

Pesquisa Nacional de Saúde do Escolar 2015

Nota metodológica n. 01

**Informações para utilização dos microdados
Amostras 1 e 2**

Versão - 14 de maio de 2018

Presidente da República
Michel Miguel Elias Temer Lulia

Ministro do Planejamento, Desenvolvimento e Gestão
Esteves Pedro Colnago Júnior

**INSTITUTO BRASILEIRO
DE GEOGRAFIA E
ESTATÍSTICA - IBGE**

Presidente
Roberto Luís Olinto Ramos

Diretor-Executivo
Fernando José de Araújo Abrantes

ÓRGÃOS ESPECÍFICOS SINGULARES

Diretoria de Pesquisas
Claudio Dutra Crespo

Diretoria de Geociências
João Bosco de Azevedo

Diretoria de Informática
José Sant`Anna Bevilaqua

Centro de Documentação e Disseminação de Informações
David Wu Tai

Escola Nacional de Ciências Estatísticas
Maysa Sacramento de Magalhães

UNIDADE RESPONSÁVEL

Diretoria de Pesquisas

Coordenação de População e Indicadores Sociais
Barbara Cobo Soares

Ministério do Planejamento, Desenvolvimento e Gestão
Instituto Brasileiro de Geografia e Estatística - IBGE
Diretoria de Pesquisas
Coordenação de População e Indicadores Sociais

Pesquisa Nacional de Saúde do Escolar 2015

Nota metodológica n. 01

**Informações para utilização dos microdados
Amostras 1 e 2**

Versão - 14 de maio de 2018



Rio de Janeiro

2018

A presente Nota metodológica tem por objetivo fornecer ao usuário orientações para a utilização dos microdados da Pesquisa Nacional de Saúde do Escolar (PeNSE 2015).

Encontram-se documentados aqui temas não abordados na publicação da PeNSE, que refletem dúvidas decorrentes da manipulação dos dados (no formato de microdados e outras ferramentas de disseminação de resultados) para a produção de estimativas sobre a população de estudo.

Os temas abordados contemplam assuntos de interesse dos usuários que entraram em contato com o IBGE, por meio dos diversos canais de atendimento. São eles:

- 1) Documentos disponibilizados.
- 2) Dicionário de dados.
- 3) Comparações com edições anteriores.
- 4) Cálculo de estimativas com intervalos de confiança.
- 5) Interpretação dos dados.
- 6) Divulgação dos resultados.
- 7) Comparações da Amostra 2 com o GSHS.
- 8) Antropometria referida (justificativas para não disponibilização de dados).

1) Documentos disponibilizados:

Estão disponíveis na página da PeNSE/2015¹ (<https://www.ibge.gov.br/estatisticas-novoportal/sociais/saude/9134-pesquisa-nacional-de-saude-do-escolar.html?=&t=downloads>) duas notas de especificação de tabelas e indicadores divulgados, os questionários da escola e do aluno, os microdados (amostras 1 e 2), esta Nota metodológica de utilização dos dados e os dicionários de dados (amostras 1 e 2). Cabe ressaltar que os questionários e manual de instrução também estão disponíveis na seção '*Conceitos e métodos*' (<https://www.ibge.gov.br/estatisticas-novoportal/sociais/saude/9134-pesquisa-nacional-de-saude-do-escolar.html?=&t=conceitos-e-metodos>) .

As tabelas disponibilizadas² foram produzidas seguindo especificações constantes nos documentos abaixo relacionados, disponíveis no mesmo local:

- *Nota metodológica – especificações de tabelas e indicadores - Amostra 1.*
- *Nota metodológica - especificações de tabelas e indicadores - Amostra 2.*

Adicionalmente, no diretório de microdados da PeNSE 2015, estão disponíveis para download os arquivos abaixo especificados, nas extensões 'sas', '.csv' e '.xlsx'.

- *PENSE AMOSTRA1 ALUNO*: quesitos do questionário do aluno para Amostra 1.
- *PENSE AMOSTRA1 ALUNOESCOLA*: quesitos do questionário do aluno relacionados aos da escola para Amostra 1.
- *PENSE AMOSTRA1 ESCOLA*: quesitos do questionário da escola para Amostra 1.
- *PENSE AMOSTRA2ALUNO*: quesitos do questionário do aluno para Amostra 2.
- *PENSE_AMOSTRA2_ALUNOESCOLA*: quesitos do questionário do aluno associados aos da escola para a Amostra 2.
- *PENSE AMOSTRA2 ESCOLA*: quesitos do questionário da escola para a Amostra 2.

Os arquivos supracitados foram agrupados em duas pastas compactadas:

¹ A Pense/2015 é composta por dois planos amostrais distintos, planejados para fornecer estimativas representativas dos seguintes grupos populacionais: 1) escolares frequentando o 9º ano do ensino fundamental em escolas públicas e privadas de todo o território nacional (*Amostra 1*) e 2) escolares de 13 a 17 anos de idade frequentando desde o 6º ano do ensino fundamental até a 3ª série do ensino médio em escolas públicas e privadas de todo o território nacional (*Amostra 2*). Para maiores informações, consulte a seção "*Notas técnicas*" da publicação da PeNSE/2015.

² Estão disponíveis, para a Amostra 1, tabelas de indicadores relativos a características da escola e dos alunos. Para a Amostra 2, foram disponibilizados apenas tabelas de indicadores relativos a características dos alunos.

'PENSE_2015_AMOSTRA1.zip' e 'PENSE_2015_AMOSTRA2.zip'. Além dos arquivos supracitados, cada pasta contém o respectivo dicionário de dados.

Os microdados contém exclusivamente observações provenientes de **questionários válidos**³. Cabe ressaltar a existência de diferenças nos quantitativos de **questionários coletados**⁴ e válidos para as amostras 1 e 2, apresentados no Quadro 1. Conforme consta na seção '*Notas Técnicas*' da publicação da PeNSE (*páginas 23 e 28*), durante o processo de cálculo e ajuste dos pesos amostrais, foi atribuído peso zero aos questionários inválidos.

Quadro 1: **Quantitativos de questionários coletados e válidos nas amostras 1 e 2 PeNSE/2015**

Total de questionários	Amostra 1	Amostra 2
Coletados	102.301	16.608
Válidos (disponibilizados nos microdados)	102.072	16.556

Fonte: IBGE, Diretoria de Pesquisas, Coordenação de População e Indicadores Sociais, Pesquisa Nacional de Saúde do Escolar, 2015.

2) Dicionário de dados:

Em 2015 foram introduzidos 'saltos'⁵ no questionário da PeNSE com o objetivo de diminuir seu tempo de aplicação e minimizar a ocorrência de respostas inconsistentes. Adicionalmente, nas turmas selecionadas para a Amostra 2, os 'saltos' restringiram o acesso de escolares menores de 13 anos de idade aos quesitos dos seguintes blocos temáticos: uso de cigarro, bebidas alcoólicas, drogas ilícitas, saúde sexual e reprodutiva.

No dicionário de dados, os 'saltos' do questionário receberam codificação '**-1**' e a descrição **Pulo no questionário**'. Adicionalmente, cabe destacar ainda a ocorrência, exclusivamente no questionário da escola, de observações codificadas como '**Dados**

³ Foram considerados **válidos**, os questionários nos quais houve registro, pelo escolar, dos seguintes quesitos: 1) Concordância em participar da pesquisa (variável VB00004); 2) Sexo (variável VB01001); e 3) Idade (variável VB01003). Para maiores informações, consulte a seção "**Notas técnicas**" da publicação da PeNSE/2015.

⁴ Foram considerados **coletados**, os questionários nos quais houve registro de resposta para, ao menos, uma pergunta do questionário, feito pelo escolar das turmas selecionadas e que aceitou receber o *smartphone* na data de coleta. Para maiores informações, consulte a seção "**Notas técnicas**" da publicação da PeNSE/2015.

⁵ O '**salto**' no questionário é caracterizado pela existência de uma pergunta principal cuja opção de resposta escolhida condicionará a possibilidade ou não de responder um ou mais quesitos posteriores.

não disponíveis (-2)'. Esta descrição foi atribuída a variáveis da escola cujos dados não passaram no processo de crítica/validação, não sendo, portanto, disponibilizados.

3) Comparações com edições anteriores:

A PeNSE 2015 passa a compor, em conjunto com as edições de 2009 e 2012, a série histórica de indicadores de saúde sobre o perfil de escolares matriculados e frequentando regularmente escolas das redes pública e privada no território brasileiro. A ampliação da representatividade da amostra e da abrangência geográfica de disponibilização dos resultados (Quadro 2), implicam em cuidado nas comparações entre as três edições.

Quadro 2: Abrangência geográfica da Pesquisa Nacional de Saúde do Escolar 2009 a 2015

Abrangência Geográfica	2009	2012	2015	
			Amostra 1	Amostra 2
Brasil		X	X	X
Grandes Regiões		X	X	X
Unidades da Federação			X	
Municípios das Capitais	X	X	X	

Fonte: IBGE, Diretoria de Pesquisas, Coordenação de População e Indicadores Sociais, Pesquisa Nacional de Saúde do Escolar, 2015.

Ou seja, comparações na série histórica entre as três edições da PeNSE são possíveis apenas para Municípios das Capitais. Para os anos de 2012 e 2015, são comparáveis resultados de Brasil e Grandes Regiões. Cabe destacar que, no que se refere a **2015**, **comparações** eventualmente realizadas devem utilizar **exclusivamente** dados da **Amostra 1**, já que a Amostra 2 (escolares de 13 a 17 anos) foi realizada na última edição.

Ainda no tocante à comparabilidade, outra recomendação refere-se ao enunciado das perguntas do questionário: eventuais alterações na redação da pergunta e/ou opções de respostas podem inviabilizar comparações de resultados das diferentes edições.

Para facilitar a utilização dos dados, variáveis correspondentes a quesitos do questionário de 2015, cuja redação é igual a 2012, receberam a mesma codificação. Novos quesitos inseridos receberam numeração imediatamente posterior àquela disponível em 2012, dentro do bloco temático correspondente.

As variáveis cujos quesitos sofreram alguma modificação na redação da pergunta e/ou das opções de resposta, em relação a 2012, foram sinalizadas com a letra "a". Por exemplo, a variável VB03011a⁶ corresponde a um quesito existente no questionário de 2012, mas que sofreu alguma modificação.

As variáveis VB03001a1, VB03001a2, VB03002a1 e VB03002a2, do bloco temático '*Atividade Física*', receberam, além da letra 'a' a numeração '1' ou '2' para indicar que esses quesitos de 2015 são originados de apenas uma pergunta do questionário de 2012. Ou seja, as perguntas 3001 e 3002, de 2012, tiveram sua redação alterada e foram desmembradas em duas (3001a1, 3001a2, 3002a1, 3002a2) no questionário de 2015.

4) Cálculo de estimativas com intervalos de confiança.

Para o cálculo de estimativas, deve-se realizar a expansão da amostra, utilizando a variável de peso amostral. No cálculo dos intervalos de confiança, além do peso, devem ser utilizadas as informações do desenho amostral, unidade primária de amostragem (UPA) e estrato geográfico. O Quadro 3 descreve os nomes das variáveis (nos microdados) associadas às informações supracitadas.

⁶ A variável **VB03011a** refere-se, no questionário da edição 2015, ao quesito '*NOS ÚLTIMOS 7 DIAS, em quantos dias você fez atividade física por pelo menos 60 minutos (1 hora) por dia? (Some todo o tempo que você gastou em qualquer tipo de atividade física, EM CADA DIA)*'. Em 2012, a variável **B03011** correspondia, no instrumento de coleta, ao enunciado '*NOS ÚLTIMOS 7 DIAS, em quantos dias você fez atividades física por pelo menos 60 minutos (uma hora) por dia? (Some todo o tempo que você gastou em qualquer tipo de atividade física)*'.

Quadro 3: Variáveis utilizadas nos cálculos de estimativas e intervalos de confiança PeNSE - 2009 a 2015.

Informação	Nome da variável		
	2009	2012	2015
Unidade Primária de Amostragem (UPA)	ID (Escola)	UPA	ESCOLA
Estrato Geográfico	CAPITAL	ESTRATOGE0	ESTRATOGEOREG
Peso	PESO_AJU_SEX0	PESO	PESO

Fonte: IBGE, Diretoria de Pesquisas, Coordenação de População e Indicadores Sociais, Pesquisa nacional de Saúde do Escolar, 2015.

5) Interpretação dos dados:

A interpretação dos dados da PeNSE deve levar em consideração a natureza amostral da pesquisa. Os indicadores produzidos a partir da amostra são estimativas que devem ser interpretadas acompanhadas de seus respectivos intervalos de confiança (espera-se que esses intervalos incluam o resultado do fenômeno, na população, com o nível de confiança especificado de 95%). Outra implicação do uso de resultados de pesquisas amostrais é a relação inversamente proporcional entre precisão da estimativa e extensão do intervalo de confiança: à medida que o intervalo de confiança aumenta, a precisão da estimativa diminui.

Na prática, para a definição do tamanho amostral, foi selecionado um conjunto de indicadores para os quais se desejava uma precisão especificada. Outros indicadores/níveis de desagregação podem apresentar intervalos de confiança muito amplos, sendo bastante imprecisos, uma vez que seria inviável obter uma amostra com tamanho suficiente para fornecer estimativas com precisão adequada para todos os indicadores e níveis de desagregação desejados.

Adicionalmente, a delimitação da população-alvo da pesquisa (Amostra 1: escolares matriculados e frequentando regularmente o 9º ano do Ensino Fundamental; Amostra 2: escolares de 13 a 17 anos matriculados e frequentando escolas) e as características da operação de campo implicam algumas considerações na interpretação dos resultados da PeNSE:

- Os diretores/responsáveis pelas escolas selecionadas, para ambas as amostras,

responderam um questionário sobre infraestrutura do ambiente escolar e políticas relacionadas à saúde desenvolvidas na escola.

- Como se trata de uma amostra planejada para fornecer estimativas para escolares, os indicadores produzidos a partir do **questionário da escola**, fornecem **resultados** com a seguinte interpretação: " **Alunos que estudam em escolas com determinada característica...**" A PeNSE ***não é uma amostra de escolas***, mas sim, de alunos, que, por uma questão operacional, foram selecionados em escolas.
- No caso da Amostra 2, embora a população de interesse seja de escolares de 13 a 17 anos, a estratégia de coleta de dados foi selecionar turmas/séries onde os escolares dessa faixa etária majoritariamente se encontram, aplicar o questionário em todos os alunos da turma e analisar exclusivamente dados de escolares de 13 a 17 anos. Embora estejam disponíveis dados de menores de 13 e maiores de 17 anos, os resultados podem não ser representativos dessas idades, pois a amostra não foi planejada para essa finalidade.

6) Divulgação dos resultados:

Conforme destacado na seção '*Notas técnicas*' da publicação (página 30), além dos microdados, os resultados da PeNSE estão disponibilizados no Banco Multidimensional de Estatísticas (BME) e no Sistema IBGE de Recuperação Automática (SIDRA).

O SIDRA disponibiliza dados com filtros semelhantes aos das tabelas de indicadores. A seleção das variáveis/tabelas no SIDRA requer atenção para que as estimativas obtidas sejam acompanhadas dos respectivos intervalos de confiança.

Figura 1: Tela de seleção de filtros de variáveis - PeNSE/SIDRA

Escolares frequentando o 9º ano do ensino fundamental, total e percentual, em que a localidade onde a escola está situada foi considerada, nos últimos 12 meses, área de risco em termos de violência (roubos, furtos, assaltos, troca de tiros, consumo de drogas, homicídios, etc.), por dependência administrativa da escola	
Unidade Territorial (1)	
	Variável (6)
	④ Dependência administrativa da escola (1)
④ Ano (1)	

Variável [6/6]

☒ Escolares frequentando o 9º ano do ensino fundamental em que, a localidade onde a escola está situada foi considerada, nos últimos 12 meses, área de risco em termos de violência

☒ Escolares frequentando o 9º ano do ensino fundamental em que, a localidade onde a escola está situada foi considerada, nos últimos 12 meses, área de risco em termos de violência

☒ Escolares frequentando o 9º ano do ensino fundamental em que, a localidade onde a escola está situada foi considerada, nos últimos 12 meses, área de risco em termos de violência

☒ Percentual de escolares frequentando o 9º ano do ensino fundamental em que, a localidade onde a escola está situada foi considerada, nos últimos 12 meses, área de risco em termos de violência

☒ Percentual de escolares frequentando o 9º ano do ensino fundamental em que, a localidade onde a escola está situada foi considerada, nos últimos 12 meses, área de risco em termos de violência

☒ Percentual de escolares frequentando o 9º ano do ensino fundamental em que, a localidade onde a escola está situada foi considerada, nos últimos 12 meses, área de risco em termos de violência

Dependência administrativa da escola [1/3]

☒ Total

☐ Pública

☐ Privada

Ano [1/2]

☐ 2015

☒ 2012

Unidade Territorial [1/62]

Níveis territoriais

☒ Brasil [1/1]
 ☐ Grande Região [0/5]
 ☐ Unidade da Federação [0/27] (Ano 2015)
 ☐ Conjunto dos Municípios Capitais / não Capitais das Unidades da Federação [0/2]
 ☐ Município [0/27]

Início

☐ ☐

Fonte: Sistema do IBGE de Recuperação Automática (SIDRA), disponível em: www.ibge.gov.br.

Na Figura 1 é apresentada uma tela do SIDRA, com a variável '*Escolares frequentando o 9º ano do ensino fundamental em que, a localidade onde a escola está situada, foi considerada, nos últimos 12 meses, área de risco em termos de violência (roubos, furtos, assaltos, consumo de drogas, homicídios etc)*'. São disponibilizadas seis opções de seleção de componentes do conjunto estimativa/intervalos de confiança (destacados em vermelho). Estes **não devem ser confundidos com opções de variáveis**, pois tratam do mesmo quesito e se encontram descritos a seguir:

1) Escolares frequentando o 9º ano do ensino fundamental em que, a localidade onde a escola está situada foi considerada, nos últimos 12 meses, área de risco em termos de violência (roubos, furtos, assaltos, consumo de drogas, homicídios etc) **(Pessoas)**.

2) Escolares frequentando o 9º ano do ensino fundamental em que, a localidade onde a escola está situada foi considerada, nos últimos 12 meses, área de risco em termos de violência (roubos, furtos, assaltos, consumo de drogas, homicídios etc), considerando um intervalo de confiança de 95% - **limite inferior (Pessoas)**.

3) Escolares frequentando o 9º ano do ensino fundamental em que, a localidade onde a escola está situada foi considerada, nos últimos 12 meses, área de risco em termos de violência (roubos, furtos, assaltos, consumo de drogas, homicídios etc), considerando um intervalo de confiança de 95% - **limite superior (Pessoas)**.

4) Percentual de escolares frequentando o 9º ano do ensino fundamental em que, a localidade onde a escola está situada foi considerada, nos últimos 12 meses, área de risco em termos de violência (roubos, furtos, assaltos, consumo de drogas, homicídios etc) **(Percentual)**.

5) Percentual de escolares frequentando o 9º ano do ensino fundamental em que, a localidade onde a escola está situada foi considerada, nos últimos 12 meses, área de risco em termos de violência (roubos, furtos, assaltos, consumo de drogas, homicídios etc), considerando um intervalo de confiança de 95% - **limite inferior (Percentual)**.

6) Percentual de escolares frequentando o 9º ano do ensino fundamental em que, a localidade onde a escola está situada foi considerada, nos últimos 12 meses, área de risco em termos de violência (roubos, furtos, assaltos, consumo de drogas, homicídios etc), considerando um intervalo de confiança de 95% - **limite superior (Percentual)**.

Além das estimativas em números absolutos de escolares (identificada como 'pessoas') e dos percentuais com os respectivos intervalos de confiança para cada variável, são permitidas seleções por dependência administrativa da escola, ano de

realização da pesquisa e unidades territoriais.

O BME disponibiliza os dados da PeNSE para tabulações utilizando diversos níveis de desagregação e filtros na construção de indicadores. Todavia, os **resultados não são acompanhados dos respectivos intervalos de confiança** e, portanto, não fornecem a precisão das estimativas.

7) Comparações da Amostra 2 com o GSHS:

Em 2015 foram disponibilizados, pela primeira vez, resultados provenientes de dois planos amostrais distintos, planejados para fornecer, respectivamente, dados representativos de escolares frequentando o 9º ano do ensino fundamental (Amostra 1) e de escolares de 13 a 17 anos de idade (Amostra 2).

A Amostra 2 foi planejada para possibilitar a comparação, considerando as idades dos escolares, com resultados de outros inquéritos nacionais no âmbito do sistema *Global School-based Student Health Survey* - GSHS, desenvolvido pela Organização Mundial da Saúde, com assistência técnica do Centers for Disease Control and Prevention (CDC - Atlanta). O GSHS é aplicado em países da África, América Latina, Ásia, Europa e Pacífico Ocidental e tem como público-alvo os escolares de 13 a 17 anos utilizando a metodologia de questionário autoaplicável.

Para compor o sistema GSHS, o inquérito nacional deve incorporar no instrumento de coleta, ao menos seis dos 10 blocos temáticos que compõem o módulo básico. Cada bloco temático possui entre três e sete perguntas que devem ser reproduzidas sem qualquer modificação na redação. A PeNSE 2015 incorporou em seu questionário, comum às duas amostras, os seguintes blocos temáticos: 1) álcool (seis quesitos); 2) alimentação (sete quesitos); 3) higiene pessoal (quatro quesitos); 4) atividade física (quatro quesitos); 5) fatores de proteção (seis quesitos); 6) comportamento sexual (cinco quesitos). Em 2015, a pergunta sobre ir ou voltar para a escola a pé ou de bicicleta, inserida no bloco '*atividade física*', foi desmembrada em dois quesitos abordando separadamente a ida para a escola e a volta da escola.

Não obstante o supracitado, a comparação direta entre as estimativas produzidas com base nos quesitos comuns ao módulo básico do GSHS deve ser realizada com

restrições. Recomenda-se a leitura prévia e criteriosa dos relatórios finais e dos metadados/documentação do inquérito do país que se deseja fazer comparação a fim de obter informações sobre eventuais parâmetros que afetem a comparabilidade internacional, tais como: tamanho da amostra, desenho amostral, exemplos utilizados na redação das perguntas, opções de respostas utilizadas nas perguntas, filtros utilizados na produção dos indicadores etc.

8) Antropometria referida (justificativas para não disponibilização de dados):

Ao final do preenchimento do questionário da PeNSE, os escolares das amostras 1 e 2 foram convidados a registrar peso e altura autorreferidos. Todavia os quesitos relacionados não foram disponibilizados devido à elevada taxa de não-resposta que afetou a qualidade desses dados.