecionadas nesse primeiro estágio foi visitada para construção de uma lista atualizada de turmas do 9º ano do ensino fundamental existentes em 2009. Após a obtenção dessas listas, foram selecionadas as turmas USA do 9º ano do ensino fundamental a serem efetivamente pesquisadas em cada uma das escolas selecionadas no primeiro estágio.

Foi escolhida uma turma em cada escola selecionada que tivesse uma ou duas turmas do 9º ano do ensino fundamental, e duas turmas em cada escola com três ou mais turmas do 9º ano do ensino fundamental. Em cada uma das turmas do 9º ano do ensino fundamental selecionadas, todos os escolares responderam ao questionário da pesquisa, eliminando-se, desta maneira, a necessidade de um terceiro estágio de seleção (seleção de escolares dentro da turma selecionada), o que aumentaria as dificuldades operacionais da pesquisa sem correspondente ganho de precisão.

Como sabido, o emprego de planos amostrais conglomerados geralmente resulta em redução de custos para amostras de igual tamanho total em comparação com a Amostragem Aleatória Simples - AAS (*AAS*), por concentrar a amostra nos conglomerados selecionados, reduzindo seu espalhamento geográfico. Por outro lado, o impacto sobre a precisão costuma ser negativo, no sentido de que amostras conglomeradas de igual tamanho que uma *AAS* leva a estimadores com maior variância.

Uma medida usual do impacto do emprego de amostragem conglomerada (ou de amostragem usando planos complexos, em geral) é o chamado Efeito do Plano Amostral - EPA (*epa*), definido como a razão entre a variância do estimador sob o plano conglomerado dividido pela variância do estimador sob *AAS* de igual tamanho.

Portanto, para estimar uma proporção da ordem de 50%, com uma margem de erro k e nível de confiança de 95%, pode-se estimar um tamanho (em número de escolares) para um plano amostral conglomerado em dois estágios e seleção com probabilidades proporcionais a uma medida de tamanho pela fórmula:

$$n_{PPT_2} = epa \times \frac{1,96^2}{4 k^2} \cdot \frac{1}{1 + \frac{1,96^2}{N 4 k^2}},$$

onde n_{PPT_2} é o tamanho da amostra de escolares num determinado estrato geográfico, N é o total de escolares no cadastro nesse estrato geográfico e *epa* é uma estimativa do efeito de conglomeração, pelo fato de se utilizar uma amostragem de conglomerados ao invés de uma amostra aleatória simples de escolares.

Os valores de *epa* para proporções em variáveis socioeconômicas da pesquisa do Sistema Nacional de Avaliação da Educação Básica - SAEB-2003 foram analisados e decidiu-se utilizar o terceiro quartil destes valores em cada estrato de interesse (capitais e Distrito Federal) como fator de ajuste para o dimensionamento da amostra da pesquisa aqui considerada. Desta forma, assegura-se que cerca de 75% das variáveis teriam estimativas com precisão igual, ou melhor, que aquela especificada para dimensionamento da amostra. Exemplificando, se o valor de *epa* for igual a 3,5 num certo estrato de interesse, considerando uma margem de erro de 0,03, o número de escolares na amostra desse estrato é dado por:

$$n_{PPT2} = epa \times n_{AAS} = 3,5 \times 1.068 = 3.738$$

onde n_{AAS} seria o tamanho de uma amostra aleatória simples de escolares para satisfazer as condições exigidas de precisão. Vê-se que o efeito de conglomeração funciona como um fator de correção para o tamanho da amostra no caso de se optar por uma amostra de conglomerados.

Para determinar o tamanho da amostra no segundo estágio, ou seja, o número de turmas selecionadas em cada estrato, basta dividir n_{PPT2} pelo número médio de escolares por turma do 9º ano do ensino fundamental, conforme obtido do cadastro de seleção da amostra de escolas. Supondo que esse número médio, para o estrato do exemplo acima, fosse de 30 escolares por turma, seria preciso uma amostra de aproximadamente 125 turmas para poder obter os 3 738 escolares necessários. O número de escolas do primeiro estágio foi obtido dividindo-se o número de turmas da amostra pelo número médio de turmas das escolas do cadastro em cada estrato.

Tabela 1 - Tamanhos de amostra planejados e coletados, por estágio de seleção, segundo os municípios das capitais e Distrito Federal - 2009

Municípios das capitais e Distrito Federal	Planejado			Coletado				
	Escolas	Turmas	Escolares matri- culados (1)	Escolas	Turmas	Escolares		
						Matri- culados (2)	Fre- quentes (2)	Presentes no dia da pesquisa
Total	1 507	2 270	72 596	1 453	2 175	72 782	68 735	63 411
Porto Velho	56	86	2 363	53	76	2 361	2 123	2 120
Rio Branco	46	72	2 495	43	67	2 256	2 256	1 967
Manaus	42	70	2 391	38	59	2 323	2 039	1 977
Boa Vista	49	82	2 340	48	80	2 316	2 029	1 997
Belém	46	72	2 440	46	73	2 678	2 626	2 189
Macapá	58	88	2 988	56	84	2 813	2 805	2 498
Palmas	41	54	2 000	41	56	1 843	1 627	1 615
São Luis	66	92	3 058	65	79	3 092	2 711	2 670
Teresina	60	82	2 604	59	78	2 582	2 276	2 276
Fortaleza	52	76	2 575	46	72	2 588	2 584	2 340
Natal	62	90	2 853	61	86	2 980	2 933	2 597
João Pessoa	70	98	3 061	67	87	2 768	2 356	2 321
Recife	56	92	3 336	52	79	3 044	3 044	2 561
Maceió	46	70	2 334	40	59	2 304	1 985	1 948
Aracaju	64	88	2 777	62	82	2 787	2 614	2 344
Salvador	48	78	2 466	45	78	2 567	2 216	2 198
Belo Horizonte	68	114	3 467	65	111	3 433	3 115	3 105
Vitória	64	92	2 922	61	83	2 519	2 489	2 259
Rio de Janeiro	56	84	2 866	56	91	3 397	3 257	2 984
São Paulo	52	88	2 835	52	92	3 178	3 086	2 680
Curitiba	44	80	2 578	44	80	2 684	2 679	2 397
Florianópolis	67	88	2 519	66	88	2 553	2 544	2 225
Porto Alegre	52	74	2 001	52	67	1 934	1 873	1 727
Campo Grande	60	78	2 180	58	76	2 420	2 151	2 141
Cuiabá	50	72	2 029	50	76	2 285	2 243	2 014
Goiânia	76	114	3 856	73	112	3 727	3 727	3 291
Distrito Federal	56	96	3 262	54	94	3 350	3 347	2 970

Fonte: INEP, Censo Escolar 2007.

(1) Censo Escolar 2007. (2) Fornecido pelas escolas selecionadas.

A Tabela 1 mostra os tamanhos de amostra calculados (planejados) e coletados por estágio de seleção (escolas, turmas e escolares) para as capitais e o Distrito Federal. Lembre-se que o cadastro de seleção da amostra foi composto pelas escolas listadas pelo Censo Escolar 2007, sendo a coleta de dados realizada entre março e junho de 2009. As informações sobre o número de escolares matriculados e escolares que frequentam regularmente as aulas nas turmas selecionadas do 9º ano do ensino fundamental das escolas foram fornecidas pelas escolas no momento da pesquisa.

Do total de escolares presentes na data da pesquisa, nas turmas selecionadas, 501 escolares se negaram a participar, restando 62 910 escolares, que efetivamente preencheram o questionário. Para a geração das tabelas apresentadas neste volume, decidiu-se utilizar apenas as informações dos escolares que concordaram em participar da pesquisa e que responderam à variável sexo, totalizando 60 973 escolares.

Os pesos amostrais foram calculados de maneira a representar os escolares matriculados no 9º ano do ensino fundamental que frequentam regularmente as aulas. Como no dia da pesquisa eventualmente alguns escolares deixaram de ir à escola e os que responderam à pesquisa podem não ter informado o sexo, o peso do escolar i , da turma j , da escola k , para um determinado estrato (município da capital ou Distrito Federal), foi calculado por:

$$w_{ijk} = w_k \frac{T_k F_{jk} R_{jk}}{t_k R_{jk} S_{jk}} = w_k \frac{T_k F_{jk}}{t_k S_{jk}}, \forall i = 1, 2, \dots, S_{jk}$$

onde:

w_k é o peso da escola k , dado pelo inverso de sua probabilidade de seleção;

T_k é o número total de turmas do 9º ano do ensino fundamental na escola k ;

t_k é o número de turmas do 9º ano do ensino fundamental selecionadas na escola k ($t_k=1$ se $T_k < 3$ ou $t_k=2$ caso contrário);

F_{jk} é o número total de escolares que frequentam regularmente as aulas na turma j da escola k ;

R_{jk} é o número total de escolares que frequentam regularmente as aulas e responderam à pesquisa, na turma j da escola k ; e

S_{jk} é o número total de escolares que frequentam regularmente as aulas, que responderam à pesquisa e informaram o sexo, na turma j da escola k .

Nesta publicação, são apresentadas as estimativas para várias características populacionais, para cada estrato geográfico considerado, e os respectivos intervalos de confiança estimados, considerando um nível de confiança de 95%.

As estimativas de variância utilizadas para construir os intervalos de confiança foram calculadas pelo método do Conglomerado Primário¹.

Cabe ainda ressaltar que, nos resultados tabulados, o percentual de não informados foi distribuído proporcionalmente entre os respondentes.

¹Para maiores detalhes ver: PESSOA, D. G. C.; SILVA, P. L. do N. Análise de dados amostrais complexos. In: SIMPÓSIO NACIONAL DE PROBABILIDADE E ESTATÍSTICA, 13., 1998, Caxambu. *Anais...* São Paulo: Associação Brasileira de Estatística, 1998. 187 p.

Coleta dos dados

A concepção desta pesquisa e os instrumentos desenvolvidos decorreram das atividades do grupo de trabalho criado pela Coordenação Geral de Doenças e Agravos Não Transmissíveis - CGDANT, da Secretaria de Vigilância em Saúde, composto por pesquisadores e técnicos com experiência reconhecida na área de pesquisa em escolares e técnicos do Ministério da Saúde, Ministério da Educação e do Instituto Brasileiro de Geografia e Estatística - IBGE. O questionário para coleta de dados foi baseado nos instrumentos utilizados no Global School-based Student Health Survey/Organização Mundial da Saúde - GSHS/OMS, no Youth Risk Behavior Surveillance System/Centers for Disease Control and Prevention - Yrbss/CDC, no Estudo sobre a Condição de Saúde e Nutrição dos Escolares da Cidade do Rio de Janeiro, no Inquérito de Tabagismo em Escolares - VIGESCOLA e nos questionários usados nos seguimentos das coortes de nascimento da Universidade Federal de Pelotas.

A pesquisa foi realizada utilizando o microcomputador de mão, Personal Digital Assistant - PDA, no qual foi inserido o questionário estruturado autoaplicável com módulos temáticos que variam em número de perguntas. O questionário concebido permitirá a inclusão de outras questões relevantes no futuro, respeitando a mesma estruturação e forma de abordagem. Os assuntos contemplados nos questionários foram: características sociodemográficas, alimentação, imagem corporal, atividade física, tabagismo, consumo de álcool e outras drogas, saúde bucal, comportamento sexual, violência, acidentes, segurança, antropometria e apreciação geral do questionário.

Foram realizados pré-testes para o uso do PDA com escolares do 9º ano do ensino fundamental em oito escolas (públicas e privadas) nos Municípios de Mesquita (Rio de Janeiro), Rio de Janeiro (Rio de Janeiro), Belém (Pará), Recife (Pernambuco) e Luziânia (Goiás), para avaliar a compreensão, adequação dos termos e expressões, abordagem da turma e o tempo para as respostas. Os resultados dos pré-testes mostraram boa aceitação, habilidade no uso do equipamento e aceitação do escolar sobre a pesquisa e PDA.

O questionário que foi a campo não tinha críticas que levasse a saltos entre perguntas. A observação dessa característica é importante visto que ela influenciou na constituição do plano tabular.

A coleta dos dados foi realizada entre os meses de março e junho de 2009.

Aspectos éticos

Em resposta à solicitação da Secretaria de Vigilância em Saúde, do Ministério da Saúde - SVS/MS, através do Parecer de Emenda nº 005 de 10 de junho de 2009, registrado na Comissão Nacional de Ética em Pesquisa - CONEP, sob o nº 11.537 se manifestou pela aprovação do projeto de pesquisa proposto e considerou: "Está claramente delineado e respeita, em princípio, as normas e diretrizes previstas em âmbitos internacional e nacional para pesquisas envolvendo sujeitos humanos, em particular, adolescentes".

Dessa forma, medidas foram tomadas para proteger o adolescente e deixá-lo confortável para responder à pesquisa. A participação foi voluntária, o escolar tinha a possibilidade de deixar de responder qualquer pergunta ou todo o questionário. Todas as informações do escolar foram confidenciais. A escola também não foi identificada.

A realização da pesquisa foi precedida do contato com as Secretarias Estaduais e Municipais de Saúde e Educação e com a direção das escolas selecionadas em cada município.