The ethics of in vivo calibrations in oral health surveys

Eticidade da calibração in vivo em inquéritos de saúde bucal

Flávia Reis de AndradeI, Paulo Capel NarvaiII, Miguel Ângelo MontagnerI

ABSTRACT: Objectives: To analyze the ethics of in vivo calibration, using the discourse of the administrators of the National Oral Health Survey (SBBrazil 2010) as a starting point. Method: This is a qualitative research involving semi-structured individual interviews with 12 members of the Steering Group and Technical Advisory Committee of the Ministry of Health, and two coordinators, one State and the other Municipal. The discourse of the collective subject technique was used for data analysis. Results: When asked about the experiences of SBBrazil 2010, which included ethical aspects, respondents identified the forms of standardization and training of teams who collected field data. For them, there is little scientific evidence to ethically support the way the training stage, including calibration, is carried out in oral health epidemiological surveys, as a certain unease can be predicted in participants of these studies. Conclusion: The ethics of a research also derives from its methodological rigor; the training process; and calibration in particular, is a fundamental technical and ethical requirement in surveys such as the SBBrazil 2010. The unease of the volunteers in face of test repetition does not ethically invalidate the in vivo calibration, but mechanisms to minimize it must be developed.

Keywords: Health surveys. Ethics. Training. Data collection. Validity of tests. Oral health.
**INTRODUCTION**

The conduction of health surveys is not an attribution of a single generation, but of many. This type of research produces results that lead to knowledge and reflection about the health status of populations in the present time, but also in the future. Therefore, this activity requires an observation from the referential point of view of ethics, which is understood as a study or reflection of human actions, embodied by daily situations that require determining what and how to do things\(^1\).\(^2\).

Even though the first national oral health survey was conducted in Brazil only in 1986\(^3\), the use of epidemiology in public dental services dates back to the 1950s, with the first school dental programs whose activities began, every year, with the execution of an epidemiological survey in each school where the program was developed. The survey was based on the organization and evaluation of planned actions\(^4\). After the pioneer initiative of 1986, the Brazilian Federal Government conducted national oral health epidemiological population surveys in 1996 and in 2003, and the latter was called National Oral Health Survey (SBBrasil), resulting in the Project SBBrasil. This epidemiological action led to an important accumulation of experience and to the consolidation of a methodology recognized in the country and abroad, resulting in several analyses expressing the important Brazilian scientific contribution in this field\(^5\).\(^6\). In 2010, these contributions led to the conduction of the National Oral Health Survey (SBBrasil 2010), the most recent national experience in the area. Professionals and researchers in the field consider the realization of another edition of Project SBBrasil in 2020.
The conduction of epidemiological surveys requires investments of different kinds, such as time and financial resources. In order to be ethical, they should produce knowledge that justifies these costs. In health services, these studies need to be valid enough to meet the needs of administration. The control of the temporal tendency of diseases and conditions is one of the main benefits resulting from health surveys. When regular and methodologically coherent, surveys allow analyzing the behavior of an epidemiological phenomenon of a certain period. This is something relevant, once, according to Burt (1997), public policies should be based on tendencies, and not on “statistical results” of a single survey. Therefore, the construction of a baseline for further observation is one of the reasons that justify the conduction of surveys. So, a health survey is questionable in terms of ethics when it does not produce data to be used as future parameters for comparison.

Surveys with an observation bias generate databases with many limitations, thus restricting the possibilities of several types of study. Therefore, they cause ethical problems. To prevent them, the objective is to reduce the sources of bias as much as possible. A strategic method in this sense is the analysis of agreement, known among specialists as “calibration.” The training researchers go through is a pre-requisite for any qualified survey. Even when inclusion or exclusion criteria are standardized and well known by examiners, they diverge when conducting tests under concrete conditions. These divergences occur between examiners and even between a specific examiner and him/herself, at different moments in time. Well-planned and executed studies know that these errors are part of surveys, therefore, they are inevitable. But such errors can and should be controlled. The way to control them is getting to know the dimension of these errors, measuring the error between examiners and intraexaminers, which can be done by using different statistics, such as Kappa and correlation coefficients, among others. Based on that, the objective of this article was to analyze the ethics of in vivo calibration based on the declarations of administrators of SBBrasil 2010. Cohen and Segre define ethics as the “aptitude to exercise the ethical function.” The objective was to contribute with the debate around a theme that is not so analyzed and object of concern by those who are in charge of oral health surveys.

**METHOD**

This study is part of a research about ethical problems and dilemmas that came up in SBBrasil 2010, conducted with administrators of this survey. This qualitative study involving the conduction of an individual interview, with semi-structured script, with members of the Managing Group (MG) of the survey. In SBBrasil 2010, each federation unit was assigned to one of the eight Control Group of the Ministry of Health in Oral Health Surveillance, who were in charge of the study that constituted the MG.

The members of this group were considered as key informants because they participated in the fieldwork, which is the phase of the survey that includes the largest number
of ethical facts. The interviewers were identified in the final report of SBBrazil 2010, which lists names and institutions of the participants of the MG. Besides, the “snowball” technique was also used. That is, at the end of each interview, the MG members indicated possible participants of the study. The suggestions included members of CTA-VSB (Steering Group and Technical Committee to structure and implement the strategy of oral health surveillance in the National Oral Health Policy) and state and municipal coordinators of the survey.

According to Lefèvre and Lefèvre\textsuperscript{11}, in cases in which the study population is limited, like this analysis, the researcher can “personally compose his/her sample, choosing all, or almost all individuals to be analyzed, according to the characteristics to be studied”. Based on that, 14 individuals were interviewed whose professional work in SBBrazil 2010, included 11 states — Goiás, Mato Grosso, Minas Gerais, Paraíba, Paraná, Pernambuco, Rio Grande do Norte, Rio Grande do Sul, Rio de Janeiro, Santa Catarina and São Paulo — and the Federal District. Nine of them composed the MG, and three were part of the CTA-VSB. Besides, two coordinators took part in the study, being one of the state and another one of the state. There was one refusal.

Interviewers chose the location, the date and the time of the interviews, which were recorded and transcribed. Total time of recording was approximately 14 hours and 30 minutes, which defines mean duration of 1 hour and 3 minutes for each interview. Data collection was conducted by a single researcher. The questions in the script were elaborated based on the theoretical referential and on the research hypotheses, but the calibration training was not considered in any of them. The question in the script that originated this study was: “Throughout your participation in Project SBBrazil 2010, did you experience or heard of any situation which, from your point of view, can be considered as an ethical dilemma?”

Given the reduced and peculiar universe of the study — initially restricted to the nine members of the MG, considered indispensable for the accuracy of the script, once the contents referred to specific themes of exclusive domain of this group of highly qualified experts — the pre-test of the data collection instrument with the two first interviewers was considered methodologically adequate. They were included as participants of the study. This option is based on the requirement of ensuring the presence of a set of possible opinions about the problem studied\textsuperscript{12}.

The answers resulted in 11 central ideas (CIs), but this study only approached one of them, since it is the only one referring specifically to the matter of calibration training. Three CIs were object of analysis in another article\textsuperscript{13}. The Discourse of the Collective Subject (DCS) technique was used for data analysis, which, based on the Theory of Social Representations, recommends gathering the individual discourses of the participants of the study to reveal the thought of a group. A sequence of the DCS established by Lefèvre and Lefèvre\textsuperscript{11} was adopted. The project of this study was approved by the Research Ethics Committee at the School of Health Sciences of Universidade de Brasília.
RESULTS

Of the interviewees, eight were female, and most were professionals related to state or municipal health secretariats (n = 5) or universities (n = 8). At the time of data collection, only one worked in the Ministry of Health. The set of lines of the interviewees, whose contents specifically referred to the calibration training, generated a DCS in the following terms:

“We have little scientific information discussing ethical issues about the forms of standardization and training of oral health teams. From the ethical point of view, to what extent is it respectful to examine and reexamine the same person? The experience of calibration is very tiring for the user, because he is assessed, analyzed, examined by at least ten dentists. In SBBrasil 2010, some mothers were tired, because the dentists examined the same child several times. But if we changed the child who was being examined, how would it be possible to calculate agreement? So, in order for some children to stay until the end of the training, they got some toothpaste from the coordinator, you know? They got snacks for the kids, but it was exhausting for them. They also got vacancies [for dental appointments in the Health Unit] so that they all could participate, especially those of five years ago, but they spent long periods waiting, because the examination goes wrong, the concordance is low, then it is necessary to retake the examination. After everyone examined [the same child], there is an analysis of agreement and divergence. Whenever there is doubt, the professionals go back to see the same child, to reassess the cases in which there were no deals. Then, comes the question: How can we have been working for so long with the same diseases and cannot see a homogeneous pattern of diagnosis? We always need calibration to level, to develop epidemiological surveys. So, the process of calibration is also part of ethics. How can we let five-year olds, and we know how they are, be examined ten times? If adults get bothered, let alone a child. So, what chances would we get? In the future, there could be another resource so that kids, in particular, would not have to stay [available for tests and retests for so long], because it is tiring.”

This central idea was generated based on the statement of three interviewees. It is worth to mention that the idea of the collective subject is not a result of the number of participants in the study, which generates a specific DCS, but of what the discursive formulation represents in the studied population, even if coming from a single participant.

DISCUSSION

Conducting epidemiological population surveys means considering ethical aspects of several orders, which are expressed in different dimensions of the investigation. The most frequently mentioned dimension is that referring to the consent of the participants in
the study, and is especially related with the stage of data collection during fieldwork. However, it is also important to mention the dimension including the stage of preparation and training of researchers who will go to the field. In this stage, people in charge of the survey are working with matters concerning the phenomenon known as “observation bias”\textsuperscript{14}, essential for the reliability of data obtained from the research participants, which together will compose the database of the survey. Based on that, different studies can be conceived and conducted.

According to Tauil\textsuperscript{15}, validity and accuracy are essential so that the conclusions of a study can be considered acceptable and, therefore, ethical. So, a valid and accurate study is, in theory, ethical. The validity of the results of an oral health survey depends mostly, but not exclusively, on the reliability of measurements\textsuperscript{16}. The calibration of the examiners is a resource that contributes with the reliability of the data, and therefore with the ethical aspect of epidemiological surveys. So, even if there is respect to human dignity, when the data produced in surveys are not reliable, the ethical dimension is compromised. It is not about stating, however, that once they are reliable, the data ensure this dimension, since they are a “sine qua non” condition, but not “enough condition to”. Even though the ethics of the preparation of field teams exists, there are questions about the in vivo calibration, as mentioned by the interviewees. The ethical dilemma experienced by the participants of surveys such as SBBrasil is owed to the fact that without calibration, bias is inevitable. But if there is calibration, even if in accordance with the best technique, the training causes discomfort, at several levels, on participants who collaborate with the study in this stage. Therefore, it is adequate to submit each experience that uses this strategy of team training to an ethical analysis.

In historic terms, since the 1950s efforts have been made in order to establish parameters, facilitate the conduction and allow the comparison of oral health epidemiological surveys\textsuperscript{17}, especially the documents published by the World Health Organization (WHO)\textsuperscript{16}. In this sense, the calibration training of SBBrasil 2010 was based on the best theory on the subject, emphasizing one of the publications of this agency, called calibration of examiners for oral health epidemiological surveys, from 1993. It suggests that the training of teams includes the performance of oral tests in individuals\textsuperscript{18,19}. The consensus technique, proposed by Frias et al.\textsuperscript{20}, was used.

Therefore, the consideration was that “each dental condition that has been registered differently in any of the retests was a matter of observation and discussion among the dentists, until they came to a conclusion”\textsuperscript{20}. Volunteers were then examined several times by all of the participating teams, and maintained in the location until professionals got to a consensus, which was reported by the interviewees as an ethical dilemma. About that, Allen and Steele\textsuperscript{21} point out that one of the main difficulties in the calibration of a large number of examiners is the tolerance of volunteers to the repetition of examinations. According to the authors, this situation is particularly a problem regarding data about periodontal conditions. For them, this provides an “ethical dimension” to the whole process of training and calibration.
In SBBrasil 2010, training with a theoretical and a practical part, was predicted to last 32 hours, being 8 periods of 4 hours each. Roncalli et al. registered that “in each workshop there were up to ten teams at the same time,” depending on the population of the city. In the DCS referring to this matter, they mention the fact that each volunteer was examined by at least ten dental surgeons. However, in the Manual of Examiners Calibration, the recommendation was to divide the teams in half. For instance, one capital containing the maximum number of teams per workshop, that is, team, would have two blocks of five in the practical part of the calibration exercise. However, there was no orientation establishing the examination of each volunteer by the two blocks of professionals as a condition. Therefore, in calibrations involving a large number of professionals, the ethical alternative is to conduct the epidemiological examination of each volunteer by only one of the blocks of teams. That could be seen as a difficult factor by some calibration instructors, since it would increase the number of volunteers to be selected, but it would be ethical, once it would contribute with the minimization of discomfort.

Facing the difficulties of in vivo calibration, there are the techniques of in lux and in vitro calibration, conducted, respectively, with the projection of images and exfoliated or extracted teeth. The former was used in SBBrasil 2010 to train for two conditions: fluorosis and dental trauma. Both are scientific validated techniques, but not always are they adequate for public health, especially for the epidemiological diagnosis of cavity, main oral health issue in Brazil; therefore, this condition is usually included in population surveys. Besides, about its limitations, reported by Agbaje et al., in vitro calibration is the technique that gets closer to the conditions that will be found by examiners during the fieldwork. The educational value contained in the act of doing — which includes, but is not restricted to, the act of seeing — is another differential from other techniques, especially the in lux. It was an important innovation brought to SBBrasil 2010, which tends to be maintained in the future editions of the project.

To choose one or another technique, it is always important to consider the discomfort and the benefit, because, as established in Resolution n. 466/12, “the ethics of the research implies on [among other aspects] searching whenever the benefits expected prevail over the risks and/or predictable discomforts.” In the case of SBBrasil 2010, national research financed with public resources, it is important to consider the cost-benefit notion. The methodology used in all national epidemiological surveys before SBBrasil 2010 was used again. Therefore, at the time of building a technical project, there was a set of scientific and empirical knowledge accumulated with time which allowed to assume, using the words of Kalamatianos and Narvai, that the “efforts made [would allow] reaching the results expected”. After all, in order for a public health action to be considered ethical, it is necessary to find evidence that it will meet the proposed objectives, therefore justifying the costs and risks or discomfort to which the patients will be submitted. So, considering the programmatic, institutional and political resources available, the in vivo calibration is an ethically acceptable option, as long as it is performed according to the technical guidelines, with adequate proportion of teste per team. The idea is
to minimize the discomfort so that the assumption of reasonability be respected in the balance of ethical principles.

The acceptance or rejection of something, from the point of view of ethics, implies considering the dimension of the individual, with the reference of their place in the world, necessarily in the inter-relationship of social life. So, ethically acceptable options, in the context of scientific investigation, require the simultaneous evaluation of what the actions of a specific study mean to the individual taking part in it, to the researchers and to society as a whole. The in vivo calibration is an ethically acceptable option in the epidemiological population survey analyzed in this article, once both the participant and the researchers and society found benefits. Besides, there was no violation of autonomy, and the scientific product resulting from the investigation (SBBrasil 2010) contributed, through its results, with the qualification of public health actions. This aspect meets the ethical requirement of justice in the scientific research, since the benefits are not restricted to one or another population group, but to everyone. Nevertheless, a similar analysis is suggested with volunteers from the calibration training.

In Brazil, oral health epidemiological studies have been improving the methods and techniques used, since the first studies were conducted in the middle of the 20th century. Since then, the successive standardizations performed by WHO have been used in our country, in order to improve the epidemiological instruments and the conditions in which they are applied. The stage of calibration training, in the processes of planning and organizing population surveys, have benefitted from these improvements. Brazilian epidemiological surveys have been incorporating this stage almost without exceptions, and, for that, the initiatives of Permanent Health Education, made stronger in Brazil after the creation of the Unified Health System, have been important bringing specific policies to this objective. The fact that SBBrasil 2010 was executed with the participation of professional teams from SYS, from states and cities, is a powerful indicator of this statement. It is not an overstatement to say that studies as broad and deep as SBBrasil 2010, with institutions from SUS, would not be possible without the participation of these teams. It is also worth to mention the fact that the professionals who take part in the population survey improve their knowledge and create experiences that are totally related with the purposes of Permanent Health Education policies. In the specific case of SBBrasil 2010, it is important to mention the strategic role of research support played by higher education institutions, especially public ones, but also some private ones.

**CONCLUSION**

In a survey, ethics is not only about the approval of a project by the Brazilian system of research ethics review with human beings. From the political and institutional decision to perform a study until the publication of data, all stages invariably have an ethical
The standardization of the epidemiological diagnosis is essential to provide reliability in the data produced by oral health surveys. The quality of the training and the preparation of field teams is decisive to guarantee the relevance of conclusions of the survey itself, and of the several analyses based on its database. Therefore, if ethics in research is also a result of its methodological rigor, the process of training and calibration constitutes basic technical and ethical requirements, as observed in surveys such as SBBrasil 2010. The discomfort of the volunteers caused by the repetition of examinations does not discredit \textit{in vivo} calibration in terms of ethics, but it is necessary to establish mechanisms to minimize it.

\textbf{REFERENCES}


Received on: 11/18/2015
Final version presented on: 04/25/2016
Accepted on: 05/16/2016