Revista Brasileira de Farmacognosia
BRAZILIAN JOURNAL OF PHARMACOGNOSY
www.journals.elsevier.com/revista-brasileira-de-farmacognosia/

Original article

Sampling problems in Brazilian research: a critical evaluation of studies on medicinal plants

Patrícia M. Medeiros[a,], Ana H. Ladio[b], Ulysses P. Albuquerque[c,*]*

[a]*Grupo de Etnobiologia e Ecologia Humana, Instituto de Ciências Ambientais e Desenvolvimento Sustentável, Universidade Federal do Oeste da Bahia, Barreiras, BA, Brazil*
[b]*Laboratorio Ecotono, Instituto de Investigaciones en Biodiversidad y Medio Ambiente, Universidad Nacional del Comahue, Centro Regional Universitario Bariloche, San Carlos de Bariloche-Río Negro, Argentina*
[c]*Laboratório de Etnobiologia Aplicada e Teórica, Departamento de Biologia, Universidade Federal Rural de Pernambuco, Recife, PE, Brazil*

ARTICLE INFO

Article history:
Received 24 October 2013
Accepted 3 January 2014

Keywords:
Cientometrics
Research design
Ethnobotany
Ethnopharmacology
Methodological aspects

ABSTRACT

This work compiled Brazilian articles regarding medicinal plant use by local communities in order to analyze the most common sampling problems and if research characteristics can influence the presence of sampling irregularities. We focused on studies about medicinal plants that present a species-indications list and had a quantitative nature. The proportion of works with and without sampling problems was evaluated considering the journal impact factor, period of publication, community status (urban x rural), sample type, presence of testing hypothesis and presence of research questions. We found that an alarming proportion of papers had some kind of sampling problems (48.39% serious and 19.35% moderate). The most common problems were related to: lack of information regarding the sample size or the universe, small sample sizes and selection of specialists based on obscure criteria. We could not find a significant influence between our tested variables and the occurrence of sampling problems, except for the community status (urban x rural). Results indicate that a significant amount of intracultural diversity is not properly captured, taking into consideration both the population as a whole and a group of interest in the community (= healers).

Introduction

The use of sampling techniques based on the Hypothetico-deductive method is common in scientific investigations, mainly due to the difficulty or cost of dealing with the whole universe (U). The proper use of these techniques allows the researchers to make conclusions about a population based on just a part of it. Furthermore, sampling can be applied to situations whose research objectives are not to draw a general profile for the entire universe, but rather to deal with specific components of this universe.

However, sampling misuse and negligence on the principles behind its application are common in scientific research (Bartlett et al., 2001; Albuquerque and Hanazaki, 2009). In fields

* *Corresponding author.*
E-mail: patricia.muniz@ufob.edu.br (P.M. Medeiros); upa@db.ufrpe.br (U.P. Albuquerque).

http://dx.doi.org/10.1016/j.bjp.2014.01.010

that deal with interviews or information about humans, the scenario is not different, as evidenced by some studies that evaluate and discuss sampling and its problems (Marks, 1951; Kitson et al., 1982; Malhotra et al., 1996; Woodberry, 1998; Bartlett et al., 2001; Lee, 2010).

Accordingly, this research contributes to the discussion of sampling issues regarding the context of medicinal plants research, in order to evaluate the Brazilian studies with a quantitative approach. We sought to identify recurring sampling problems in these studies. We also intend to verify if research characteristics influence the occurrence of sample problems. The appearance of sampling problems can have important consequences concerning bioprospecting studies based on an ethnodirected approach.

Questions, hypothesis and their explanations

The following questions and hypotheses were formulated:

a) *Are more recent publications more concerned about sampling quality than older publications?* Hypothesis: More recent publications present less sampling problems than older ones. This is expected since literature has increased, the number of manuals that offer methodological support for ethnobotanical and ethnopharmacological research (Oliveira et al., 2009). In fact there was an increase in the number of Brazilian and international publications that deal with sampling issues in this scientific fields or correlated fields (Bernard, 2006; Albuquerque et al., 2014).

b) *Does the journal's impact factor influence the presence of sampling problems?* Hypothesis: studies published in journals with higher impact factors apply a higher sampling quality. This might be true since high-impact journals are considered to publish studies with meticulously performed sampling design.

c) *Do studies developed in urban areas have proportionally more sampling problems than studies developed in rural areas?* Hypothesis: Studies from urban areas have proportionally more sampling problems. We drew this hypothesis since urban areas are often more populated, they need a larger use of sample strategies, given that it is not always possible to interview the entire population

d) *Do the different types of informant's selection (purposive, random or mixed) present differences to what concerns sampling quality?* Hypothesis: sampling quality is higher for purposive (intentional) samples. We believe in this assumption because random samples often mean interviewing more people, which is more difficult to reach.

e) *Do studies that test hypothesis present less sampling problems than studies that do not test them?* Hypothesis: studies that explicitly test hypothesis are more concerned with sample quality and present less sampling problems.

f) *Do studies that present a clear research question have less sampling problems than studies that do not present it?* Hypothesis: studies that present a clear research question are more concerned with sample quality and present less sampling problems.

Sampling in qualitative and quantitative research

In most cases, the use of sampling techniques in quantitative research is associated with hypothesis testing, search for behavior patterns and comparisons within a study or with other studies. In these cases, the sample size is determined by the number required to allow valid inferences about the population or group of interest (Marshall, 1996). Therefore, a good sampling design is an essential step to reach objectives linked to the search for patterns and trends. The misuse of sampling can constitute a source of bias regarding acceptance of hypotheses that should be rejected and vice versa, in addition to hiding behavior patterns and inappropriately capturing the internal diversity of a given universe (Freiman et al., 1978; Bartlett et al., 2001).

Sample size is not the only requirement to ensure representativeness. It is necessary to follow the principles of randomness, so that the sample is not biased by a particular group. These assumptions are usually considered in quantitative research aimed at generalizing findings to the universe.

The qualitative investigations, in turn, have some particular features such as frequent use of theoretical samples, and other types of intentional samples (Marshall, 1996). Qualitative studies often assume that some cases (or respondents) provide "richer" information than others, and so the election of them would increase the researcher's ability to understand a given scenario (Marshall, 1996). In many cases, small samples are used in qualitative studies in order to allow more detailed information, so that this approach does not intend to generalize findings (Marshall, 1996; Curtis et al., 2000). Despite the peculiarities of qualitative research, many authors argue that this approach is not free from evaluations of sample quality. The sample choice must be consistent and explicit about the research objectives (Curtis et al., 2000; Dixon-Woods et al., 2004). However, as sampling in qualitative studies cannot be evaluated in quantitative terms, we decided to focus this work only on quantitative investigations, which clearly follow a positivist orientation.

Statistical assumptions in quantitative ethnobotany

According to the statistical assumptions, a sample is considered to be representative when it is randomly chosen and the probability of misrepresenting the universe is 5% or less (Bernard, 2006). Problematic sampling designs can lead the internal diversity not to be properly captured and it can be a source of bias to conclusions on, for example, the distribution of botanical knowledge within a community.

In ethnobotanical sampling, it is common to consider the community members (total or > 18 years) as the sample unit for generalization, as well as the heads of family (men and/

or woman responsible for the residence) or the residences (in the latter case, one or more people per residence would give information about the use of this plant). The rarefaction curve technique is used (Begossi, 1996) to verify the sample adequacy without the need for a representative N value of U considering 5% error. In this case, the number of informants interviewed is considered sufficient when there is stabilization of the curve, so that most plants and uses assigned to them have already been cited.

Moreover, it is also common in the ethnobotanical research the use of intentional samples. It can reduce the total timeof study by directly seeking relevant people to the query. Therefore, the bias inherent to this sampling method contributes to its efficiency (Tongco, 2007), although the conclusions cannot be extrapolated to the community as a whole, the same way that a probabilistic sample performed with a sample size lower than the necessary would not be considered as representative. An example of appropriate use of non-probabilistic sample consists of interviews with key informants to get information about plants with bioactive potential, dealing directly with people who have the most knowledge on the subject. Therefore, the application of an intentional sample can serve specific approaches and is very interesting when we need to study in detail a specific group, but its characteristics do not allow generalizations for the entire community since it covers only a portion of it. An example of the misuse of an intentional sample is to interview only local specialists and consider their knowledge as representative of the whole community, given that a precise diagnostic would identify people less knowledgeable.

Methods

Data collection

In this study literature search of articles was performed. In many cases the studies do not explain their epistemological orientation. Therefore, in order not to make the mistake of classifying the studies' sample quality, regardless their epistemological affiliations, only quantitative studies were considered; as the use of quantitative tools aims to generalize findings for a population or group of interest (the last one with the help of intentional samples) and requires a representative or coherent sample. We considered as quantitative those studies which calculated ethnobotanical indices or used statistical tests.

Our corpus included original works on medicinal plants that had a list of known and/or used[1] species with their therapeutic indications. It was considered for this study both articles exclusively on medicinal plants and studies with various uses (edible, construction, fodder, fuelwood etc.), including the medicinal one.

The literature search was conducted at the following databases and publishers: Scielo (www.scielo.org), Scopus (www.scopus.com) and Scirus (www.scirus.org). The keywords "Medicinal Plants + Brazil", "Plantas Medicinais + Brasil" "Ethnobotany + Brazil" and "Etnobotânica + Brasil" were used. Literature search was performed until 2011.

In addition to this database search, a more detailed search was performed using the keyword "Brazil" in the journals "Ethnobotany Research and Applications", "Economic Botany", "Journal of Ethnopharmacology" and "Boletin Latinoamericano de Plantas Medicinales y Aromáticas". For the journals "Acta Botanica Brasilica", "Revista Brasileira de Plantas Medicinais" and "Revista Brasileira de Farmacognosia", the keyword "Etnobotânica" was used. For articles not available online we made an effort to acquire them by e-mailing authors and using library services. In this way, we believe that most of the articles have been addressed in this study. To ensure the entry of a significant portion of works, a revision of references was made of each paper in search of new entries that were not available online, or that could not be found with the keywords used.

For each study obtained and included in our *corpus*, the following information was recorded: decade of publication, journal's impact factor (according to JCR 2011), origin of journal (national or international), place where the study was carried out (urban or rural area), type of sample (intentional or probabilistic), universe (U - total number of people, residences, heads of family, vendors etc. in the community or another place), number of respondents (N), nature of the study nature (quantitative or qualitative) and if the study formulated questions or hypotheses.

Although it is common to consider urban and rural as a continuum, we opted to perform a dichotomous classification in order to make the analysis possible. We used the information available in the articles as a parameter.

A sample was considered as random (or probabilistic) when all the members of a given community or community's strata had the same probability of being recruited to participate. It was considered as purposive (or intentional) when respondents where chosen regardless of a random selection. Finally, a mixed sample was considered when the research included two or more different sampling technique, being at least one of them random and one purposive. More information about sampling strategies can be found in Albuquerque et al. (2014).

Quantitative studies were so considered when they presented at least one of the following: statistical testing or application of ethnobotanical indices. Only quantitative studies were included in this survey (more information about these indices can be found in Medeiros et al., 2012). For case studies that fit into more than one classification (e.g. urban and rural), or do not have information on this, we decided to exclude them from the analysis of the referred category, considering them, however, for remaining analyzes.

Studies were classified, according to the risk of bias in their samples, into: low risk, moderate risk and high risk. The criteria for classification are available in Chart 1. In some cases the study fit into more than one criterion. Table 1 clarifies the classification procedures in such situations.

[1] Some plants may have their medical use known to the people of a community without necessarily being used by them.

Chart 1
Criteria to establish the risk of bias in ethnobotanical studies regarding medicinal plants performed in Brazil based on sampling quality.

1) When sample is extracted from the total number of people or from an age interval
Low
a) When the sample size (N) reaches the universe (U)
b) When N is representative of U, with sample randomness and considering a margin of error of up to 5%
c) When N is at least 80% of U, considering that some respondents may refuse to participate of interviews or may not be in their household even after successive trials.
Moderate
a) When N is extracted from U, with sample randomness and a margin of error higher than 5% and lower than 10%.
b) When N is at least 80% of the value is needed for representativeness, considering a margin of error of up to 5%.
c) When N could be considered to be representative of U (with a margin of error of up to 10%) if only the numbers are considered, but in situations in which sample is occasional or when there is no specificity about randomness.
High
a) When N is extracted from U with a margin of error higher than 10%.
b) When N is less than 80% of the value is necessary for representativeness, considering a margin of error of up to 5%.
c) When there is no information about the universe (U), or when there is no information about the sample (N).

2) When sample is based on the heads of family (one or two per household)
Low
a) When all heads of family were interviewed.
b) When a representative number of heads of family were interviewed, with sample randomness and a margin of error up to 5%.
c) When N is at least 80% of the heads of family
Moderate
a) When N is extracted from the universe (U) of family chiefs, with sample randomness and a margin of error higher than 5% and lower than 10%.
b) When N is at least 80% of the value necessary for the representativeness of the heads of family, considering a margin of error of up to 5%.
c) When N could be considered to be representative of the number of heads of family (with a margin of error of up to 10%), if only the numbers are considered, but in situations which sample is occasional or when there is no specificity about randomness.
High
a) When N is extracted from the universe (U) of heads of family with a margin of error of up to 10%
b) When N is less than 80% of the value needed for representativeness of the family chiefs, considering a margin of error of up to 5%.
c) When there is no information about the number of heads of family, or when there is no information about the sample (N).

3) When sample is based on households
Low
a) When one member of each household was interviewed.
b) When a representative number of households had one of its members interviewed, with sample randomness and a margin of error up to 5%.
c) When N is at least 80% of the households.
Moderate
a) When N is extracted from the universe (U) of households, with sample randomness and a margin of error higher than 5% and lower than 10%.
b) When N is at least 80% of the value necessary for sample representativeness, considering a margin of error of 5%.
c) When N could be considered to be representative of the households (with a margin of error of up to 10%) if only the numbers are considered, but in situations in which sample is occasional or when there is no specificity about randomness.
High
a) When N is extracted from the universe (U) of households with a margin of error higher than 10%.
b) When N is less than 80% of the value which is necessary for household representativeness, considering a margin of error of up to 5%.
c) When there is no information about the number of households (U), or when there is no information about the sample (N).

4) When sample is intentional, focusing on a group of interest (e.g. midwives, healers, local specialists)
Low
a) When sample corresponds to the totality of the specific group.
b) When sample is representative of the specific group, with sample randomness and a margin of error or up to 5%).
c) When sample is at least 80% of the specific group.
d) In cases of local specialists, when the snowball technique is used and there is an indication of the total number of dwellers.
e) In cases of local specialists, when they are selected based on clear and well established criteria.
Moderate
a) When N is extracted from the universe (U) of the specific group, with sample randomness and a margin of error higher than 5% and lower than 10%.
b) When N is at least 80% of the value necessary of representativeness of the specific group, considering a margin of error of up to 5%.
c) When N could be considered to be representative of the specific group (with a margin of error of up to 10%) if only the numbers are considered, but in situations which sample is occasional or when there is no specificity about randomness.
d) In cases of local specialists, when there is no indication of the universe (U), but the snowball technique is applied to select the key respondents.
High
a) When N is extracted from the universe (U) of the specific group with a margin of error higher than 10%.
b) When N is less than 80% of the value which is necessary for representativeness of the specific group, considering a margin of error of up to 5%.
c) When there is no information about the specific group (U), or when there is no information about the sample (N), except for the use of the snowball technique, when there is no information about the U.
d) In cases of local specialists, when they are selected based on arbitrary or obscure criteria.

5) When rarefaction curves are used
Low
a) When there is information about N and U and when the rarefaction curve stabilizes, regardless of sample representativeness and the criteria for respondent selection.
Moderate
a) When there is no information about U, but the rarefaction curve stabilizes
b) When there is no information about N and U and when the rarefaction curve gets close to stabilization.
High
a) When there is no information about N, regardless of the rarefaction curve behavior.
b) When the rarefaction curve gets far from stabilization.
c) When the study claims to have performed a rarefaction curve, but does not exhibit its results and does not affirm that there was stabilization.

6) When participatory methods are used
Low
a) When the number of participants corresponds to a representative amount of the population or specific group (with a margin of error of up to 5%, but not considering the precepts of randomness, that mostly does not apply to participatory methods).
Moderate
a) When the number of participants is not representative of the population or specific group.
b) When there is no information about the universe (population as a whole or specific group), but there is information about the number of participants.
High
a) When there is no information about the number of participants.

7) Diffuse selection criteria
High
a) When there is no information about N or U.
b) When there are various diffuse criteria for selecting the same sample.

Table 1
Prevalence criteria in situations when there is an overlap of classifications according to the risk of bias for ethnobotanical studies about medicinal plants developed in Brazil.

Situation	Prevalence	Example
A study is classified by two or more routes and all of these routes lead to the same risk of bias.	All classifications are considered	4–High–c and 4–High–d *Prevalence of a high risk*
Study with composite samples (for more than one place or more than one type of respondent).	More restrictive criterion	1-Moderate-a and 1-High-a *Prevalence of a high risk*
Prevalence of a high risk		
Studies with rarefaction curves	Criteria for rarefaction curve	1-Moderate-a and 5-Low-a *Prevalence of a low risk*

Data analysis

The analytical strategy of this study was based on comparing the proportion of studies with high, moderate and low risk of bias according to the following variables: decade of publication, presence/absence of an impact factor, place where the study was developed, type of sample, presence/absence of hypothesis testing, presence/absence of explicit question formulation of ethnobotanical indices. These analyses were performed using the Williams G-test (Zar, 1996), in 2xn and 3xn contingency tables. Studies with high risk of bias x studies with moderate risk of bias x studies at low risk of bias were integrated in 3xn tables. Studies with some risk of bias x studies with no risk of bias were considered in 2xn tables.

The influence of journal's impact factor on the risk of bias (present x absent) was tested with a simple logistic regression.

Results and discussion

General aspects about sample quality

We found 126 ethnobotanical studies on medicinal plants that exhibited a species-indication list. However, only 31 could be considered as quantitative, which is why our results will only refer to those 31 studies.

Among them, 48.39% were classified as having high risk of bias, while 19.35% had moderate risk of bias and only 32.26% had low risk of bias. This result demonstrates that the majority of quantitative Brazilian ethnobotanical studies devoted to medicinal plants are not concerned about the representativeness of their samples or do not make clear the criteria used for selecting informants. Regarding the generalization issues, it is likely that the internal diversity of knowledge, beliefs, or use of medicinal plants is not being properly captured and, considering studies that adopt intentional samples, these can be made arbitrarily or the criteria adopted may not be clear in the article.

The main reasons for works to be classified as having high risk of bias were: (1) when N is extracted from U with a margin of error higher than 10% (1-high-a for 16.13% of studies); (2) when N is less than 80% of the value necessary for representativity, considering a margin of error of up to 5% (1-high-b for 16.13% of studies), and (3) lack of information about the universe or the sample, for samples taken from the total number of people or an age interval (1-high-c for 16.13% of studies) (Table 2). The main reason for a study being fitted as

Table 2
Percentage of studies, according to the categories, to define the risk of bias for ethnobotanical studies about medicinal plants developed in Brazil.

Qualification	% of studies
1-High-c	16.13
1-High-a	16.13
1-High-b	16.13
4-Moderate-d	9.68
4-High-d	12.9
4-High-a	6.45
4-High-b	6.45
3-Low-a	6.45
4-Low-d	6.45
5-Low-a	6.45
4-High-c	6.45
1-Moderate-c	3.23
3-Moderate-a	3.23
1-Low-b	3.23
1-Low-c	6.45
5-Moderate-b	3.23
3-Low-c	6.45
7-High-b	3.23
2-Low-c	3.23
1-High-c	16.13
1-High-a	16.13
1-High-b	16.13
4-Moderate-d	9.68

having moderate risk of bias was: (1) Lack of indication of the universe (U) when the snowball technique is applied to select the key informants (4-moderate-d for 9.68% of studies). The other reasons for a study to be classified as having a moderate risk had less than 5% of occurrence.

Regarding studies that reached a low risk of bias, most of them were classified by having encompassed the whole universe (one person per household), 80% of the universe or by having adequately used the snowball technique (Table 2). This result contributes to the acknowledge of sample fragilities in Brazilian ethnobotanical studies, since many studies were classified as having a low risk of bias just because they did not "need" to use complex sampling techniques.

It is possible that not always a study classified as having high or moderate risk of bias presents in fact samples issues. Some studies do not present information about the type of sampling, about how it was done or under which basis (universe) it was extracted. In this sense, Kitson (1982, p. 966) states that *"researchers invariably seem to find it more interesting and exciting to write about what was learned from a study than how it was learned"*. Sometimes the simple attention to the adequate description of the research design can save the study from "distrusts" related to the representativeness or coherence of the sample. This is related to the positivist need for precise information about the research design in order to make it testable and replicable. For other scientific paradigms, such evaluation and replicability are not possible because it is assumed that the nature of interpretation is influenced by the researcher subjectivity. Therefore, the evaluation of sample quality only makes sense from a hypothetico-deductive perspective.

Studies profile x risk of bias

Factors that do not influence the risk of bias: most of our hypothesis were rejected, since most factors analyzed showed no significant influence on the study's risk of bias. The proportion of studies with high, moderate and low risk of bias, as well as the proportion of studies with high risk and other risks, did not differ significantly to what concerns the decade of publication, the presence of an impact factor, type of sample and the presence/absence of questions and hypotheses ($p > 0.05$ for all possible combinations).

Because of the increase in the number of textbooks in the area, we expected that more recent works would have less sampling problems when compared to the oldest publications. However it seems that an increase of the "offer" in terms of access to information about research design quality was not accompanied by an increase in the attention to sampling aspects, since more recent studies tend to have the same sampling problems presented in the older ones.

Furthermore, the publication of a work in a journal with impact factor does not assure that it will be free from sampling problems, as shown by the logistic regression ($p > 0.05$). This result offers additional support to the advocates for the impact factor not be used to evaluate the study's quality (Seglen, 1997).

Another important result is the lack of differences in terms of sampling problems when comparing purposive, random and mixed samples. It means that there is no type of sampling problem-free. The assumptions and applicability of a random sample are quite different from an intentional sample, since the former is related to a greater statistical guidance and the latter has to do with interests in specific community groups. However, none of them is being properly employed.

It is also possible to infer that questions or hypotheses may not be answered or tested satisfactorily, since studies with the characteristic of generating questions and hypotheses do not differ from the others regarding the sample quality.

Factors that influence the risk of bias: G test in 2x2 table (rural x urban areas regarding the proportion of risk x no risk of bias) showed significant differences, so that the risk of bias is proportionally more present in studies developed in urban areas ($G = 4.36$, $p < 0.05$). This result can be explained by the fact that urban communities often present a greater number of people and households. Therefore, they need a more complex sample design, which can lead to a larger amount of errors. Additionally, many studies developed in rural areas do not need to employ samples, as all adults or family chiefs can be interviewed due to the small population size.

The G test for 3x2 did not indicate significant differences in the proportion of studies with high risk of bias x moderate x low risk of bias according to this factor ($G = 4.4$, $p > 0.05$).

Final remarks

The high occurrence of sampling problems in quantitative studies of medicinal plants (which have lists of species and indications) reveals problems capturing the internal diversity of knowledge, beliefs and practices, and also points to a misappropriation of intentional sampling techniques. Moreover, the lack of information on the sampling procedure is also quite evident, and among all the problems listed here, it is the most easily solvable, since the inclusion of this information can free the study of suspicions concerning its sample design.

Those sampling problems can bias results and influence decisions and proceedings that involve plant knowledge and use. This is especially problematic in the field of biodiversity conservation, as an inadequate sampling can give misleading clues about the conservation status of medicinal plants and the strategies needed to be placed. Inappropriate samples can also give misleading clues about the most popular plant species for a given community, which are usually chosen species for pharmacological studies.

We assume that a study that follows a positivist orientation may forgo effective sampling techniques, since their conclusions do not intend to be extrapolated to a community or a particular group (*e.g.* local experts) but only applied to those interviewed, which is not usually part of the rhetoric of those studies. Therefore, the easiest way to solve the problem is to assure the coherence of research design.

We must consider that this study brought a diagnosis based on a *corpus*: quantitative studies about medicinal plant knowledge or use that display lists of plants with their therapeutic indications. The results of the diagnosis may only be applicable to this study group. Therefore, the development of further studies is recommended in order to diagnose the status of ethnobotanical studies to what concerns the controversial issue of sampling.

Finally, although our study focused only on sampling problems, we consider that other issues can significantly bias research results, such as the data collection procedures and botanical identification. Identifying specimens from ethnobotanical surveys is not easy, since many of those specimens are cultivated and sterile. As misidentification can completely alter research results, we suggest that further investigation should focus on this issue.

Authors' contributions

PMM contributed in idealizing/drawing the study, collecting and analyzing data and writing the article. AHL contributed in idealizing/drawing the study and correcting the manuscript, UPA contributed in idealizing/drawing the study and correcting the manuscript.

Conflicts of interest

The authors declare no conflicts of interest.

Acknowledgment

The authors thank CNPq for financial support (471989/2012-6), the fellowship given to PMM and for the productivity grant given to UPA.

REFERENCES

Albuquerque, U.P., Hanazaki, N., 2009. Five problems in current ethnobotanical research - and some suggestions for strengthening them. Hum. Ecol. 37, 653-661.

Albuquerque, U.P., Lucena, R.F.P., Cunha, L.V.F., Alves, R.R.N., 2014. Methods and Techniques in Ethnobiology and Ethnoecology. Springer.

Bartlett, J.E., Kotrlik, J.W., Higgins, C.C., 2001. Organizational research: determining appropriate sample size in survey research. Inform. Technol. Learning Perform. J. 19, 43-50.

Begossi, A., 1996. Use of ecological methods in ethnobotany: diversity indices. Econ. Bot. 50, 280-289.

Bernard, R., 2006 Research methods in anthropology: qualitative and quantitative approaches. 4. ed. Oxford: AtlaMira Press.

Curtis, S., Gesler, W., Smith, G., Washburn, S., 2000. Approaches to sampling and case selection in qualitative research: examples in geography of health. Soc. Sci. Med. 50, 1001-1014.

Dixon-Woods, M., Shaw, R.L., Agarwal, S., Smith, J.A., 2004. The problem of appraising qualitative research. Qual. Saf. Health Care 13, 223-225.

Freiman, J.A., Chalmers, T.C., Smith Jr., H., Kuebler, R.R., 1978. The importance of beta, the type II error and sample size in the design and interpretation of the randomized control trial - survey of 71 negative trials. New Engl. J. Med. 299, 690-694.

Kitson, G.C., Sussman, M.B., Williams, G.K., Zeehandelaar, R.B., Shickmanter, B.K., Steinberger, J.L., 1982. Sampling issues in family research. J. Marriage Fam. 44, 965-981.

Lee, B., 2010. Exploring a new research method in diversity research. Procedia - Social and Behavioral Sciences 7, 494-503.

Malhotra, N.K., Agarwal, J., Peterson, M., 1996. Methodological issues in cross-cultural marketing research: a state-of-art review. Int. Market Rev. 13, 7-43.

Marks, E.S., 1951. Some sampling problems in educational research. J. Educ. Psychol. 42, 85-96.

Marshall, M., 1996. Sampling for qualitative research. Fam. Pract. 13, 522-525.

Medeiros, M.F.T., Silva, O.S., Albuquerque, U.P., 2012. Quantification in ethnobotanical research: an overview of indices used from 1995 to 2009. Sitientibus. Série Ciências Biológicas 11: 211-230.

Oliveira, F.C., Albuquerque, U.P., Fonseca-Kruel, V.S., Hanazaki, N., 2009. Avanços nas pesquisas etnobotânicas no Brasil. Acta Bot. Bras. 23, 590-605.

Seglen, P.O., 1997. Why the impact factor of journals should not be used for evaluating research. BMJ 314, 498-502.

Tongco, M.D.C., 2007. Purposive sampling as a tool for informant selection. Ethnobot. Res. Appl. 5, 147-158.

Woodberry, R.D., 1998. When surveys lie and people tell the truth: how surveys oversample church attenders. Am. Sociol. Rev. 63, 119-122.

Zar, J.H., 1996. Biostatistical analysis. 3. ed. New Jersey: Prentice-Hall International Editions.