BMC Public Health

RESEARCH

Open Access

Precision assessment of a hypertension prevalence survey

Reinaldo José Gianini[1*], Natália Ferreira Caneto[1], Natália Murate Junqueira[1], Leticia Gouvea Rodrigues[1], Beatriz Zurma Parri[1] and Cibele Isaac Saad Rodrigues[1]

Abstract

Background Population surveys are crucial for public policy planning and provide valuable representative data. In the health sector studies to identify and assess the prevalence of Arterial Hypertension (AH), a chronic non-communicable disease (NCD), along with its associated risk factors have been conducted.

Objectives This study aims to assess the effectiveness of a population health survey in estimating the prevalence of arterial hypertension (AH) in the Sorocaba municipality between August 2021 and June 2023. Methods: The analyzed performance indicator is the precision (design effect - deff) of AH prevalence in adults (≥ 18 years) and their exposure to primary risk factors. The total sample included 1,080 individuals from the urban area, deemed sufficient to estimate a deff of 1.5. This cluster-based study utilized census sectors as clusters, with data collected through household interviews, standardized questionnaires, and measurements of blood pressure and biometric parameters. The deff calculation formula used was weighted variance / raw variance. The Research Ethics Committee approved this study, with registration CAAE 30538520-1-0000-5373.

Results The deff values ranged from 0.44 for chronic obstructive pulmonary disease to 1.63 for asthma, with a deff of 1.00 for AH prevalence. Conclusion: The study demonstrated good precision in its results, with high receptivity and cooperation from participants. The cost-effectiveness of the research deemed appropriate. The technique of selecting households within clusters (census sectors) based on detailed mapping and demographic data from the Instituto Brasileiro de Geografia e Estatística (IBGE) proved to be practical and efficient, suitable for replication in other municipalities and for studying other NCDs.

Keywords Hypertension, Prevalence, Cluster sampling, Health surveys, Data accuracy, Epidemiology

*Correspondence:
Reinaldo José Gianini
reinaldognn@gmail.com
[1]Faculdade de Ciências Médicas e da Saúde, Pontifícia Universidade Católica de São Paulo, Sorocaba, Brazil

Background

Population health surveys are epidemiological tools used to construct reliable health information and are carried out with probabilistic and representative samples. These surveys may be international, national, regional or local [1]. When primary data are obtained, the methodology is defined by the researcher, who has control over the data collection and application of the instrument, usually through a questionnaire. On the other hand, when secondary data are collected, the information databases are public and were originally created for administrative, financial or epidemiological purposes [2]. In healthcare, these surveys deal with the description and analysis of the prevalence of diseases and their risk factors, including reported or diagnosed morbidity. They assess, through questionnaires, the functioning of healthcare from the user's point of view, knowledge about morbidity and lifestyles, reliable estimates about the prevalence of a given event, risk and protective factors, conditions of health, the development of health systems, the surveillance of various chronic diseases, and their determinants [3, 4].

The need for health surveys arises from the limitations of the available recorded health data, which may be continuous, as yearly report, or with defined periodicity. These data can be classified as timely, partial, incomplete, unsatisfactory, fragmented, outdated or those that not allows disaggregation [5, 6].

The National Health Survey (PNS), one of the main Brazilian population surveys, uses a probabilistic cluster sample. It is obtained in 3 selection stages: primary units – census tracts; secondary units – households; and tertiary units – individuals aged 18 or over. In 2013, the PNS included 6,081 census sectors, 81,767 households and 62,986 individuals [7].

These surveys can be performed via telephone, mail, the internet, home visits or interviews. Regarding diagnosed morbidity, these methods of surveys include subsamples with clinical examinations and biological specimen collections [8].

In developed countries, population-based surveys have been conducted since the 1960s. In the case of Brazil, the Ministry of Health has made substantial investments in this area since the 1990s, such as financing the National Research Health Supplement by Household Sample (PNAD) since 1967, which became continuous, repeated every three months, since 2012 [9].

Other important surveys in Brazil were the National Cancer Institute (INCA) [10], the Municipality of São Paulo Survey [11], the ISA SP CAPITAL [12], and the VIGITEL which is conducted annually since 2006 [13].

At the international level, the following surveys have been conducted: the National Health Interview Survey (USA) [14], the European Health Interview [15], the General Health Survey for England [16], the National Population Health Survey [17] (Canada), and the World Health Survey [18].

Regarding the prevalence of arterial hypertension (AH) in adults ≥ 18 years old, the 2013 PNS reported a prevalence of 21.4% (95% CI 20.8–22.0) using self-reported criteria, 22.8% (95% CI 22.1–23.4) for measured AH, and 32.3% (95% CI 31.7–33.0) for measured hypertension and/or reported medication use. The prevalence for adults ≥ 75 years old was 55% [19].

For complex samples like the PNS, the method involves cluster sampling with multiple selection stages, each with unequal draw probabilities. Estimating the design effect (deff) on the collected sample data is recommended. Deff is an important indicator in both planning the sample size and evaluating the precision of survey results. The formula to calculate the design effect is typically: $DEFF = \frac{Var_{clustersampledesign}}{Var_{simplerandomsample}}$ [20]. A deff lower than 2 is recommended [21].

Studying the prevalence of chronic non-communicable diseases in Brazilian populations, such as AH, has epidemiological value, as long as the statistical criteria established for defining the prevalence and associated risk factors are valid. To this end, calculating the deff helps in assessing the performance of the methodology used.

Objectives

To assess the performance of a population health survey carried out in the municipality of Sorocaba from 2021 to 2023, which aimed to estimate the prevalence of AH. The performance indicator used is the precision (deff) of the AH and its main risk factors in adults.

Methods

This survey aimed to analyze the urban resident population of the municipality of Sorocaba, aged 18 years and older, estimated at 535,843 individuals in 2022. The total population of Sorocaba in the same year was estimated at 723,682 individuals [22]. The survey was performed using a stratified cluster sampling method Sampling was performed in clusters composed of 6 strata represented by the regionalized division of the Municipal Health Department of Sorocaba: Center South, East, Center North, North, Northwest and Southwest Regions. In each census sector, data were collected from 18 individuals divided into 6 domains: 6 individuals over 60 years of age (3 men and 3 women), 6 individuals aged 40 to 59 years (3 of each sex), and 6 individuals aged 18 to 39 years (3 of each sex). In each stratum, 10 census sectors were selected. The total sample included 1,080 individuals, satisfying the following parameters: expected prevalence of arterial hypertension = 33.7% with an estimated error of 3.65%; significance level of 0.05 considering a two-sided test with Z = 1.96; design effect (deff) = 1.5; and loss of information (non response) = 10%. The selection of 60 census

sectors was conducted according to the following procedure: listing all sectors belonging to a given region in descending order according to the size of the population of each sector, calculating the interval to be used in systematic random sampling (total sectors of the region/10), drawing the sector from the first interval and systematically repeating the interval to select the 9 remaining sectors in that region.

The selection of households was carried out using the following procedure: mapping the residential streets of the census sector, defining the route, calculating the interval between households to be selected (total households/18), drawing a random number between 1 and 18 to select the first household to be included and systematically repeating the interval for the selection of the remaining 17 individuals. To ensure data collection, the recruitment of individuals followed the order of the rarest domain to the most frequent, namely, elderly males, elderly females, young males, young females, adult males and adult females. Collection was extended on holidays, Saturdays and Sundays to ensure the inclusion of workers.

As this is a survey sample, considering each domain, the weight of the sampling unit is the inverse of sampling fractions, and was composed of the following: *a*. Individual weight = total of individuals in the census sector divided by 3; *b*. Sector weight = region population divided by sampled sector population, multiplied by 10; *c*. Weight of the region = total population divided by population of the region, multiplied by 6. The final weight of each sample unit = $(a \times b \times c)$ [20].

For data collection, a standardized form containing sociodemographic data and risk factors for AH was used as an instrument. In addition to the measurements of blood pressure (BP), age and sex of the individuals selected for the sample, additional information was collected, such as skin color/race/ethnicity, marital status, education, occupation/type of work, religion, family income in Reais, number of people in the family, family morbidity history, use of health services, chronic diseases and comorbidities, medical treatment, practice of physical activity, ingestion of alcoholic beverages, and smoking. Every human participant gave its informed consent after reading the Free and Informed Consent Form, which is attached in a supplementary file.

Data were collected from August 2021 to June 2023 during household visits. Interviews were conducted using a standardized questionnaire with open-ended and multiple-choice questions. In addition, blood pressure was measured in accordance with the recommendations of the Brazilian AH Guidelines 2020. Participants who had an average systolic blood pressure ≥ 140 mmHg and/or diastolic blood pressure ≥ 90 mmHg were considered to have AH or who had a BP below these limits and were receiving regular treatment with anti-hypertensive medications [23–25].

The STATA 16.0 program and its survey command were used for the analyses. The calculation of the deff was carried out using the *stat effects* command with the deff option after *svy mean* command for continuous quantitative variables, or *svy*,

proportion for qualitative variables [26]. The deff is the ratio between the weighted variance, which considers both inter-cluster and intra-cluster variance, and the raw variance, which can be thought of as intra-cluster variance assuming the entire sample is one cluster. [20].

This project was submitted and approved by the Research Ethics Committee (CEP) of the Faculty of Medical and Health Sciences (FCMS) of the Pontifical Catholic University of São Paulo (PUC-SP), registered as CAAE 30538520-1-0000-5373.

Results

Between the second half of 2021 and the first half of 2023, data were collected from 1,080 individuals in 60 sectors as planned. The intracluster weights assigned to the individuals interviewed ranged from 1.7 to 161. All sectors planned for collection were included. The sector weights ranged from 1.93 to 159.1. The region weights ranged from 0.14 to 3.35. The greatest data loss was for family income (34%); however, for the other variables, the losses did not exceed 0.4%. The interviewers, always in pairs, strictly followed the collection protocol, trying to cover the entire area of the census sector. Vertical or horizontal condominiums were included, applying the same proportionality rule (draw) for choosing the residential unit. Non-regular situations, such as collective housing or clusters of precarious housing, were registered, analyzed and included, and the same sample selection rule was applied. The interviewers, students of the Scientific Initiation Program of Medicine, supervised by professors, followed the construction of the database in STATA, based on data collected in an EXCEL spreadsheet, and they participated in the first analyses.

Table 1 shows that for the global deff according to social and demographic variables, exception to Ethnicity/Brown, they were less than 1.5, which was the value estimated in the sample planning, indicating better precision of the results. The Family monthly income and Occupation have the lowest deff (1.03), indicating good accuracy in information. As the sampling was planned for representing the global prevalence of AH among people from 18 or more years old, it can be observed that the deff by sex and age categories show more variability, with higher values for 18–39 years and lower values for 60 years or more; this trend is similar for both genders.

Table 2 shows that the global highest deff, defined according to variables related to lifestyle, was for

Table 1 Design effect values (deff), according to sociodemographic variables. Sorocaba-SP/2021-23

Variable	Type	Category*	DOMAINS		Global deff
Sex			**M**	**F**	
Age			**18–39 40–59 60or+**	**18–39 40–59 60or+**	
Marital Status	PQ	Separated	1.65 0.88 0.38	1.36 0.96 0.49	1.32
Ethnicity	PQ	Brown	1.50 0,78 0.48	1.61 1.20 0,58	1.52
Religion	PQ	Christians (not catholic)	1.13 0.58 0.,43	1.12 1.20 0.45	1.29
Education	PQ	High school incomplete	1.58 0.84 0,49	1.62 0.90 0.49	1.48
Family monthly income	Q	Brazilian Reais	1.76 0.94 0.72	1.70 1;32 0.75	1,03
Number of people in the household	Q	total number of people in the household	1.73 0.84 0.47	1.75 1.05 0.52	1.27
Occupation	DQ	Inactive	1.28 0.67 0.34	1.11 0.80 0.16	1.03

*For polytomous variables the category with the highest deff is described. PQ=Polytomous Qualitative; DQ=Dichotomous Qualitative; Q=Quantitative

Table 2 Design effect value (deff), according to variables related to lifestyle. Sorocaba-SP/2021-23

Variable	Type	Category	DOMAINS		deff
Sex			**M**	**F**	
Age			**18–39 40–59 60or+**	**18–39 40–59 60or+**	
Smoking	DQ	Yes	1.53 0.79 0.25	1.15 0.94 0.48	1.12
Time of smoking	Q	years	1.55 0.94 0.43	1.72 1.37 0.60	0.60
Ex-smoker	DQ	Yes	1.35 0.53 0.29	1.33 1.07 0.54	0.94
Time as ex-smoker	Q	years	1.43 1.16 0.55	2.11 1.67 0.71	0.67
Alcoholism	DQ	Yes	1.26 0.72 0.28	1.46 0.73 0.48	1.49
Alcohol amount	Q	Grams per day	1.92 1.10 0.49	1.93 1.29 0.59	1.43
Ex-alcoholism	DQ	Yes	1.28 1.28 0.43	1.51 1.03 0.49	1.05
Use of Illicit drugs	DQ	Yes	1.28 1.22 0.37	1.45 1.23 0.53	1.30
Use of Depressant drugs	DQ	Yes	1.38 0.85 n.c.	1.50 1.07 n.c.	1.32
Use of Stimulant drugs	DQ	Yes	1.37 1.90 0.39	1.55 n.c n.c	1.55
Use of Disruptive drugs	DQ	Yes	1.55 0.86 0.38	1.50 1.23 0.52	1.45
Physical activity	DQ	Yes	1.19 0.74 0.30	1.53 0.83 0.44	1.42
Quantity of physical activity	Q	Hour per week	1.78 1.14 0.44	2.20 1.59 0.55	1.61
Level of physical activity	Q	Moderated Stress load	1.97 0.77 0.46	2.03 1.12 0.55	1.60
Low sodium diet	DQ	Yes	1.55 0.61 0.50	1.12 0.77 0.27	1.50
Salt shaker on the table	DQ	Yes	1.68 0.66 0.48	1.48 0.97 0.62	1.25

DQ=Dichotomous Qualitative; Q=Quantitative; n.c=non calculable

Quantity of physical activity (1.61); deff values lower than 1 were observed for time of smoking (0.60) and time as an ex-smoker (0.67). Ex-smokers and ex-alcoholics have deff values close to 1. Again, it is verified the same trend by sexo and age, higher deff for 18–39 years. Use of depressant drugs and Use of stimulant drugs show non-calculable deff because had not enough sample number.

Table 3 shows that according to the variables related to family and personal morbidity history, the global deff varied from 0.44 for chronic obstructive pulmonary disease (COPD) to 1.63 for asthma. Those values were equal to or close to 1 for dyslipidemia, and hyperthyroidism. Family member with some morbidity has deff lower than 1 for all domains.

Table 4 shows that according to variables related to the clinical examination, the global deff varied from 0.88 for arm circumference to 1.28 for stress, and the deff was very satisfactory for HA, at 1.00. Height, and metabolic syndrome factors presented deff values close to 1. The same pattern is verified according to domains: higher deff for youngers.

As shown in Table 5, according to variables related to the drug treatment of AH the global deff varied from 0.86 for the Use of calcium channel blockers, to 1.41 for the use of sympatholytics.

Discussion

This research can be considered a successful study representing the urban population of Sorocaba. It required significant effort but minimal resources. The strategy of implementing it as a scientific initiation program with students supervised by medical teachers utilizing available software at the institution was effective. In terms of urban population aged 18 and above, it can be inferred that the selection bias was minimal, evident from both data analysis and comparison with other surveys [20, 26–29].

Table 3 Value of the design effect (deff) according to variables related to family and personal morbidity antecedents. Sorocaba-SP/2021-23

Variable	Type	Category	DOMAINS		deff
Sex			**M**	**F**	
Age			**18–39 40–59 60or+**	**18–39 40–59 60or+**	
Family member with some morbidity	DQ	Not	0,54 0.43 0.28	0.64 0.60 0.36	1.11
Family member with stroke	DQ	Yes	1.82 0.74 0.54	1.49 0.75 0.40	1.26
Family member with acute myocardial infarction	DQ	Yes	1.66 0.95 0.51	1.40 1.17 0.59	1.42
Family member with diabetes mellitus	DQ	Yes	1.28 0.48 0.43	1.15 0.97 0.42	1.29
Family member with high blood pressure	DQ	Yes	0.64 0.36 0.33	0.72 0.44 0.34	1.15
Comorbidities (interviewed)	DQ	Yes	1.32 1.18 0.50	1.92 1.47 0.88	1.37
Dyslipidemia (interviewed)	DQ	Yes	1.40 1.05 0.43	1.45 1.38 0.49	1.05
Ischemic Heart Disease (interviewed)	DQ	Yes	n.c. 0.76 0.58	1.55 1.33 0.59	0.89
Stroke (interviewed)	DQ	Yes	1,52 0.88 0.81	n.c. 1.31 0.45	0.89
Chronic Kidney Disease (interviewed)	DQ	Not	0.12 0.13 0.14	0.09 0.15 0.10	0.85
Obstructive Arterial Disease (interviewed)	DQ	Not	0.12 0.13 0.18	0.07 0.15 0.06	0.85
Type 2 diabetes mellitus (interviewed)	DQ	Yes	1.45 1.01 0.32	1.50 1.23 0.45	0.86
Asthma (interviewed)	DQ	Yes	1.73 0.95 0.37	1.87 1.52 0.44	1.63
Chronic Obstructive Pulmonary Disease (interviewed)	DQ	Not	0.12 0.13 0.22	0.07 0.15 0.10	0.44
Hypothyroidism (interviewed	DQ	Yes	1.31 0.90 0.47	1.30 1.48 0.65	1.16
Hyperthyroidism (interviewed	DQ	Yes	1.07 n.c n.c	1.55 1.07 0.51	1.05
Gastritis (interviewed)	DQ	Yes	n.c 0.95 0.39	1.52 1.05 0.52	1.15
Osteoporosis (interviewed	DQ	Not	0.12 0.14 0.21	0.75 0.15 0.14	0.58
Common Mental Disorder (interviewed	DQ	Yes	1.38 0.76 0.58	1.26 1.36 0.37	1.21

DQ = Dichotomous Qualitative; n.c = non calculable

In terms of precision, the research demonstrated high satisfatory levels, with a global deff of less than 1.64 for all variables analyzed. Specifically for arterial hypertension, the deff was 1.00 in our study, indicating high precision. For example, in the study by Alves [20], it was 1.37 for the group of teenagers and 1.52 for the group of elderly people. From the data, it can be inferred that one of the axes of precision was objective information, subject to verification, notably variables related to clinical examination, as opposed to subjective information such as variables related to lifestyle. The other axis of precision refers to the frequency of the event in the population and its intra-cluster and inter-cluster variance. If the clusters are more similar than expected for a random distribution of a given variable, the deff may be less than 1. Furthermore, if the variable is widely distributed in the population, there is a small cluster effect, and the prevalence of the event studied differs significantly inside strata, so it is likely that the deff is less than 1 [27]. It can occur in both, very frequent and rare events, which explains the deff of osteoporosis (0.58) and of COPD (0.44), respectively.

Concerning the study limitations, it is important to note that information on family income was highly unreliable, that information on drug use was probably influenced by modesty or fear, and that information on seeking health services was compromised by the respondents' memories [30, 31, 32].

As for the domains, the tendency for higher deff for younger people indicates the need to increase this subsample, while deff below one for older people indicates a smaller amount needed for this subsample, and this occurs in both men and women. This is because there is a sample planned for satisfying a maximum global deff of 1.5 for AH, and not planned for each domain. Furthermore, some very specific conditions demonstrate that different sample calculations would be necessary if there were interest.

As suggestions for future studies, the comorbidities of the interviewees could have been diagnosed in some situations. For example, in patients with diabetes, asthma and COPD, capillary blood glucose tests and spirometry could be employed. Furthermore, considering that it is a resident population and that the bond established in the first contact was of high quality, it would be possible to schedule new visits to obtain more information, including various clinical protocols, such as adherence to treatment.

Conclusion

A survey was conducted with the urban population of Sorocaba between 2021 and 2023 for measuring the prevalence of AH in adults. The results showed good precision, with a global deff of 1.00 for this variable. There was good receptivity and collaboration from interviewees, and the research had an adequate cost–benefit ratio. The technique used for selecting households in the respective clusters (census sectors), based on detailed mapping and

Table 4 Design effect value (deff) according to variables related to the clinical examination. Sorocaba-SP/2021-23

Variable	Type	Category	DOMAINS		deff
Sex			**M**	**F**	
Age			**18–39 40–59 60or+**	**18–39 40–59 60or+**	
Weight	**QT**	**kilos**	**1.47 0.72 0.57**	**1.67 0.78 0.77**	**1.10**
Height	QT	centimeters	1.29 1.19 0.42	1.45 1.07 0.70	1.07
Arm circumference	QT	centimeters	1.58 0.86 0.39	1.60 1.07 0.68	0.88
Abdominal circumference	QT	centimeters	1.37 0.87 0.55	1.54 1.07 0.65	1.08
Average Systolic Blood Pressure	QT	mmHg	1.56 1.04 0.40	1.52 0.96 0.54	0.90
Average Diastolic Blood Pressure	QT	mmHg	1.52 0.88 0.38	1.45 1.23 0.54	1.10
Average Heart Rate	QT	beats/minute	1.53 1.09 0.53	1.17 1.07 0.50	1.08
Body mass index	QT	kg/m^2	1.40 0.89 0.69	1.44 1.12 0.77	1.19
Hypertension	QL	> or < 140/90 mmHg	1.25 0.48 0.29	1.30 1.06 0.22	1.00
Metabolic Syndrome Factors	QL	Yes	1.15 0.44 0.39	0.90 0.59 0.25	1.02
Stress	QL	Yes	1.05 0.57 0.25	0.86 0.57 0.42	1.28

QL = QUALITATIVE QT = QUANTITATIVE

Table 5 Design effect value (deff) according to variables related to the drug treatment of arterial hypertension. Sorocaba-SP/2021-23

Variable	Type	Category	DOMIAINS		deff
Sex			**M**	**F**	
Age			**18–39 40–59 60or+**	**18–39 40–59 60or+**	
Medical insurance	Q	Yes	0.94 0.75 0.33	0.73 0.80 0.32	1.13
Diagnosis of hypertension	Q	Yes	1.71 0.73 0.37	1.36 1.02 0.31	0,91
Hypertension treatment	Q	Yes	n.c 1.23 0.63	1.42 1.29 0.79	1.08
Use of angiotensin II AT1 receptor blockers	Q	Yes	2.21 0.93 0.51	1.32 1.55 0.51	1.07
Use of angiotensin II converting enzyme inhibitors	Q	Yes	1.5 1.12 0.39	1.49 1.42 0.50	0.96
Use of calcium channel blockers	Q	Yes	1.48 1.05 0.48	n.c 1.23 0.45	0.86
Use of thiazide diuretics	Q	Yes	1.36 0.76 0.59	1.47 0.99 0.62	0.88
Beta blocker use	Q	Yes	1.36 0.84 0.52	1.44 1.28 0.44	0.93
Loop diuretic use	Q	Not	0,12 0.14 0.19	0.07 0.15 0.10	0.90
Use of spironolactone	Q	Not	0.12 0.15 0.21	0.07 0.14 0.09	0.90
Use of sympatholytic	Q	Not	0.12 0.14 0.21	0.09 0.14 0.09	1.41
Use of direct-acting vasodilators	Q	Not	0.12 0.14 0.21	0.75 0.14 0.10	0.96
Time since last consultation	QT	months	1.37 0.92 0.34	1.99 1.47 0.62	1.04
Recent search for health services	Q	Yes	1.08 0.80 0.27	1.26 0.81 0.51	1.17

Q = Qualitative QT = quantitative

demographic information from the IBGE, proved to be practical and efficient.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12889-024-19626-z.

Supplementary Material 1

Supplementary Material 2

Acknowledgements
Thanks for support from Sorocaba City Hall.

Author contributions
Reinaldo José Gianini (RJG)planned, supervised, analyzed and interpreted data, and wrote this article. Natália Ferreira Caneto planned, analyzed and interpreted data Natália Murate Junqueira (NMF) planned, analyzed and interpreted data Leticia Gouvea Rodrigues (LGR) planned, analyzed and interpreted data Beatriz Zurma Parri (BZP) planned, analyzed and interpreted data Cibele Isaac Saad Rodrigues (CISR) planned, supervised, analyzed and interpreted data, and wrote this article.

Funding
This research received funding from Pontifícia Universidade Católica de São Paulo, with grant numbers 17807 and 15504.

Data availability
The data that support the findings of this study are not openly available due to reasons of sensitivity and are available from the corresponding author upon reasonable request.

Declarations

Ethical approval
This project was submitted and approved by the Research Ethics Committee (Comitê de Ética em Pesquisa) of the Faculdade de Ciências Médicas e da Saúde of the Pontifícia Universidade Católica de São Paulo, registered as CAAE 30538520-1-0000-5373.

Consent to participate
Every human participant has given their consent to participate.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Received: 16 April 2024 / Accepted: 29 July 2024
Published online: 12 August 2024

References

1. Silva VSTM, Pinto LF. Nationwide population-based household surveys in health: a narrative review. Ciênc saúde Coletiva. 2021;26(09):4045–58. https://doi.org/10.1590/1413-81232021269.28792020.
2. Kindratt TB. Big Data for Epidemiology: Applied Data Analysis Using National Health Surveys, 2022.
3. Malta DC, Leal MC, Costa MFL. Inquéritos Nacionais De Saúde: experiência acumulada e proposta para o inquérito de saúde brasileiro. Rev bras Epidemiol. 2008;11(supl1):159–67. https://doi.org/10.1590/S1415-790X2008000500017.
4. http://doi.org: 10.1136/jech.2008.077172.
5. Mota E. Integrating population surveys to the National Health Information System. Ciênc saúdecoletiva. 2006;11(4):870–86. https://doi.org/10.1590/S1413-81232006000400010.
6. Andrade FRe, Narval PC. Population surveys as management tools and health care models rev. Saúde Pública. 2013;47(Suppl3):154–60. https://doi.org/10.1590/S0034-8910.2013047004447.
7. Instituto Brasileiro de Geografia e Estatística - IBGE. Pesquisa Nacional De Saúde [National Health Survey] – 2013. https://www.ibge.gov.br › estatisticas › sociais › saude.
8. Monteiro CA, Florindo AA, Claro RM, Moura EC. Validity of indicators of physical activity and sedentariness obtained by telephone survey. Rev Saude Publica. 2008;42(4):575–81. https://doi.org/10.1590/s0034-89102008000400001.
9. Pesquisa. Nacional por Amostra de Domicílios – PNAD [National Household Sample Survey] – 2008. www.datasus.gov.br/informações de saúde/inquéritos e pesquisas/suplemento saúde.
10. Instituto Nacional do Câncer INCA. Inquérito Domiciliar [Household Survey]. 2004 http://www.inca.gov.br/inquerito/
11. Inquérito de Saúde [Health Survey] do Municipio de São Paulo ISA-SP-Capital. 2008. http://extranet.saude.prefeitura.sp.gov.br/areas/ceinfo/inquerito-de-saude-isa-capital
12. Zaitune MPA, Barros MBA, César CLC, Carandina L, Goldbaum M, Alves MCGP. Factors associated with global and leisure-time physical activity in elderly individuals: a health survey in São Paulo (ISA-SP), Brazil. Cad Saude Publica. 2010;26(8):1606–18. https://doi.org/10.1590/s0102-311x2010000800014.
13. Vigilância de fatores. de risco e proteção para doenças crônicas por inquérito telefônico VIGITEL – 2008. www.datasus.gov.br/informações de saúde/inquéritos e pesquisas.
14. Centers for Disease Control and Prevention – National Center for Health Statistics. Natl Health Interview Surv. 2019. https://www.cdc.gov/nchs/nhis/about_nhis.htm
15. Kuulasmaa K, Tolonen H (Prepared), editors. What is EHES and why it is needed? National Institute for Health and Welfare, 2013. Discussion Paper 2013_007. URN:ISBN: 978-952-245-844-5, URL: http://urn.fi/URN:ISBN:978-952-245-844-5
16. National Health Services. Narional Health Survey for England 2019 (NS). http://healthsurvey.hscic.gov.uk/support-guidance/public-health/health-survey-for-england-2019.aspx
17. Canadian Research Data Centre Network. National Population Health Survey, 1994–2011. https://crdcn.ca/data/national-population-health-survey/
18. World Health Organization – WHO. World Health Surveys 2024.WHO.Int. https://apps.who.int/healthinfo/systems/surveydata/index.php/collections/whs
19. Andrade SSA, Stopa SR, Brito AS, Chueri PS, Szwarcwald CL, Malta DC. Self-reported hypertension prevalence in the Brazilian population: analysis of the National Health Survey, 2013. Epidemiol Serv Saúde. 2015;24(2):297–304. https://doi.org/10.5123/S1679-49742015000200012.
20. Alves MCGP, Escuder MML, Goldbaum M, Barros MBA, Fisberg RM, Cesar CLG. Sampling plan in health surveys, city of São Paulo, Brazil, 2015. Rev Saúde Pública. 2018;52. https://doi.org/10.11606/S1518-8787.2018052000471.
21. Barata RB, Moraes JC, Antonio PRA, Dominguez M, World Health Organization. Rev Panam Salud Publica. Immunization coverage survey: empirical assessment of the cluster sampling method proposed by the. 2005;17(3):184 – 90. https://doi.org/10.1590/S1020-49892005000300006 22. Instituto Brasilerio de Geografia e Estatistica. Censo 2022. IBGE. https://www.ibge.gov.br/estatisticas/sociais/trabalho/22827-censo-demografico-2022.html
22. Instituto Brasilerio de Geografia e Estatistica. Censo 2022. IBGE. https://www.ibge.gov.br/estatisticas/sociais/trabalho/22827-censo-demografico-2022.html
23. Barroso WKS, Rodrigues CIS, Bortolotto LA, Mota-Gomes MA, Brandão AA, Feitosa ADM, et al. Diretrizes Brasileiras De Hipertensão arterial – 2020. Arq Bras Cardiol. 2021;116(3):516–658. https://doi.org/10.36660/abc.20201238.
24. Rodrigues CIS. (coord) – Diretrizes brasileiras de hipertensão VI. Diagnósticoeclassificação.Braz.J.Nephrol2010;32(suppl1). https://doi.org/10.1590/S0101-28002010000500004
25. Malachias MVB, Souza WKSB, Plavnik FL, Rodrigues CIS, Brandão AA, Neves MFT, et al. 7ª Diretriz Brasileira De. Hipertensão Arterial Arq Bras Cardiol. 2016;107(3 Suppl 3):1–103.
26. STATA 16.0 (R). Copyright 1985–2019 StataCorp LLC. Statistics/Data Analysis. Texas USA.
27. Thomson DR, Rhoda DA, Tatem AJ, Castro MC. Gridded population survey sampling: a systematic scoping review of the field and strategic research agenda. Int J Health Geogr. 2020;19:34. https://doi.org/10.1186/s12942-020-00230-4.
28. USAID. Sampling and household listing manual. Demographic and health survey methodology. Calverton, USA: ICF International; 2012.
29. Grais RF, et al. Don't spin the pen: two alternatives methods for second-stage sampling in urban cluster surveys. Emerg Themes Epidemiol. 2007;4(8). https://doi.org/10.1186/1742-7622-4-8.
30. Silva NNe, Roncalli AG. Sampling plan, weighting process and design effects of the Brazilian oral Health Survey. Rev Saude Publica 2013:47(supl.3):3–11. https://doi.org/10.1590/S0034-8910.2013047004362
31. Galobardes B, Lynch J, Smith GD. Measuring socioeconomic position in health research. Br Med Bull. 2007;81–2. https://doi.org/10.1093/bmb/ldm001.
32. Rothman KJ, Greenlan S, Lash TL. Modern epidemiology. Lippincott Williams & Wilkins; 2008.

Publisher's Note