Planning and execution of a population-based health survey by means of a multidisciplinary research consortium

Planejamento e execução de um inquérito populacional de saúde por meio de consórcio de pesquisa multidisciplinar

Samuel Carvalho Dumith1,2, Renata Gomes Paulitsch2, Marina Xavier Carpena1, Monica de Fátima Rossato Muraro1, Milena de Oliveira Simões1, Karla Pereira Machado1, Mariane da Silva Dias1, Andressa Carine Kretschmer2, Manoela Maciel Olić2, Leandro da Silveira Pontes2, Lulie Rosane Odeh Susin1

1 Graduate Program in Public Health, School of Medicine, Federal University of Rio Grande (FURG). Rio Grande, RS, Brazil.
2 Graduate Program in Health Sciences, School of Medicine, Federal University of Rio Grande (FURG). Rio Grande, RS, Brazil.

How to cite this article:

ABSTRACT
AIMS: To describe the methodological and operational aspects of a health survey, conducted by means of a multidisciplinary research consortium.

METHODS: A cross-sectional, population-based study was conducted with individuals 18 years or older, living in the urban area of the municipality of Rio Grande, in Rio Grande do Sul state, Brazil. The sampling was probabilistic, with census tracts as the primary sample unit. The research instrument consisted of a standardized and previously tested questionnaire, with closed questions, applied by face-to-face in dwellings. A brief questionnaire was administered to 10.5% of subjects for quality control purposes. Fieldwork lasted from April to July 2016, and the study was approved by an ethics committee.

RESULTS: Of the 70 sampled census tracts, 711 households were randomly selected, comprising 1,429 eligible individuals. Overall, 1,300 individuals (91.0%) of 676 households (95.1%) answered the questionnaire. Losses and refusals were more prevalent for males and downtown’s sectors. The questionnaire reproducibility was satisfactory (mean κ = 0.80).

CONCLUSIONS: The main methodological aspects of a multidisciplinary research consortium were pointed out, which may be of interest to other researchers. We highlight the importance of this type of study to yield information on various health conditions of the population.

KEYWORDS: epidemiology; cross-sectional studies; health surveys; methodology; health.

RESUMO
OBJETIVOS: Descrever os aspectos metodológicos e operacionais de um inquérito de pesquisa sobre saúde, realizado por meio de um consórcio de pesquisa multidisciplinar.

MÉTODOS: Um estudo transversal de base populacional foi realizado com indivíduos de 18 anos ou mais, residentes na zona urbana do município de Rio Grande, RS, Brasil. A amostragem foi probabilística, tendo como unidade primária amostral os setores censitários. O instrumento de pesquisa consistiu num questionário padronizado, com questões fechadas e previamente testado, aplicado face a face no domicílio. Um questionário resumido foi aplicado a 10,5% dos indivíduos para fins de controle de qualidade. O trabalho de campo se estendeu de abril a julho de 2016 e a pesquisa foi previamente aprovada por comitê de ética em pesquisa.

RESULTADOS: Dos 70 setores censitários amostrados, foram selecionados 711 domicílios e 1.429 indivíduos. Respondem o questionário 1.300 indivíduos (91,0%) de 676 domicílios (95,1%). As perdas e recusas foram mais prevalentes para os indivíduos do sexo masculino e dos setores do centro da cidade. A reprodutibilidade do questionário foi satisfatória (κ médio = 0,80).

CONCLUSÕES: Apontaram-se os principais aspectos metodológicos de um consórcio de pesquisa multidisciplinar, que poderão ser de interesse para outros pesquisadores. Destaca-se a relevância deste tipo de estudo para a produção de informações sobre diversas condições de saúde da população.

DESCRITORES: epidemiologia; estudos transversais; inquéritos populacionais; metodologia; saúde.
INTRODUCTION

The Federal University of Rio Grande (FURG) School of Medicine currently sponsors graduate programs in health sciences (GPHS) and public health (GPPH). Both graduate programs have an accredited master’s degree (M.Sc.) program and the GPHS also has a doctorate degree (Ph.D.) program. In 2015, a joint initiative of the GPHS and GPPH was established to develop a population-based study, in which the research data were collected through a consortium, that is, a collaborative project among researchers [1].

A research consortium consists of a set of different research goals to be accomplished through a collaborative [1] fieldwork based on individual projects [2]. This work format is reported to strengthen efforts and optimize resources [1] while making the fieldwork simpler and enabling the inclusion of a greater sample size and the attainment of accurate outcomes [2]. It differs from other methodological approaches by providing all participants with opportunities for new knowledge acquisition and active learning, as well as for developing the ability to work as a team [3].

A research consortium facilitates the execution of large-scale population-based surveys through which several health indicators are evaluated in a representative sample of the population, with the possibility of generalization of inferences [1]. With this, information concerning the health-disease process at a population level is obtained (e.g. magnitude and distribution of diseases) and used to map out different epidemiological realities. Ultimately, the outcomes obtained in large-scale studies may help drive the planning of health actions and other public policies for a better quality of life of the population [4].

As epidemiological surveys are commonly time- and resource-consuming, their execution could be impracticable if carried out in an individualized way, which reinforces the importance of the collaborative work [3]. Of note, successful research consortia have been established in other graduation programs in Brazil [2, 3].

It should be mentioned that most scientific journals have restrictive submission rules for article length due to space limitation, which results in lack of details concerning the study operationalization [4]. Here, we describe the methodological aspects of a research consortium implemented in 2016 in a municipality in southern Brazil.

METHODS

The research consortium

The project entitled “Health of the Riograndina Population” (“Saúde da População Riograndina”) was executed through a research consortium. The study was conceptualized in July 2015, when the first meeting was held. The group was composed of graduate students from various fields of knowledge, namely: nutrition, physical education, dentistry, psychology and physiotherapy.

The selected topics were discussed by the students and their supervisors according to their areas of interest. The relevance and adequacy of each topic, as well as study feasibility within the context of the research consortium, were discussed during the group meetings. After each student elaborated their research project, a single comprehensive project of the research consortium was organized, which contemplated the different objectives of each individual project and a shared methodological section.

The work team consisted of two coordinators, whom were GPPH and GPHS professors, ten supervisors (graduate students) and nine interviewers hired to collect the data. In addition, two undergraduate medical students joined the team to assist in entering data.

Study location, population and design

The municipality of Rio Grande is located in the southernmost part of the state of Rio Grande do Sul and has approximately 200,000 inhabitants, of which about 95% reside in the urban area. Its economy is mostly concentrated in port activity. The Human Development Index of Rio Grande was 0.744 in 2010, with Gross Domestic Product per capita of BRL 34,997.50 (Brazilian Reais) in 2015 [6].

The population of this study consisted of individuals aged 18 years or older residing in the urban area. Individuals institutionalized in nursing homes, hospitals and prisons, and those with physical and / or cognitive inability to respond to the questionnaire, were not considered eligible.

This was a cross-sectional population-based study. This design is adequate for the evaluation of health-related outcomes and investigation of associated factors. In population-based surveys,
the cross-sectional design enables data collection in a short period of time and with lower resource demands [7, 8].

Sample size calculation and sampling

Two sample size calculations were performed, one for the study of prevalence of outcomes (main variables of each individual project) and another for the study of the factors associated with the outcomes. In order to estimate the prevalence of different health outcomes, the sample size was calculated considering the following parameters: expected prevalence of 10%, margin of error of two percentage points and 95% confidence level, resulting in a sample of 860 individuals. An additional 50% was included to take into account the effect of sample design (estimated to be 1.5), totaling 1,290 individuals. The effect of sample design appears when the unit of the sample differs from the unit of analysis. This occurs mainly in conglomerate and multiple-phase sampling. Another 10% was added for possible sample dropouts or refusals, resulting in a final sample size of 1,420 individuals.

In the study of associated factors, the sample size was calculated considering a 95% confidence level, 80% statistical power, 10% prevalence of the outcome, frequency of exposures between 20% and 60%, and a prevalence ratio of 2, resulting in an n of 784 individuals. An additional 50% was included for the effect of sample design (estimated to be 1.5) plus 15% due to possible confounding factors, resulting in 1,294 individuals. Lastly, an additional 10% were added for possible sample dropouts or refusals, resulting in a final n of 1,423 individuals.

The sampling process took place in two stages and was based on information from the 2010 Demographic Census of the Brazilian Institute of Geography and Statistics [5]. First, the census tracts were selected, followed by households and, lastly, by residents. To obtain a sample size of approximately 1,420 individuals, a total of 710 households were considered, as the average number of residents aged 18 years or older per household was two [5].

Seventy-two census tracts were systematically selected out of the 293 eligible tracts, which accounted for 25% of the total of eligible tracts, with an average of 10 households per tract. As the primary sampling unit was “tracts”, followed by “households”, the design effect was taken into account in the sample size calculation. To minimize this effect, we preferred to select more tracts and fewer households in each census tract.

All households in the urban area (77,835) of Rio Grande were initially allocated in descending order of the average monthly income of the head of the household. The first household was randomly selected, followed by a “leaping” systematic selection of 1,080 households. With this, 72 census tracts were selected, corresponding to 23,439 households, from which 711 (30 neighborhoods) were drawn after a “leap” of 32 households. Two census tracts were excluded for not having households drawn. All individuals aged 18 years or older from the selected households were considered eligible to be part of the sample.

Research instrument

The study information was obtained through a single precoded and standardized questionnaire, which was previously tested. This questionnaire was composed mostly of closed questions on socioeconomic, demographic and behavioral characteristics, morbidity, nutritional status, access to and use of healthcare services, quality of life, physical activity practice, characteristics of the environment, as well as mental and oral health.

The variables were divided into 15 blocks containing 252 questions. There was a block (socioeconomic aspects) and a set of questions about food insecurity that were answered only by the head of the household. As for the variable “index of goods”, the head of the household was also asked about the ownership or not of some items described by the Brazilian Association of Research Companies, such as color television set, bathroom, automobile, motorcycle, DVD player, microwave oven, refrigerator and washing machine [9]. Among the validated research instruments that were included in the questionnaire are the World Health Organization Quality of Life (WHO-QoL) [10], Brazilian Food Insecurity Scale (EBIA) [6], International Physical Activity Questionnaire-short version (IPAQ-short version) [11], Perceived Stress Scale-14 (PSS-14) [12] and Patient Health Questionnaire-9 (PHQ-9) [13]. The application of the questionnaire lasted 30 min on average.

An instruction manual was developed to assist in the application of the questionnaire and completion of the answers. The manual contained guidelines on how questions should be asked, the meaning and synonyms of some terms, as well as instructions on how to proceed in case of questions raised by the interviewees. Each student was responsible for preparing the instructions for their questions, and then two other students assembled the manual.
Selection of interviewers

The following criteria were used for selection of interviewers: a) female sex; b) complete high school; c) availability at night and on weekends; d) participation in the training; e) succeed in the training assessments. The decision to select only women to conduct the interviews was based on the higher likelihood of residents to open their house door to female interviewers and to participate in the survey.

The interviewer positions were posted on the FURG’s website and on fliers disseminated in local technical schools and universities, job advertisement boards and in a newspaper of great circulation in the municipality of Rio Grande. Among the 165 applicants, 56 were pre-selected and 20 were selected as potential interviewers whom next underwent a training process.

The training of the interviewers was carried out for five days (30 h) and addressed the following topics: presentation of the project; introduction to the research instrument and instruction manual; handling the questionnaire for data collection; reading the questionnaire; simulations and dramatizations of interviews; theoretical test including related questions; dressing code of the interviewers (compulsory use of the identification badge, T-shirts with identification and a cap); pilot study, in which each interviewer conducted an interview under supervision.

Of the twenty interviewers selected, only sixteen completed the training and nine were hired for the study. However, only four interviewers remained from the midpoint to the end of the fieldwork and they performed approximately 80% of all interviews. This setback, created by the dropout of some interviewers, ended up extending the data collection timeline for another two months (April to July 2016) beyond the period originally planned.

Logistics

Dissemination of the study

Prior to data collection, the study was disseminated via interviews in local radio stations, publications on local newspapers and TV newscasts, and on a social media page on the internet. During the fieldwork, there was a continuous effort to disseminate the study so that to increase participation among the selected sample subjects.

Prior household visits

The selected households were visited by the study supervisors prior to data collection. The purpose of this visit was to verify if the household was residential and occupied, to determine the number of residents aged 18 or older, as well as to inform the residents about the study and arrange a schedule for the interview.

During the visit, the supervisors asked the residents for contact numbers and took notes of some house coordinates to facilitate the work of the interviewers. In case the house was uninhabited or did not exist, the house immediately next to the right (looking at it from the front), would be selected.

Pilot study

Prior to data collection, a pilot study was carried out in a census tract not included in the sample. The study questionnaire, manual of instructions, organization of the fieldwork and performance of twenty interviewers were tested. Following the pilot study, nine interviewers out of the sixteen who completed the training were selected for data collection.

Data collection

Each supervisor (graduate student) was responsible for accompanying the interviewer throughout data collection to monitor their work and clarify any questions and/or issues. The receipt of blank questionnaires and the delivery of completed questionnaires by the interviewers occurred in person on Monday mornings. The supervisors were responsible for receiving, reviewing and coding the questionnaires and entering the data. If necessary, they called the interviewees in case a question had not been filled out.

As an initial strategy, the interviewers collected data in locations close to where they lived. With the dropout of some interviewers, the data collection strategies had to be defined on a weekly basis. From the midpoint to the end of the fieldwork, the interviewers pulled together to work within the same neighborhood accompanied by at least two supervisors. For a few weeks, a car made available by the university was used. Since this car did not operate at night nor on Saturdays and Sundays, the study supervisors pulled together to work at night and on weekends. This procedure accelerated the data collection timeline and prevented further delays.

Quality control

The quality control was performed concomitantly with the data collection. On a weekly basis, a fieldwork supervisor randomly drew 10% of the questionnaires applied in the previous week and called the residents to ask standard questions. This procedure was accomplished to check whether the interviewers...
were indeed visiting the households to perform the interviews.

**Data typing and processing**

The data typing started from the second week of data collection, when the first batch of questionnaires was received. The first data entry was carried out by the supervisors and the second by two undergraduate research fellows. Prior to typing, all questionnaires were reviewed and coded.

The questionnaires were numbered according to their corresponding census tract (1 to 70), household (1 to 30) and resident (1 to 9) and stored in batches. Each batch contained questionnaires from five census tracts, and a total of fourteen batches were obtained. The data were entered into Epi-Data 3.1 program and then transferred to the statistical package Stata 11.2, where the database was cleaned and new variables were created.

The questionnaire (which contained eleven key questions) was resubmitted to 10.5% of the sample. The *kappa* coefficient was used for calculation of concordance. To estimate the fraction of the total variability of measures due to variations between individuals, we used the intraclass correlation coefficient (ICC), a variant for the calculation of the sample design effect, and the size of the cluster. In our study, the conglomerate was the average number of individuals per census tract.

**Ethical issues**

This research project was approved by the Research Ethics Committee in the Health Center of the FURG in March 2016, under protocol no. 20/2016. The study followed the guidelines of the Resolution 466/12, and all subjects signed an informed consent form to participate in the study. Those who could not read and/or write authorized their participation by depositing their fingerprints on the form. Participants who requested some type of assistance (e.g. medical, nutritional or psychological) were properly referred to the corresponding service(s).

**Funding**

The study expenses were budgeted at BRL 24,000 (Brazilian Reais), of which BRL 1,500 were allocated for preparation of support materials (T-shirts, caps, badges, office supplies and magnets with the study logo); BRL 18,000 were used to pay for the interviewers’ stipends; BRL 4,000 were used to pay for transportation and food expenses; and BRL 500 were used to pay one of the typists.

Part of the study expenses were sponsored by the graduate programs, which contributed with the amount of BRL 11,500. The master and doctoral students involved in the research consortium also contributed with BRL 1,000 each, totaling BRL 10,000. Lastly, one of the study coordinators sponsored the amount of BRL 2,500. The printing of the questionnaires (BRL 1,500) was performed by the FURG printing office.

**RESULTS**

A total of 676 (95.1%) out of 711 eligible households were surveyed, and 1,300 (91.0%) out of 1,429 individuals eligible for the study were interviewed. Among the 129 non-respondents, 99 were refusals (77%) and 30 were dropouts (23%). The refusals and dropouts were more prevalent in males (12% versus 6.5% in females, *p* < 0.001). There was no difference in the mean age of participants and non-participants (*p* = 0.64). Moreover, there was a greater number of dropouts/refusals in the downtown district, which accounted for 20% of the total of eligible residents.

Two participants answered only half of the questionnaire and two others gave up before the end of the interview. In addition, for some questions, the participant did not know/remember the answer or did not want to answer them. The variable with the highest number of ignored answers was the income of each household member, with 9.9% of non-responses (n = 128 individuals).

The questionnaire containing eleven key questions was resubmitted to 137 participants (10.5% of the sample). The mean *kappa* coefficient for the eleven questions was 0.80, which indicates a substantial agreement [14]. The study topics investigated, as well as the number of respondents, the effect of sample design and the ICC values are described in Table 1.

Until the moment this article was finished, the scientific production resulting from this research project consists of three articles in press, six articles submitted, and twelve ongoing studies, as well as several dissertations, theses and presentations in national and international conferences. Concomitantly, the main results of the study were made available to participants by means of a letter sent via direct mail to their respective addresses. The results were also disclosed in the Municipal Health Department and in the local press (newspapers, radio stations and television newscast).
DISCUSSION

Cross-sectional population-based studies are of great importance to map out the health status of the community at a given time frame. Our study surveyed a representative sample of the urban population aged 18 years and older in the municipality of Rio Grande. To the best of our knowledge, there is only one report of a population-based survey carried out in Rio Grande in the year 2000, which addressed the access of the population to healthcare services [15]. In that study, 1,260 individuals aged 15 years and older, from 540 households within 45 census tracts, were interviewed [15].

Individual response variations were observed in the interviews, with obesity and self-perception of healthy eating being the most and the least frequent topics, respectively. It is believed that the report of obesity and chronic respiratory diseases may have been influenced by recall bias, since obesity was deduced from self-reported weight and height and respiratory diseases from the memory of medical diagnoses. Moreover, the variable bruxism was reported by only a few interviewees, which can be explained by the lack of knowledge of the population about this condition, even to perceive its associated signs and symptoms.

The strengths of our study include the work performed at night and on the weekends, which allowed for an interaction between supervisors and interviewers; the collaboration of the University by printing the questionnaires, assisting in the dissemination of the study and granting permission for the interviewers to use badges and uniforms as identification resources; the variety of topics surveyed; the use of probabilistic sampling to obtain a representative sample of the adult and elderly population of the urban area of Rio Grande; the methodological rigor in data collection and database creation; and the feedback sent to the study population.

Nevertheless, it is necessary to consider some limitations inherent to the cross-sectional study design, such as restriction to certain research topics [3] and the possibility of memory bias, in which interviewees fail to provide some answers. It should be emphasized that due to the large number of team members in a research consortium, there may be a limited space for each team’s questions so that the studied topics are not analyzed in-depth. Another limitation found in our study was the lack of measurements, such as weight and height, as well as clinical and biochemical tests, due to the difficulty in work logistics and limited financial resources.

Our recommendations for further studies are that interviewers should be paid based on their productivity. Here, the interviewers were paid a fixed monthly amount, which probably contributed to some of them failing to meet the established goals. Another recommendation is to collect information about freezer and dishwasher ownership instead of refrigerator and television (which are in almost 100% of the households) for the creation of the socioeconomic indicator named “the goods index” [16]. Lastly, it is also worth mentioning that a wide dissemination of the study is important to increase residents’ receptiveness.

In our study, we used information from the demographic census of the Brazilian Institute of Geography and Statistics carried out in 2010 and published in 2011. There were several changes in the urban scenario from that period up to the beginning of this survey. Some residential establishments did

Table 1. Description of the study topics with the corresponding number of responses, sample design effect and intraclass correlation coefficient. “Health of the Riograndina Population” Research Consortium (n = 1,300). Rio Grande, RS, Brazil, 2016.

<table>
<thead>
<tr>
<th>Study topic</th>
<th>Number of responses</th>
<th>Effect of sample design</th>
<th>Intraclass correlation coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Depression</td>
<td>1,295</td>
<td>1.70</td>
<td>0.022</td>
</tr>
<tr>
<td>Bruxism</td>
<td>1,280</td>
<td>1.08</td>
<td>0.001</td>
</tr>
<tr>
<td>Obesity</td>
<td>1,242</td>
<td>1.01</td>
<td>0.006</td>
</tr>
<tr>
<td>Nutrition counseling</td>
<td>1,296</td>
<td>1.35</td>
<td>0.017</td>
</tr>
<tr>
<td>Food insecurity</td>
<td>673*</td>
<td>1.49</td>
<td>0.049</td>
</tr>
<tr>
<td>Self-perception of healthy eating</td>
<td>1,297</td>
<td>1.32</td>
<td>0.011</td>
</tr>
<tr>
<td>Physical education classes</td>
<td>1,296</td>
<td>2.36</td>
<td>0.058</td>
</tr>
<tr>
<td>Chronic respiratory diseases</td>
<td>1,279</td>
<td>0.90</td>
<td>0.003</td>
</tr>
<tr>
<td>Practice of leisure physical activity</td>
<td>1,290</td>
<td>2.22</td>
<td>0.047</td>
</tr>
</tbody>
</table>

* The unit of analysis was the household.
not exist any longer and new ones had been built. Hence, it is recommended to identify the selected census tracts (a process known as “area hit”) so as not to compromise the representativeness of the sample. Another suggestion is to sort out the households by the code of the census tracts rather than by ordering the average monthly income of the head of the household. This recommendation is due to the fact that some census tracts with richer and poorer districts were overrepresented in the study sample.

The research consortium system is considered to be effective for the training of graduate students and for enabling a faster and cheaper execution of a project addressing several topics through a single research instrument. Undoubtedly, it is a useful resource to drive the implementation of evidence-based public health interventions and measures.

Taken altogether, the information presented in this methodological article may support the design and execution of other studies investigating health issues in a given population. Here, we addressed the steps, advantages and limitations of conducting a population-based survey through a research consortium involving graduate students with a common interest: to establish a diagnosis of the population health.

NOTES

Acknowledgements

Samuel C. Dumith is a research productivity fellow at the National Council for Scientific and Technological Development (CNPq, Brazil). Andressa C. Kretschmer, Marina X. Carpena, Milena O. Simões, and Monica F. R. Muraro received a masters scholarship from the Coordination for the Improvement of Higher Education Personnel (CAPES, Brazil).

Funding

This study was supported by the Foundation for Research Support of the State of Rio Grande do Sul (FAPERGS) (First Projects Program – ARD/PPP 2014, process nº 16/2551-0000359-9).

Conflicts of interest disclosure

The authors declare no competing interests relevant to the content of this study.

Authors’ contributions

All the authors declare to have made substantial contributions to the conception, or design, or acquisition, or analysis, or interpretation of data; and drafting the work or revising it critically for important intellectual content; and to approve the version to be published.

Availability of data and responsibility for the results

All the authors declare to have had full access to the available data and assume full responsibility for the integrity of the study results.

REFERENCES


