Rev Saúde Pública 2013;47(Supl 3):1-8
Original Articles
DOI:10.1590/S0034-8910.2013047004362

Nilza Nunes da Silva[I]

Angelo Giuseppe Roncalli[II]

Sampling plan, weighting process and design effects of the Brazilian Oral Health Survey

ABSTRACT

OBJECTIVE: To present aspects of the sampling plan of the Brazilian Oral Health Survey (SBBrasil Project). with theoretical and operational issues that should be taken into account in the primary data analyses.

METHODS: The studied population was composed of five demographic groups from urban areas of Brazil in 2010. Two and three stage cluster sampling was used. adopting different primary units. Sample weighting and design effects (deff) were used to evaluate sample consistency.

RESULTS: In total. 37.519 individuals were reached. Although the majority of deff estimates were acceptable. some domains showed distortions. The majority (90%) of the samples showed results in concordance with the precision proposed in the sampling plan. The measures to prevent losses and the effects the cluster sampling process in the minimum sample sizes proved to be effective for the deff. which did not exceeded 2. even for results derived from weighting.

CONCLUSIONS: The samples achieved in the SBBrasil 2010 survey were close to the main proposals for accuracy of the design. Some probabilities proved to be unequal among the primary units of the same domain. Users of this database should bear this in mind. introducing sample weighting in calculations of point estimates. standard errors. confidence intervals and design effects.

DESCRIPTORS: Dental Health Surveys. methods. Cluster Sampling. Epidemiologic Research Design.

I Faculdade de Saúde Pública. Universidade de São Paulo. São Paulo. SP. Brasil

II Universidade Federal do Rio Grande do Norte. Natal. RN. Brasil

Correspondence:
Nilza Nunes da Silva
Depto. Epidemiologia
Faculdade de Saúde Pública
Universidade de São Paulo
Av. Dr. Arnaldo. 715 - Cerqueira Cesar
01246-904 São Paulo. SP. Brasil
E-mail: nndsilva@usp.br

Received: 05/16/2012
Approved: 04/18/2013

Article available from: www.scielo.br/rsp

INTRODUCTION

The Brazilian Oral Health Survey (SBBrasil 2010) is a health monitoring strategy which uses primary data to produce information which aids in implementing oral health care policies. It is the second large-scale nationwide oral health survey; a similar. previous one was carried out in 2003. Two other nationwide surveys were carried out in 1986 and 1996. although only in the state capitals and assessing fewer health problems.

The SBBrasil 2010 was planned during 2009 and data collection took place between February and November 2010 in 177 municipalities. including the 27 state capitals. A total of 37.519 interviews and oral examinations were carried out on the age groups recommended by the World Health Organization (five year olds. 12 year olds. 15 to 19 year olds. 35 to 44 year olds and 65 to 74 year olds). The main oral health care problems (dental caries. periodontal disease. malocclusion. fluorosis. trauma and edentulism). as well as socioeconomic data regarding information on use of orthodontic services. self-reported oral morbidities and self-perception of oral health. The final report and the database of original data are available on the Brazilian Ministry of Health General Coordination of Oral Health site.[a]

There were 160 samples distributed according to 32 geographical domains. representing the populations of the aforementioned age groups. resident in the state capitals or in municipalities in the interior of the five regions of Brazil. Obtaining epidemiological information directly from these samples. whether for an age group. a state capital or a municipality in the interior. requires knowledge of the sampling plan. In other words. the inferences made should take into account the method designed for the inclusion. for the selection of a specific individual in the sample from the domain to which they belong. The general model used was cluster sampling in multiple stages. in which the sampling units were selected with probability proportional to the number of residences in each.

The aim of this article was to present aspects of the sampling plan. explaining theoretical and operational issues which should be taken into account when analyzing the primary data.

METHODS

Expected number of interviews and oral examinations

For those aged five and those aged 12. and for the 65 to 74 year old age group the coefficient of variation of ratios was adopted as a measure of accuracy. because most of the health problems consisted of categorical variables. The quantitative decayed. missing and filled teeth index (DMFT) proved to be inadequate as a parameter due to its low mean value and high variability. especially at ages five and 12. The results obtained through the equation $cv(p) = \frac{\sqrt{(p(1-p))/n}}{p}$ varied between 3% and 27% depending on the expected prevalence values for the population. when n = 125. As this was the minimum acceptable number for the domains of the abovementioned groups. it was verified that the absolute values of the standard error were below 5% and never above 18% of the rates of prevalence above 10%. In order to diminish the cluster sampling effect on this accuracy criterion. it was decided to double the number of interviews (deff = 2) and select 250 individuals in each domain.[4]

For the 15 to 19 and the 35 to 44 year old age groups. the sample size was calculated using the expression $n = [(s_x.1.96)/m]^2$. with 1.96 being the value of normal distribution corresponding to the 95% confidence interval estimated for the mean number of decayed. missing and filled teeth (DMFT) in each domain: (m) is the tolerable margin for error inherent to the simple random sampling process; and (s_x) estimates the standard error using data from the 2003 survey. The initial results were corrected to compensate for the response rate effect of 80% and design effect (deff) of 2.

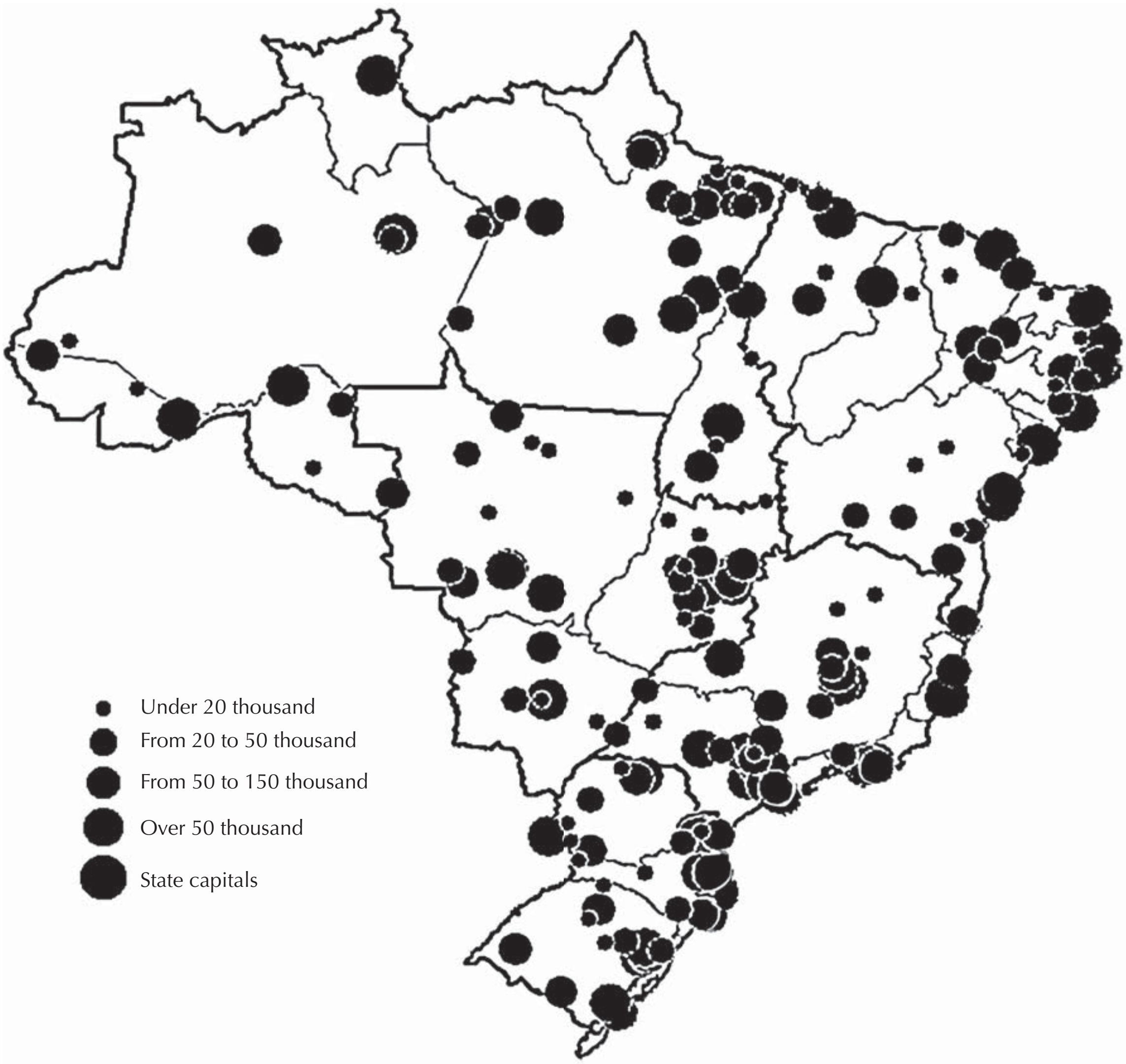
The samples of residences in the 160 domains were calculated using the equation (dom = n/r x 0.9). in which “n” is the minimum number of interviews. determined by the aforementioned criterion of accuracy. and “r” is the density of elements (of each demographic group) per residence. calculated based on data from the 2000 demographic census. The correction of 0.9 aimed to prevent loss of accuracy due to closed or vacant residences and to refusals to take part in the study.

Sampling process

The Figure shows the distribution of the state capitals and the interior municipalities included in the macro-region samples. The method used for selecting the sample followed the general model of cluster sampling in multiple stages. With probability proportional to size (PPS).[1] In the first stage. 30 census tracts were selected for each state capital and 30 municipalities for the interior of each region. These were the primary sampling units (PSUs) which were included when drawing up the files as well as in estimating standard error and confidence intervals.

In the second stage. residences in the census tracts of each capital and two census tracts in the municipalities making

[a] Ministério da Saúde. Coordenação de Saúde Bucal da Secretaria de Assistência à Saúde. Projeto SBBrasil 2010 - Pesquisa Nacional de Saúde Bucal. [cited 2013 Sep 04]. Available from: http://dab.saude.gov.br/cnsb/sbbrasil/index.html



Source: Brazilian Ministry of Health SBBrasil 2010 – Principal Results[3]

Figure: Selected state capitals and municipalities according to population size. SBBrasil 2010.

up the sample from the interior were selected. Each geographical region contained 30 tracts for each capital and 60 for the sample of municipalities in the interior. In the third stage. which only applied to the municipalities in the interior. residences were randomly selected within each of the sectors selected in the previous stage.

In the samples of residences in each demographic domain and each group. all of the elements deemed eligible were interviewed and examined. Therefore. the probability of an individual being selected was the same as the probability of their residence being selected. In the state capitals. the equation $f(c) = \frac{30.D_j}{\Sigma D_j} \cdot \frac{d}{D_j}$ calculates the theoretical probability of a residence being included in the sample for the state capitals. in which is the number of residences in the j[th] census tract and (d) is the number of residences selected within each sector.

For those residing in the interior. where the selection process had three stages. the probability of inclusion was calculated by $f(m) = \frac{30.D_{mj}}{\Sigma D_{mj}} \frac{2.D_{mj,i}}{D_{mj}} \frac{d}{D_{mj,i}}$. with being the number of residences in the municipality (j). and in the i[th] census tract situated in the territory of the municipality (j) selected in the first stage. In the interior. (d) was also the number of residences selected within the census tract.

However. in both equations. the denominator of the last fraction recorded the result of the quick count. conducted in the field. in order to update the data from the 2000 census which was used to select the municipalities or census tracts in the previous stages. The self-weighting of the samples according to the PPS method was abandoned and the equations shown were effectively calculated substituting these terms for their respective values(D'_j) and ($D'_{m.j.i}$) updated in 2010.

Table 1. Number of selected residences according to samples of individuals (n) according to age group and geographic domain. SBBrasil. 2010.

Domain	Age / Age group									
	5 years		12 years		15 to 19 years		35 to 44 years		65 to 74 years	
	n	Domiciles	n	Domiciles	n	Domiciles	n	Domiciles	n	Domiciles
Porto Velho	250	3.238	250	3.561	200	577	487	999	250	4.093
Rio Branco	250	3.376	250	3.441	481	1.379	559	1.251	250	2.884
Manaus	250	3.769	250	3.408	238	681	553	1.151	250	3.238
Boa Vista	250	3.561	250	3.653	200	568	390	825	250	2.993
Belém	250	4.397	250	3.913	200	626	780	1.544	250	2.120
Macapá	250	2.955	250	3.152	200	535	467	1.106	250	3.769
Palmas	250	3.671	250	3.238	212	589	443	919	250	5.124
São Luís	250	4.290	250	4.036	200	622	508	1.133	250	2.855
Teresina	250	4.218	250	4.290	200	642	813	1.842	250	2.662
Fortaleza	250	4.550	250	4.148	200	626	668	1.428	250	2.241
Natal	250	4.692	250	4.416	262	871	390	844	250	2.057
João Pessoa	250	4.972	250	4.442	210	679	502	1.054	250	2.007
Recife	250	5.250	250	4.663	200	699	475	968	250	1.734
Maceió	250	3.890	250	4.036	228	736	502	1.088	250	2.545
Aracaju	250	4.496	250	4.550	200	694	505	1.074	250	2.145
Salvador	250	4.782	250	4.692	200	713	398	815	250	2.368
Belo Horizonte	250	6.156	250	5.865	200	890	457	1.099	250	1.812
Vitória	250	6.113	250	5.987	200	913	476	1.181	250	1.920
Rio de Janeiro	250	6.028	250	6.028	200	966	411	1.014	250	1.489
São Paulo	250	5.637	250	5.749	200	913	415	970	250	1.904
Curitiba	250	6.493	250	5.781	204	882	480	1.122	250	2.113
Florianópolis	250	6.840	250	6.137	200	897	307	742	250	2.036
Porto Alegre	250	6.737	250	6.096	200	951	321	853	250	1.572
Campo Grande	250	5.425	250	5.229	200	779	469	1.139	250	2.237
Cuiabá	250	5.198	250	5.325	200	781	427	1.029	250	2.738
Goiânia	250	5.826	250	5.564	253	1.062	375	872	250	2.309
Brasília	250	5.106	250	4.960	200	772	526	1.157	250	2.913
Interior – North	250	3.073	250	3.289	214	577	597	1.506	250	2.261
Interior – Northeast	250	3.610	250	3.940	235	731	618	1.634	250	1.695
Interior – Southeast	250	5.306	250	5.413	211	883	581	1.485	250	1.623
Interior – South	250	6.021	250	5.605	208	892	546	1.406	250	1.509
Interior – Central-West	250	4.896	250	5.076	256	1.000	547	1.336	250	1.955

Sampling weight and Design effects

The sample weights were calculated by the inverse of the probability equations $(f)^{-1}$ and added to the files of the individuals examined. This meant attributing the data from each element included in the sample to those not included in the same PSU. This mechanism can reduce potential bias due to the disproportionality of the numbers observed in the interviews between PSUs. In theoretical terms. it means affirming that the sampling plan does not follow the principle of self-weighting. according to which the probability of an individual being included in the samples for all the domains. in each demographic group. [3] would be equal and expressed by ($f = n/N$).

The weights (w) were calculated for each primary sampling unit of the sample. including. as seen in the mathematical equations. the terms of the probability of being selected at each stage. Operationally. the results obtained for a PSU were attributed to all of the individuals included and the final file of the data contains this weight for each individual record from which it is composed.

Estimates for the measurements or proportions. standard error and confidence intervals were calculated with and without basic weight. using the “SVY” (survey) module of the Stata program. version 11.2. This application introduces design variables (defining the domains) and basic weights in the statistical process. Estimates of standard error were calculated using the Taylor linearization method. applicable to data from complex sampling plans.[1,2]

Design effects (deffs) were calculated for the estimates of each domain according to geographical region and demographic group. Comparison of these measures calculated with or without basic weight meant the effect of homogeneity and the intra-class impact on the accuracy of the sample weights could be assessed.[3]

The SBBrasil 2010 project was carried out following the standards set by the Declaration of Helsinki and was approved by the *Conselho Nacional de Ética em Pesquisa*. record no. 15.498. 7th January 2010.

RESULTS

The sample was divided into geographic regions. defined by the 27 state capitals and the 150 interior municipalities in the five macro-regions of Brazil (Figure). In total. 1.110 census tracts were selected: 30 for each state capital and 60 for each sample of the municipalities in the interior.

The number of residences selected in order to achieve the minimum number of interviews and oral examinations in

Table 2. Number of interviews achieved according to geographic domains and age groups. SBBrasil. 2010

Domain	Age / Age group									
	5 years		12 years		15 a 19 years		35 a 44 years		65 a 74 years	
	n	%	n	%	n	%	n	%	n	%
Porto Velho	180	72.0	183	73.2	163	81.5	331	68.0	211	84.4
Rio Branco	165	66.0	173	69.2	218	45.3	214	38.3	186	74.4
Manaus	204	81.6	148	59.2	146	61.3	229	41.4	181	72.4
Boa Vista	195	78.0	207	82.8	137	68.5	182	46.7	192	76.8
Belém	306	122.4	261	104.4	159	79.5	496	63.6	262	104.8
Macapá	232	92.8	226	90.4	159	79.5	346	74.1	239	95.6
Palmas	184	73.6	180	72.0	152	71.7	317	71.6	168	67.2
São Luís	128	51.2	105	42.0	110	55.0	89	17.5	159	63.6
Teresina	184	73.6	192	76.8	118	59.0	285	35.1	216	86.4
Fortaleza	234	93.6	190	76.0	113	56.5	369	55.2	255	102.0
Natal	188	75.2	162	64.8	136	51.9	175	44.9	231	92.4
João Pessoa	142	56.8	141	56.4	128	61.0	216	43.0	211	84.4
Recife	270	108.0	198	79.2	83	41.5	147	30.9	225	90.0
Maceió	167	66.8	173	69.2	107	46.9	187	37.3	184	73.6
Aracaju	234	93.6	250	100.0	181	90.5	214	42.4	192	76.8
Salvador	233	93.2	255	102.0	214	107.0	274	68.8	267	106.8
Belo Horizonte	200	80.0	262	104.8	149	74.5	260	56.9	247	98.8
Vitória	205	82.0	213	85.2	117	58.5	155	32.6	173	69.2
Rio de Janeiro	280	112.0	267	106.8	230	115.0	348	84.7	342	136.8
São Paulo	224	89.6	233	93.2	183	91.5	373	89.9	255	102.0
Curitiba	236	94.4	268	107.2	158	77.5	417	86.9	283	113.2
Florianópolis	188	75.2	238	95.2	162	81.0	220	71.7	233	93.2
Porto Alegre	225	90.0	210	84.0	251	125.5	431	134.3	304	121.6
Campo Grande	209	83.6	206	82.4	189	94.5	380	81.0	207	82.8
Cuiabá	118	47.2	156	62.4	79	39.5	159	37.2	172	68.8
Goiânia	259	103.6	269	107.6	197	77.9	250	66.7	240	96.0
Brasília	179	71.6	196	78.4	148	74.0	224	42.6	140	56.0
Interior – North	352	140.8	365	146.0	233	108.9	470	78.7	319	127.6
Interior – Northeast	341	136.4	337	134.8	217	92.3	431	69.7	307	122.8
Interior – Southeast	398	159.2	389	155.6	245	116.1	496	85.4	289	115.6
Interior – South	287	114.8	294	117.6	247	118.8	570	104.4	343	137.2
Interior – Central-West	376	150.4	365	146.0	296	115.6	479	87.6	358	143.2

Table 3. Prevalence estimates. standard errors and design effect with and without weighting for the age groups of 5. 12 and 65 to 74 years old according to study domains. SBBrasil. 2010.

Groups	Prevalence of dental caries (deciduous) 5 years old						Prevalence of dental caries (permanent) 12 years old						Prevalence of bleeding gums 65 to 74 years old					
Weighting	Without			With			Without			With			Without			With		
Domain	prop	SE	deff	prop	SE	deff	prop	v	deff	prop	SE	deff	prop	v	deff	prop	SE	deff
Porto Velho	0.632	0.036	0.927	0.632	0.036	0.926	0.753	0.039	1.341	0.744	0.037	1.203	0.152	0.030	1.439	0.150	0.028	1.275
Rio Branco	0.601	0.034	0.784	0.642	0.038	1.034	0.727	0.029	0.723	0.735	0.026	0.583	0.134	0.029	1.326	0.124	0.031	1.655
Manaus	0.570	0.033	0.946	0.567	0.036	1.069	0.678	0.048	1.541	0.662	0.057	2.082	0.282	0.041	1.525	0.295	0.041	1.430
Boa Vista	0.670	0.038	1.289	0.666	0.037	1.205	0.732	0.023	0.556	0.730	0.022	0.503	0.182	0.026	0.866	0.192	0.026	0.806
Belém	0.567	0.034	1.369	0.536	0.044	2.219	0.637	0.038	1.568	0.644	0.044	2.083	0.182	0.026	0.866	0.324	0.050	3.017
Macapá	0.669	0.034	1.202	0.644	0.038	1.441	0.726	0.038	1.591	0.731	0.043	2.162	0.268	0.043	2.289	0.261	0.040	2.019
Palmas	0.467	0.038	1.054	0.465	0.039	1.141	0.665	0.038	1.150	0.678	0.035	0.962	0.113	0.020	0.679	0.110	0.020	0.686
São Luís	0.428	0.042	1.169	0.407	0.044	1.314	0.685	0.028	0.504	0.684	0.027	0.487	0.199	0.029	1.149	0.202	0.029	1.073
Teresina	0.603	0.051	1.879	0.567	0.054	2.035	0.523	0.034	0.882	0.500	0.041	1.284	0.199	0.029	1.149	0.122	0.034	2.384
Fortaleza	0.418	0.043	1.719	0.427	0.046	2.011	0.529	0.033	0.798	0.527	0.034	0.849	0.176	0.033	1.873	0.166	0.033	2.019
Natal	0.540	0.050	1.907	0.532	0.049	1.847	0.590	0.032	0.689	0.576	0.034	0.746	0.164	0.018	0.530	0.168	0.018	0.515
Joao Pessoa	0.619	0.050	1.479	0.601	0.048	1.305	0.700	0.027	0.467	0.704	0.026	0.458	0.156	0.026	1.080	0.157	0.026	1.064
Recife	0.506	0.042	1.869	0.527	0.046	2.269	0.533	0.044	1.500	0.536	0.033	0.849	0.178	0.044	2.950	0.179	0.042	2.661
Maceió	0.599	0.041	1.137	0.605	0.042	1.229	0.616	0.039	1.116	0.621	0.039	1.041	0.223	0.039	1.607	0.213	0.038	1.541
Aracaju	0.504	0.036	1.212	0.525	0.039	1.397	0.416	0.033	1.184	0.418	0.037	1.425	0.250	0.037	1.407	0.252	0.034	1.181
Salvador	0.434	0.038	1.319	0.435	0.041	1.525	0.416	0.026	0.693	0.409	0.032	1.082	0.329	0.025	0.726	0.345	0.023	0.621
Belo Horizonte	0.545	0.035	0.976	0.546	0.038	1.144	0.439	0.039	1.601	0.436	0.037	1.472	0.178	0.032	1.799	0.184	0.037	2.202
Vitória	0.439	0.044	1.591	0.428	0.042	1.463	0.498	0.044	1.668	0.550	0.047	1.382	0.249	0.041	1.560	0.235	0.041	1.620
Rio de Janeiro	0.291	0.035	1.559	0.290	0.033	1.431	0.478	0.053	2.736	0.494	0.053	2.809	0.146	0.029	2.300	0.110	0.032	3.419
São Paulo	0.411	0.037	1.245	0.418	0.040	1.478	0.476	0.036	1.199	0.477	0.037	1.283	0.345	0.033	1.256	0.344	0.035	1.358
Curitiba	0.553	0.028	0.727	0.562	0.029	0.799	0.549	0.029	0.909	0.553	0.030	0.997	0.159	0.039	3.343	0.167	0.041	3.389
Florianópolis	0.390	0.043	1.444	0.391	0.042	1.399	0.326	0.039	1.718	0.316	0.040	1.777	0.266	0.047	2.590	0.264	0.046	2.544
Porto Alegre	0.400	0.038	1.329	0.397	0.039	1.487	0.519	0.039	1.289	0.512	0.033	0.918	0.329	0.036	1.756	0.355	0.038	1.972
Campo Grande	0.526	0.035	1.037	0.571	0.040	1.368	0.558	0.044	1.633	0.598	0.064	3.484	0.285	0.044	1.993	0.326	0.072	4.830
Cuiabá	0.648	0.051	1.204	0.632	0.077	2.636	0.623	0.023	0.330	0.623	0.025	0.389	0.180	0.041	1.963	0.179	0.044	2.207
Goiânia	0.480	0.029	0.851	0.480	0.029	0.854	0.509	0.052	2.922	0.511	0.052	2.907	0.050	0.016	1.264	0.049	0.016	1.254
Brasília	0.475	0.039	1.109	0.472	0.042	1.252	0.415	0.039	1.245	0.434	0.042	1.425	0.221	0.039	1.239	0.212	0.039	1.325
Interior – North	0.722	0.042	3.028	0.702	0.041	2.778	0.734	0.029	1.534	0.737	0.032	1.391	0.182	0.025	1.333	0.165	0.021	1.020
Interior – Northeast	0.669	0.039	2.199	0.691	0.031	1.427	0.715	0.038	2.264	0.761	0.040	2.892	0.195	0.029	1.686	0.169	0.026	1.455
Interior – Southeast	0.522	0.026	1.061	0.505	0.028	1.179	0.557	0.046	3.294	0.515	0.045	3.161	0.142	0.029	2.046	0.156	0.033	2.379
Interior – South	0.646	0.033	1.342	0.628	0.034	1.351	0.641	0.041	2.161	0.593	0.049	1.714	0.160	0.033	2.726	0.160	0.033	2.807
Interior – Central-West	0.664	0.026	1.134	0.659	0.031	1.565	0.734	0.034	2.176	0.724	0.036	2.343	0.196	0.032	2.305	0.202	0.035	2.736

prop: proportions
SE: standard error
deff: design effect

the domains can be found in Table 1. It can be seen that. in the majority of cases. the results are greater for the groups of five and 12 year olds. who have lower intra-residence density. The only exception is in the North. which showed higher samples for the elderly in Porto Velho. Macapá and Palmas. This important demographic detail should not be overlooked in sampling plans which take the residence as the sampling unit at some stage of the selection. For example. in the state capital. São Paulo. in order to achieve 250 interviews. it was necessary to select 5.637 residences for the first group and almost three times fewer (1.904) for the group aged 65 to 74. This difference is the result of unequal densities. calculated by the ratio of individuals/residence. equal to five children or 15 elderly individuals for each 100 residences.

Identifying and selecting the addresses within each census tract. supervised by the research coordination team. sought to preserve the criteria of accuracy defined in the sampling plan. However. the effective number of interviews and oral examinations achieved in each sample was rarely the same or above the defined minimums (Table 2). Only the samples in the interior of each geographical region were the minimums defined in the sampling plan preserved. achieving at least 70% of the planned interviews in all of the domains.

In the state capitals. despite the process of updating the register of residences in each census tract. circumstance due to the infrastructure and logistics of field work can be associated with the results found. Almost half of the samples in the 35 to 44 year old age group did

Table 4. Estimates for the DMFT (decayed. missing and filled teeth). standard error and design effects for the age groups 15-19 and 35-44 years old according to study domains. SBBrasil. 2010.

Groups	15 to 19 years old						35 to 44 years old					
Weighting	Without			With			Without			With		
Domain	prop	SE	deff	prop	SE	deff	prop	SE	deff	prop	SE	deff
Porto Velho	6.753	0.432	1.636	6.770	0.450	1.742	19.000	0.400	1.300	18.970	0.420	1.440
Rio Branco	4.862	0.267	0.762	4.845	0.261	0.765	19.691	0.465	0.945	19.556	0.641	1.705
Manaus	4.849	0.338	0.954	4.986	0.411	1.252	19.659	0.381	0.804	19.336	0.461	1.239
Boa Vista	5.678	0.611	2.713	5.584	0.516	1.975	18.039	0.742	2.012	17.981	0.713	1.846
Belém	4.878	0.419	1.618	4.852	0.389	1.372	16.168	0.449	2.062	15.869	0.509	2.615
Macapá	4.017	0.569	2.785	3.989	0.484	2.149	12.823	0.735	3.862	12.963	0.704	3.423
Palmas	5.313	0.702	2.532	5.333	0.476	1.420	17.417	0.437	1.317	17.484	0.422	1.248
São Luís	4.697	0.296	0.734	4.609	0.264	0.663	12.261	1.019	2.621	12.441	1.032	2.751
Teresina	4.083	0.393	1.038	4.142	0.392	1.035	15.689	0.393	0.644	15.715	0.393	0.645
Fortaleza	3.193	0.258	0.765	3.186	0.256	0.787	17.271	0.348	0.961	17.096	0.405	1.321
Natal	4.701	0.562	1.619	4.444	0.542	1.255	18.845	0.516	0.952	19.094	0.468	0.795
Joao Pessoa	6.152	0.602	1.791	6.406	0.526	1.254	17.750	0.595	1.897	17.608	0.534	1.567
Recife	3.903	0.652	1.716	4.256	0.707	1.812	15.911	0.927	1.962	15.853	0.989	2.203
Maceió	5.501	0.439	1.005	5.336	0.433	0.996	17.541	0.417	0.705	17.359	0.383	0.583
Aracaju	2.588	0.365	2.604	2.503	0.308	1.931	17.509	0.522	1.479	17.255	0.511	1.401
Salvador	2.095	0.226	1.232	2.208	0.241	1.449	14.408	0.419	1.005	14.101	0.559	1.804
Belo Horizonte	2.332	0.227	1.076	2.367	0.239	1.156	16.253	0.361	0.757	16.354	0.380	0.840
Vitória	2.667	0.327	1.152	2.838	0.334	1.131	15.897	0.541	1.035	15.548	0.684	1.554
Rio de Janeiro	3.304	0.490	2.259	2.886	0.441	2.524	15.383	0.828	3.169	15.472	0.657	2.011
São Paulo	4.209	0.534	2.544	4.208	0.511	2.351	15.874	0.406	1.093	15.867	0.404	1.107
Curitiba	2.597	0.197	0.679	2.649	0.216	0.769	17.186	0.568	2.886	17.017	0.539	2.577
Florianópolis	2.566	0.264	1.026	2.568	0.264	1.020	16.169	0.636	1.918	16.152	0.639	1.926
Porto Alegre	2.976	0.362	2.057	2.984	0.296	1.509	13.853	0.357	1.901	13.708	0.313	0.863
Campo Grande	4.872	0.477	2.467	4.872	0.349	1.353	18.372	0.389	1.201	18.810	0.352	1.062
Cuiabá	4.315	0.374	0.685	4.179	0.418	0.859	17.322	0.601	1.001	17.350	0.655	1.174
Goiânia	4.111	0.319	1.236	4.111	0.321	1.328	17.751	0.481	1.269	17.751	0.480	1.270
Brasília	3.461	0.432	1.636	3.351	0.391	1.432	16.534	0.407	1.063	16.729	0.373	0.870
Interior – North	5.945	0.407	2.206	5.961	0.388	2.054	17.382	0.634	3.212	17.517	0.577	2.816
Interior – Northeast	6.115	0.523	1.803	6.099	0.504	1.867	17.692	0.628	2.708	17.831	0.727	3.326
Interior – Southeast	3.935	0.392	2.335	4.332	0.456	2.796	17.000	0.487	2.105	16.563	0.458	1.910
Interior – South	3.988	0.452	2.568	4.504	0.364	1.798	18.823	0.636	4.033	18.299	0.728	5.094
Interior – Central-West	7.008	0.506	1.901	6.891	0.374	1.336	18.029	0.594	3.105	17.641	0.703	4.359

SE: standard error
deff: design effect

not achieve 50% of the number expected in the plan. thus losing the protection included against the cluster sampling effect (deff = 2). Of the 13 occasions on which this occurred. 10 of them were in capitals in the North or North East. Samples with a performance below 50% of the expected sample size may record abrupt departures from the accuracy criteria and estimates of standard error and deffs should be analyzed with caution.

Table 3 shows the results of estimated prevalence. standard error and deff for the five year old. 12 year old and 65-74 year old age groups. according to geographic domain. For the five year olds. the prevalence of dental caries in deciduous teeth is illustrated. represented by the proportion of individuals with deft ≥ 1 and. at age 12. the prevalence in permanent teeth is shown (proportion of individuals DMFT ≥ 1). The prevalence of bleeding gums is shown for the 65-74 year old age group. In general. the impact of the cluster sample process and of the sampling weight on accuracy is low. There were rare occurrences similar to the prevalence estimates of bleeding in the 65 to 74 years old found in Campo Grande. which doubled the deff when estimated with weights. Also. as expected in the planning. the values for the coefficients of variation do not exceed 15%

Estimates of the mean. standard errors and deffs of the DMFT in the 15 to19 year old group and the 35 to 44 year old group (Table 4) may also be considered stable in the great majority of domains and. thus. compatible with the accuracy criteria set in the design. However. the highest values for deff were reached in th 35 to 44 year old age group in Macapá. São Luís and municipalities in the interior of the macro regions for estimates without weighting. This results are due to the impact of similarity between the individuals who made up the primary sampling units and expected as. as seen before. it is in this group that the greatest deviations between the sample sizes proposed by the design and those actually achieved occurred. This is because interviews were only carried out with eligible elements residing in the selected residences included in the sample.

DISCUSSION

It can be concluded that the samples achieved in the SBBrasil 2010 come close to the principles proposed in the design. With a response rate of above 70% for the residences selected. the probability of being included for all of the individuals was the same as the probability of their residences being selected. However. due to the difference in the age group composition of the base addresses used and what was actually found in the field. these probabilities ended up being unequal between primary units in the same domain. Thus. the users of this database should bear in mind this peculiarity. introducing sampling weight in calculations of point estimates. standard errors. confidence intervals and design effects.

The results shown for the deffs consolidate the results obtained for the means and proportions in four demographic domains. The conservative design feature seems to have preserved the accuracy criteria and the cluster effect. keeping them at desired levels. The number of interviews and examinations carried out was above the minimum planned in 142 of the samples. with deff values of 2 or below. Unfortunately. the 18 samples which achieved below 50% and with deffs above 2 were concentrated in the demographic domain of the 35 to 44 year old age group and in the North and Northeast of Brazil. These results should guide future sample designs for epidemiological surveys for this demographic section of the Brazilian population. especially those which. like the SBBrasil 2010 project. work with professionals of local health care services in order to obtain samples of residences and to carry out interviews and examinations.

Thus. it is important to highlight that the SBBrasil 2010 used an epidemiological survey model which could be incorporated into daily practice by the health care services as an essential tool in health care planning and assessment activities. It should. therefore. combine operational feasibility with representative data.

REFERENCES

1. Cochran WG. Sampling techniques. New York: John Wiley & Sons; 1977.
2. Kish L. Survey sampling. New York: John Wiley & Sons; 1965.
3. Korn EL. Graubard BI. Analysis of health surveys. New York: John Wiley & Sons; 1999. (Wiley Series in Probability and Statistics).
4. United Nations. Department of Economics and Social Affairs. Statistics Division. Household sample surveys in developing and transitions countries. New York; 2005 [citado 2013 ago 13]. (Serie F. 96). Disponível em: http://stats.un.org/unsd/hhsurveys/pdf/Household_surveys.pdf

The *Pesquisa Nacional de Saúde Bucal 2010* (SBBrasil 2010. Brazilian Oral Health Survey) was financed by the General Coordination of Oral Health/Brazilian Ministry of Health (COSAB/MS). through the *Centro Colaborador do Ministério da Saúde em Vigilância da Saúde Bucal*. *Faculdade de Saúde Pública* at *Universidade de São Paulo* (CECOL/USP). process no. 750398/2010.
This article underwent the peer review process adopted for any other manuscript submitted to this journal. with anonymity guaranteed for both authors and reviewers. Editors and reviewers declare that there are no conflicts of interest that could affect their judgment with respect to this article.
The authors declare that there are no conflicts of interest.