Use of IRAMUTEQ software in qualitative research: an experience report

Utilização do *software* IRAMUTEQ em pesquisa de abordagem qualitativa: relato de experiência

Uso del software IRAMUTEQ en investigación cualitativa: informe de experiencia

Fernanda Garcia Bezerra Góes[I], Andressa Silva Torres dos Santos[II], Brenda Lucas Campos[III]
Aline Cerqueira Santos Santana da Silva[IV], Liliane Faria da Silva[V],
Luiz Carlos Moraes França[VI]

Abstract: **Objective**: to report the experience of using the software *Interface de R pour les Analyses Multidimensionnelles de Textes et de Questionnaires* (IRAMUTEQ) as a computerized tool to support the analysis of textual data in qualitative research in health. **Method**: this was an experience report that took place between March and June 2019, based on qualitative research developed through the Creative-Sensitive Method and using IRAMUTEQ. Newborns and family members of newborns from a public hospital located in the city of Rio das Ostras in the state of Rio de Janeiro, Brazil, participated in the study. **Results**: it was possible to verify that the different lexical analysis techniques available in IRAMUTEQ allowed to discriminate the semantic contents referring to the caregivers' claims about the care to be performed with the newborn at home. **Conclusion**: this computerized tool proved to be useful, adequate, and reliable, qualifying the textual analysis process.
Descriptors: Research; Nursing research; Qualitative research; Methodology; Software

Resumo: **Objetivo**: relatar a experiência de utilização do *software Interface de R pour les Analyses Multidimensionnelles de Textes et de Questionnaires* (IRAMUTEQ) como ferramenta informatizada de apoio à análise de dados textuais em uma pesquisa qualitativa em saúde. **Método**: tratou-se do relato de experiência acontecida entre março e junho de 2019, a partir de uma pesquisa qualitativa desenvolvida por meio do Método Criativo-Sensível e com o uso do IRAMUTEQ. Participaram puérperas e familiares de recém-nascidos de uma instituição hospitalar pública localizada na cidade de Rio das Ostras no Estado do Rio de Janeiro, Brasil. **Resultados**: foi possível constatar que as diferentes técnicas de análise lexical disponíveis no IRAMUTEQ permitiram discriminar, claramente, os conteúdos semânticos referentes às pretensões dos cuidadores sobre os cuidados a serem desempenhados com o recém-nascido em casa. **Conclusão**: essa ferramenta informatizada mostrou-se útil, adequada e fidedigna, qualificando o processo de análise textual.
Descritores: Pesquisa; Pesquisa em enfermagem; Pesquisa qualitativa; Metodologia; Software

[I] Nurse. Doctor. Universidade Federal Fluminense, Rio das Ostras, RJ, Brazil. ferbezerra@gmail.com https://orcid.org/0000-0003-3894-3998

[II] Nurse. Universidade Federal Fluminense, Rio das Ostras, RJ, Brazil. torresandressa@hotmail.com https://orcid.org/0000-0001-7142-911X

[III] Nursing Student. Universidade Federal Fluminense, Rio das Ostras, RJ, Brazil. brendalcampos@hotmail.com https://orcid.org/0000-0001-5832-3943

[IV] Nurse. Doctor. Universidade Federal Fluminense, Rio das Ostras, RJ, Brazil. alinecer2014@gmail.com https://orcid.org/0000-0002-8119-3945

[V] Nurse. Doctor. Universidade Federal Fluminense, Niterói, RJ, Brazil. lili.05@hotmail.com https://orcid.org/0000-0002-9125-1053

[VI] Nurse. Master. Universidade do Estado do Rio de Janeiro, Rio de Janeiro, RJ, Brazil. lcmoraesfranca@hotmail.com http://orcid.org/0000-0002-6370-115X

Resumen: **Objetivo**: reportar la experiencia del uso del software *Interface de R pour les Analyses Multidimensionnelles de Textes et de Questionnaires* (IRAMUTEQ) como una herramienta computarizada para apoyar el análisis de datos textuales en una investigación cualitativa en salud. **Método**: se trata de un relato de experiencia que se llevó a cabo entre marzo y junio de 2019, con base en una investigación cualitativa desarrollada a través del Método Creativo-Sensible y utilizando IRAMUTEQ. Participaron recién nacidos y familiares de recién nacidos de un hospital público ubicado en la ciudad de Rio das Ostras en el estado de Rio de Janeiro, Brasil. **Resultados**: se pudo constatar que las diferentes técnicas de análisis léxico disponibles en IRAMUTEQ permitieron discriminar claramente los contenidos semánticos referidos a las afirmaciones de los cuidadores sobre el cuidado a realizar con el recién nacido en el domicilio. **Conclusión**: esta herramienta computarizada resultó ser útil, adecuada y confiable, calificando el proceso de análisis textual.

Descriptores: Investigación; Investigación en enfermería; Investigación cualitativa; Metodología; Programas informáticos

Introduction

In recent decades, scientific and technological development has shown important advances in Brazil.[1] In this scenario, health professionals, including nurses, are getting to know and incorporating new research methods in their scientific investigations. However, they experience challenges to ensure that the proposed studies are relevant and sufficiently rigorous methodologically[2] and become published and, consequently, visible in journals with important impact factors.

In the health area, nursing research has the role of producing and improving knowledge, seeking the qualification of care, and improvements in the quality of life of people.[3] Therefore, the production of knowledge in the area needs to rely on robust research methods, that guarantee the systematization and the organicity of the data for adequate analysis and interpretation of the findings, by the researcher.

Among the various types of research, a qualitative investigation is a methodological approach often used among health and nursing researchers, as it presents an idealistic, subjectivist and interpretive view of reality that is multiple and subjective, enabling a more comprehensive perspective on human relationships. In this approach, the experiences of individuals and their perceptions are useful and important aspects for research,[4] which leads to the understanding of complex phenomena in the analytical process.

The researcher who uses the qualitative method seeks to apprehend the subjectivity of individuals from their context.[5] However, the raw data from this approach are composed of languages, as they encompass thoughts, beliefs, and/or opinions. Therefore, the analysis stage becomes an arduous task, given the subjectivity of the data and the large textual volume, not very structured, added to the pressing search for greater methodological rigor.[6-7]

A viable alternative is the analysis software, already consolidated in quantitative research and which, since the 1980s, has been growing among qualitative researchers, with the name Computer-Aided Qualitative Data Analysis Software (CAQDAS). Some of the advantages in the analytical process for this software are the organization and separation of information and the increase in the efficiency of the process and the ease of localization of the text segments.[8]

Among the CAQDAS, the *Interface de R pour les Analyses Multidimensionnelles de Textes et de Questionnaires* (IRAMUTEQ) stands out. It is a free textual data processing program, created by the French researcher Pierre Ratinaud, in his language, currently encompassing several dictionaries, including Portuguese, which made its use in Brazil possible since 2013.[7-9]

IRAMUTEQ aims to apprehend the structure and organization of utterances, exposing the most recurrent relationships between lexical universes, admitting that the words used in similar contexts correlate to the same lexical world.[7] It enables the analysis of interviews, documents, essays, dynamics, among others, from very simple to multivariate analyzes, offering a wide range of possibilities for textual analysis,[10] sometimes little explored by health researchers.

Scope review that analyzed 121 qualitative health research studies found that 21.5% of the studies used software, including Nvivo (18.2%), Atlas.ti (1.7%), Dedoose (0.8%), and CQPweb software (0.8%).5 Therefore, despite the free use of IRAMUTEQ, it is inferred that researchers of qualitative health research are not yet familiar with this tool or are unaware of it.

It appears that the production of evidence with the application of this software in the health area is incipient, which requires an expansion of studies for the recognition of the tool and is a

reference for its applicability, contributing to the methodological training of researchers. Thus, the objective was to report the experience of using IRAMUTEQ as a computerized tool to support the analysis of textual data in qualitative research in health.

Method

This is an experience report on the use of the IRAMUTEQ software in qualitative research, developed in a public hospital in the city of Rio das Ostras in the State of Rio de Janeiro, Brazil. Ten puerperal women and nine relatives (five aunts and four parents) of newborns aged 18 years old and over participated in the study.

As a data production technique, we used one of the dynamics of creativity and sensitivity of the Creative-Sensitive Method (CSM), entitled Concrete Dynamics.[11] Its use implied in the clarification of the caregiver through the execution of care practice, in this study, referring to postnatal care of newborns to be developed by families at home, such as body hygiene and nutrition.

The CSM is based on dialogue, listening, and solidarity, favoring the active participation of participants in the search for the collective construction of knowledge, through artistic productions based on a Debate Generating Question (DGQ). This method combines typical strategies of traditional qualitative research such as participant observation, press conference, and group discussion.[12]

Eight dynamics were carried out and totaled each one in five moments as provided by the method.[12] Each meeting took place in a private environment of the institution, between March and June 2019, with an average duration of 20 minutes. The participants' role consisted of demonstrating and talking about the care they intended to carry out with the baby at home, separating the materials to be used or not, producing a fictitious care scene, guided by the following DGQ: How do you intend to carry out the care for your newborn at home?

Different appropriate and inappropriate materials were made available for the care of the

newborn. These materials included pacifier, baby bottle, umbilical band, coin, gauze, cotton, 70% alcohol, antiseptic medicine containing povidone-iodine (povidine), cotton swabs, diapers, ointment against rash, wet handkerchief, shampoo, soap, perfume, talc, scissors, among others, in addition to a static vinyl mannequin with dimensions similar to a baby's body.

At the collective level, the productions were socialized through speech and demonstration, spontaneously, in the group. In this process, convergent and divergent issues generated themes of debate, collective analysis, and group discussion. Finally, there was a synthesis of the topics covered and the validation of the information by the participants. This type of dynamics favored the free expression of ideas, promoting the generation of a sufficient lexical volume for the processing of textual data in the selected software.

For a complete and accurate record of the speeches, all the Concrete Dynamics about the postnatal care of newborns were recorded in digital media, which after transcribing, generated a single textual corpus, which constituted the primary source of the data, submitted to lexicographic analysis, using the IRAMUTEQ software. We used the five types of analysis available: Classic Textual Statistics; Word Cloud; Similitude Analysis; Descending Hierarchical Classification (DHC) and Correspondence Factor Analysis (CFA).

For the correct use of the software, some precautions are necessary. Initially, to complete its installation, the R software was downloaded and installed free of charge on the computer, which has a language and environment for statistical analysis and graphics production and later the IRAMUTEQ, also open access. These steps must occur in the order described since IRAMUTEQ uses software R in the processing of its analyzes.

Other procedures are also necessary for the software to work properly. It is common for IRAMUTEQ to be available in another language, usually in French, so the language change was made by clicking on 'edition', then 'preferences' and 'interface language', with a change to Portuguese. In the end, it was necessary to close the software so that it could save the modifications.

Subsequently, the program was restarted and error checking was performed, again clicking on 'editing', then 'preferences', and then 'check the installation of R packages' as instructed.

This study comprises a subproject entitled "Knowledge and practices of family members about home care with low-risk newborns", approved by the Research Ethics Committee, under number 3,100,491, on December 21, 2018. The research was conducted following Resolutions 466/2012, 510/2016, and 580/2018 of the Ministry of Health. The data were produced after recruitment and invitation made personally by the researchers, consent process, and obtaining the signature of the Informed Consent Form by the participants in a reserved environment at the institution. For anonymity, we used an alphanumeric code related to the 19 study participants [M - mothers (1-10), F - fathers (1-4), and A - Aunts (1-5)], according to the order of participation.

Results

After the phases of installation and necessary adjustments for the good functioning of the program, already described, there was an analysis of the textual data that occurred in three stages: 1) preparation and codification of the textual corpus with the description of the empirical material derived from the dynamics developed; 2) processing of textual data in the software and 3) interpretation of the findings by the researchers (which were not detailed because it was not the object of analysis in this experience report).

For the first stage, we transcribed in full all the speeches of the participants, transcribed from the eight dynamics of creativity and sensitivity, with 10 puerperal women and nine family members of newborns, with an average participation of two to three per dynamics. For the composition of the textual corpus, the speeches were grouped dynamically, because of the CSM which presupposes a group process, unlike individual interviews, as the group becomes a space of multiple voices for the research participants, in the enunciation of concrete existential situations, liable to socialization in the collective space. Therefore, the eight texts analyzed in IRAMUTEQ

came from the transcriptions of the eight dynamics, which together gave rise to the textual corpus of analysis, being sorted in only one text editing file and separated by a command line, for each dynamic, which necessarily must include four asterisks, a space, an asterisk, and an individual or collective identification for the text (**** * Dynamic_1 to **** * Dynamic_8).

The researchers reviewed the text, who excluded from the text the interventions and all the verbal material they produced, remaining only the speeches of the participants and who answered the question in focus. Typing and punctuation errors were corrected, as well as the removal of words, loose phrases and language vices, such as "É ...", "Ihhh", "Aí", "Né", "Tá", "Tô", "Ah", "Oh", "Hummm", in addition to quotation marks ("), apostrophe ('), hyphen (-) and percentage (%).

Members have been kept in their algebraic form and compound words have been unified using the underline character "_", replacing spaces (examples: newborn, wet_kerchief, alcohol_70, wet_band, soap_coconut, tooth_brush, scissors_with_point, and scissors_on_point) so that the program counted them as a single term.

After preparing the textual corpus, an attentive rereading was carried out, seeking to achieve the greatest possible use, in data processing, of the lexical universe that made up the corpus of analysis, since IRAMUTEQ does not have a function for text verification and correction. Later, this single file was stored as a plain text document (.txt) with UTF-8 character encoding (Unicode Transformation Format 8-bit code units).

For the second step, the first step was to 'open a textual corpus' in IRAMUTEQ and, later, in the settings window, which opens automatically, in 'define characters' the option "utf_8_sig - all languages" was selected, as well as 'Portuguese' in the language and 'standard' in the dictionary. The other settings are kept.

The first type of analysis carried out is called "Classic Textual Statistics", which allows an initial analysis of the corpus, presenting the relationship between the frequency and the number of lexical units. Thus, without stemming (process to be selected or not for each type of analysis), that

is, without reducing the words based on their roots, the textual corpus generated was composed of eight texts, with a total of 8,124 occurrences, being 1,160 (100%) distinct words, of which 594 (51.2%) had a single occurrence (hapax), with an average occurrence per text of 1,015.5 words.

In this analysis, the Zipf Diagram (Figure 1) is generated, a visual way to demonstrate the behavior of words in the corpus, by illustrating, on the vertical axis, the frequency of occurrence of words throughout the text and, on the horizontal axis, the number order/ranking of words, with the value 1 for the most recurring, 2 for the next, and so on. That is, on the Y-axis the logarithm of the frequencies is shown, or how many times a word is present in the textual corpus, and on the X-axis the number of words is shown. Thus, we observed a decreasing curve in the diagram, in which the points located at the upper limit of the graph, closer to the Y-Axis, represent words with a high frequency of repetition, but less recurrent in the analyzed material, as indicated by the X-Axis.
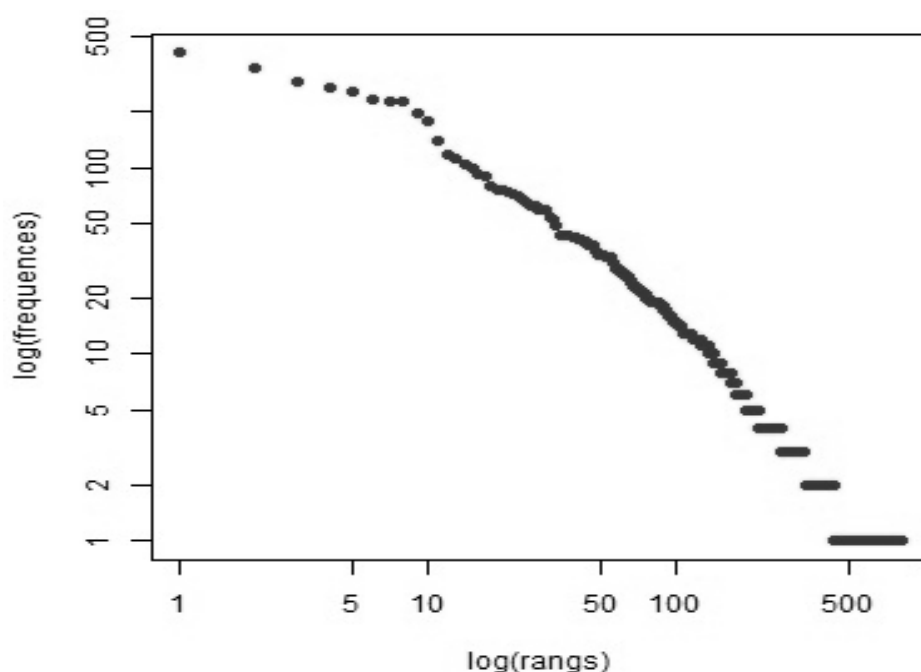

Figure 1 - Zipf diagram with the behavior of the words in the textual corpus about the caregivers' claims regarding home care for the newborn. Rio das Ostras, RJ, Brazil, 2019.

Therefore, for the left side of the curve, it appears that few words are repeated many times and on the right side many are repeated few times. Therefore, given this situation, we reinforce those words with frequency 1 (hapax) were recurrent in the text (f = 594), as is visible at the end of

Axis X (horizontal line). On the other hand, only three active forms, after stemming, appeared more than one hundred times in the transcribed dynamics, making up the group of the most recurrent ones, at the top of the Y-Axis (vertical line). In order of ranking, these active forms were: no (f = 410), use (f = 224) and because (f = 105).

When interpreting the findings, we found that the use of these terms is justified by the dynamics used, in which the participants explained how they would carry out home care with the baby, separating the materials they would or not use and explaining the reason for the choice:

> *I would not use umbilical_band. People used to put it on, but today I don't use it, because I think the navel slaps and to heal, to dry, I think it's going to get wet.* (P1)
> *I would use the coin and the umbilical_band.* (M1)

The analysis through the "Word Cloud" projects words structured in the form of a cloud, with different sizes. The largest were those that hold certain importance because they are repeated more often in the textual corpus. Thus, based on the frequency indicator, we could identify the words of greater preponderance - the largest and more centrally than the others - without using a cut-off point for the inclusion of terms in the cloud (Figure 2). It appears that the terms "no", "use" and "because" are highlighted in the set of words, corroborating with the previous analysis.
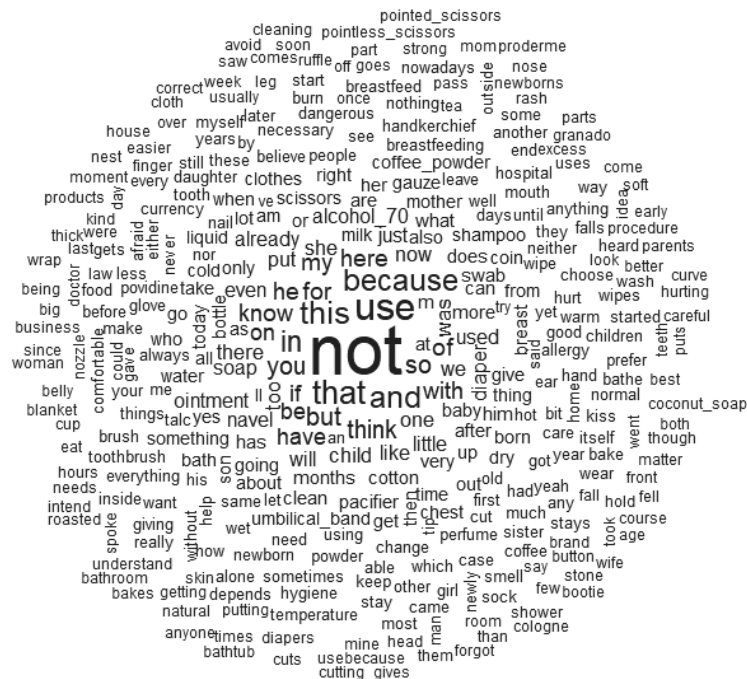
Figure 2 - Word cloud about caregivers' intentions regarding home care for the newborn. Rio das Ostras, RJ, Brazil, 2019.

We also verified that the main materials pronounced by the mothers and their families, with a frequency of up to ten repetitions, to be used or not in the care of newborns, were: diapers (f=43), alcohol 70% (f=38) , ointment (f = 33), pacifier (f = 28), baby bottle (f = 27), coin (f = 26), cotton swab (f = 26), soap (f = 23), gauze (f = 21), shampoo (f = 21), clothing (f = 20), talc (f = 19), scissors (f = 19), umbilical band (f = 18), wet wipe (f = 16), perfume (f = 13) , povidine (11) and brush (f = 10). It is noteworthy that it is possible to redeem all text segments corresponding to the terms in the software.

With the Similitude Analysis, based on graph theory, in addition to the occurrences of the words, we identified the indications of the connections (links) between them, which helps in the identification of the lexical content structure (Figure 3). For this analysis, words with recurrence of at least 10 times were included. We observed that the three words that stood out the most (no, use, and why), as already mentioned, assumed central positions in three groups of terms that generated different ramifications, with emphasis again on the term "no" as it is the most central and with the largest number of connections.

Figure 3 - Similitude analysis with the connection and indications of connectedness between the words related to the caregivers' claims regarding home care for the newborn. Rio das Ostras, RJ, Brazil, 2019.

Therefore, in the Similitude Analysis, progress is made in understanding the ramifications of the central words for other words that connect, which allowed an expansion in the interpretation of the data. Thus, the path by which the words were enunciated and their interconnections, in which the strongest and most recurrent, are represented by thicker lines, were identified.

Therefore, as the participants' statements presented references that are part of the process of caring for newborns at home, there was a thorough elucidation of 'what to use', 'how to use' and 'what not to use' in the care of their babies, including the reasons for the choices indicated, as shown in the following text segments:

> *Diaper, ointment too, I use it even if I don't have anything, I pass it to prevent it, with every diaper change. Not a bottle yet, because I intend to leave it only in the breast for up to 6 months.* (P10)
> *I would use a pacifier, I will stop by the pharmacy to buy it because here you cannot.* (P9)
> *I would use these scissors without a tip, this one has a tip, it can hurt* (T1).

Another method used for analysis was the DHC, in which the text segments are grouped

according to their respective vocabularies and the set of them is divided according to the presence or absence of the stemmed forms, constituting different lexical worlds. This method establishes a stable and definitive classification, through classes of text segments that, at the same time, a present vocabulary similar to each other, according to common lexical parameters and vocabulary different from the text segments of the other classes, through calculations of distances and proximity using chi-square tests (χ^2).

Thus, the software organizes the analysis of the data in different layouts of dendrograms in the DCH, which illustrate the relationships between the classes. For this analysis, we found 243 text segments, with a classification of 203 of them, that is, 83.54% of utilization. The stemming resulted in 803 lemmas and, among the active forms, 705 were analyzable and 262 had a frequency ≥ of three. In the DCH, six classes of different text segments were generated (Figure 4), which demonstrates the semantic content of each class.
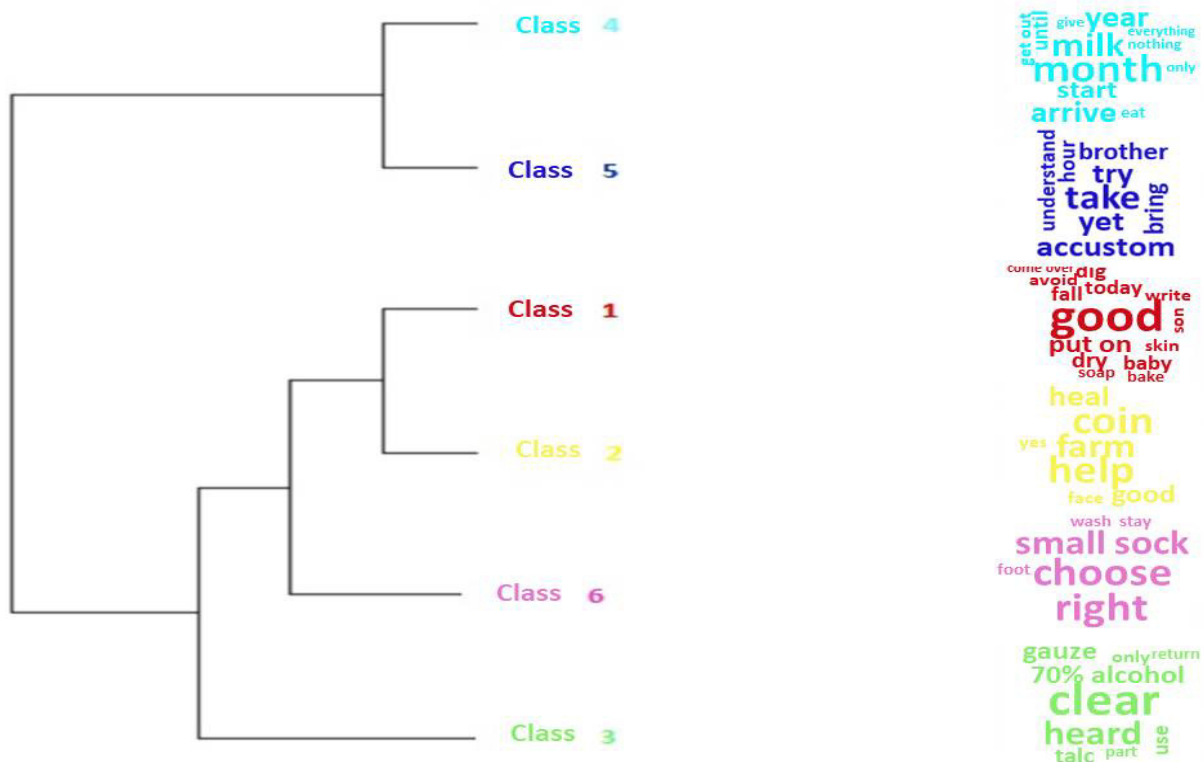

Figure 4 - Dendrogram in the Descending Hierarchical Classification of the text segments related to the caregivers' claims regarding home care for the newborn. Rio das Ostras, RJ, Brazil, 2019.

The analysis was stable and, consequently, the six classes were also stable, that is, they were composed of text segments with similar words. In the dendrogram, the textual corpus was partitioned into two independent blocks (subcorpus). The first was composed of class 3 (21.2%), a second subdivision including class 6 (17.2%), and another subdivision with classes 2 (12.3%) and 1 (13.8%), which have closer semantic content but, still, with some differentiation. The other block was formed by classes 4 (15.8%) and 5 (19.7%), being the most isolated from the others, which demonstrates greater proximity and homogeneity between the two.

In the interpretation of the data, from the DHC, the active forms of each class of text segments were retrieved, which included nouns, adjectives, adverbs, verbs, and unrecognized forms, with emphasis on those obtained in the chi-square test (χ^2) a value ≥ 3.84. Therefore, $p < 0.05$ reveals the strength of the word's association in a respective class. The lower the χ^2, the less the terms are related to the class; on the other hand, the lower the p-value, the greater this relationship, which contributes to the validation and reliability of the findings. Thus, words whose p-values were less than 0.0001 are extremely significant in each class, when they denote with more than 99.99% certainty that their allocations in the class were not by chance.

Thus, after processing the data by the software, the meaningful words in the classes and their insertions in the text segments were read. Therefore, based on the semantic content, the classes were named: Class 1: Care for the newborn's bath; Class 2: Umbilical stump care; Class 3: Materials used in newborn hygiene; Class 4: Nutrition of the newborn; Class 5: Use of artificial nipples; Class 6: Decision-making power in the care of the newborn.

We reinforce that this nomination, made by the authors, considered the semantic universe of each class. In other words, from the main associated words, their exhaustive reading, and the semantic context in which they were inserted, the meaning of the classes was extracted, leading to their appointment. The interpretation of the set of these divisions allowed us to reach an understanding of how the participants intended to carry out home care with the newborn.

The program also offers another way of presenting the results, through the CFA that represents, on a Cartesian plane, the different words and the positioning of the classes, based on the frequencies and the correlation values (χ^2) of each word in the corpus (Figure 5), favoring the visualization of the typical vocabulary of each class in different lexical worlds or semantic contexts.
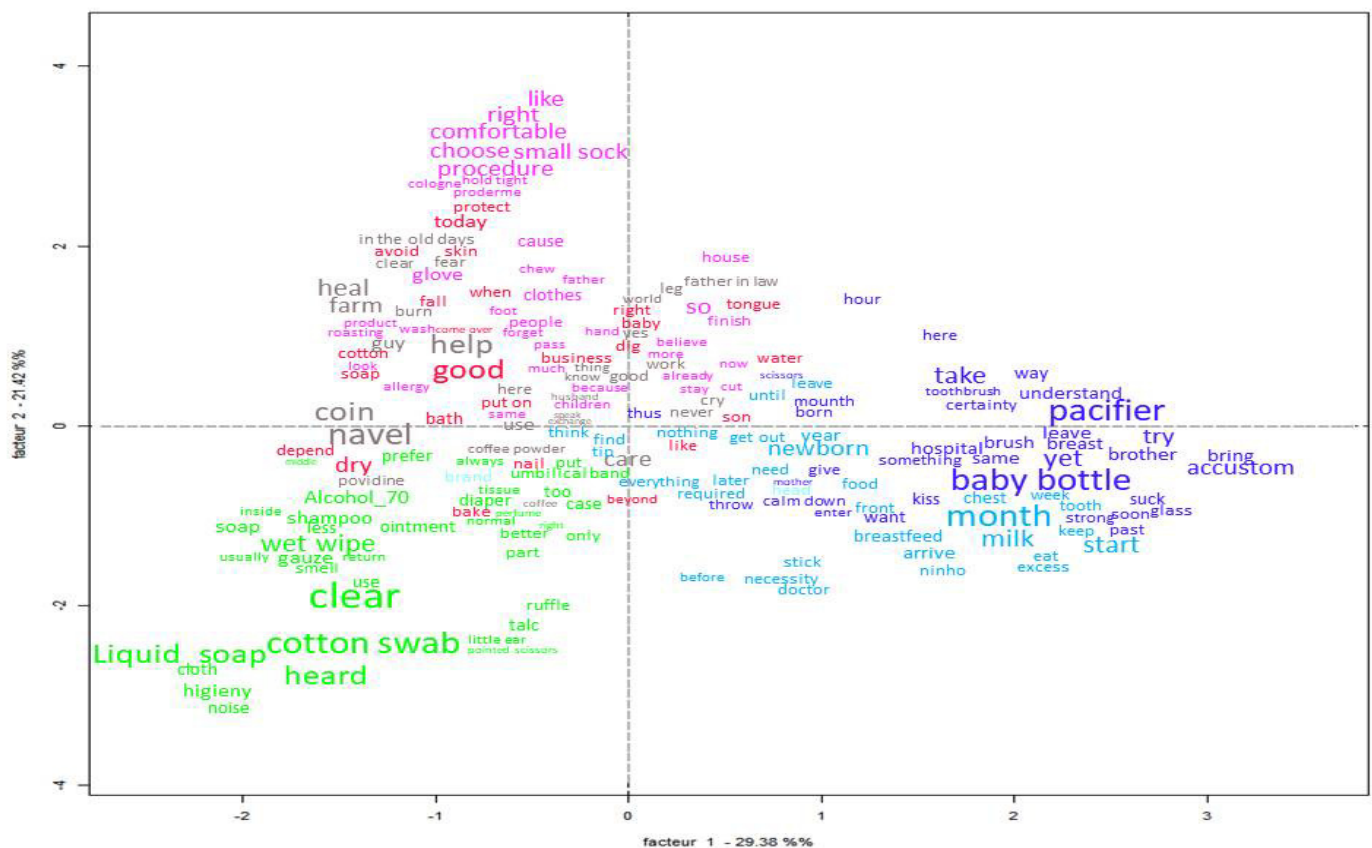

Figure 5 - Correspondence Factor Analysis (CFA) of the text segments related to caregivers' intentions regarding home care for the newborn. Rio das Ostras, RJ, Brazil, 2019.

From the CFA, in general, we found the most important words with larger χ^2, therefore, larger sizes in the image, all classes are presented in a more centralized segment, while the others expand to the periphery. Also, few words from each class are dispersed to the other quadrants. It is noteworthy that class 3 (lower left quadrant), in green, is the most isolated in the Cartesian plane, without advancing to any other quadrant; therefore, less associated with other classes and with a very particular lexical world. At DCH, this was the first class constituted, covering a list of specific materials related to the hygiene of the newborn.

We could confirm that classes 1, 2, and 6, presented here in red, gray, and pink (upper left

quadrant), appear associated with each other. This occurs as the bathing care of the newborn and the umbilical stump are interconnected in the speeches of the puerperal women and their families, as well as the questions related to the decision-making power of the participants. Therefore, the words of the three classes are mixed in the quadrant.

In the interpretative analysis, we found that, mainly, the family members who had previous knowledge in newborn care, opted for the way they believed to be correct in carrying out a care practice, demonstrating autonomy in their choices.

Finally, classes 4 and 5, in light blue and dark blue (lower right quadrant), are correlated in the factorial plane, as they have vocabularies closer to the newborn's nutrition. They include breastfeeding and the use or not of artificial teats, such as pacifiers and bottles, justifying their proximity.

The use of IRAMUTEQ in this qualitative research was of great value for the researchers involved, as it allowed the improvement of their knowledge in the handling of analytical software. Furthermore, it is reinforced that its applicability favors the operationalization of the data, mainly, regarding the organization and separation of a large textual volume quickly and singularly, when the corpus is properly and in advance prepared. Therefore, it is even easier to interpret the findings objectively, as the text segments are easily recovered. Another perceived advantage was the possibility for all researchers to have access to the software, as it was free and with a Portuguese dictionary, which enabled meetings for training and discussions in the stage of processing textual data, with the effective participation of all those involved in the research, who learned how to use this tool.

Discussion

The use of IRAMUTEQ was designed for large textual volumes. As long as the group is homogeneous, in general, between 20 and 30 texts are recommended.[7] In this study, although 19

family members participated, the textual corpus was classified as usable and suitable for carrying out this type of analysis, as it reached a percentage higher than the recommended by the literature of 75%.[7-8]

The level of use of 83.54% can be justified by the use of CSM, which, by using art as a tool for free expression of existential situations, created a space for discussion and reflection mediated by dialogue. The individuals problematized their experiential and existential practices on the subject on screen in a free and less mechanized way, [12] generating enough data for processing in IRAMUTEQ.

The number of participants was limited during the fieldwork, due to data saturation, identified during the organization of the empirical material derived from the dynamics, through the informational redundancy of the participants' statements.[13] The collection was considered saturated when no new elements were found, with the recurrence of ideas on the theme in the dynamics and without the addition of new information, [14] which was confirmed by the level of use.

The different lexical analysis techniques allowed to discriminate the semantic contents referring to the pretensions of the puerperal women and their relatives about the care to be developed with the newborn at home, including different materials and practices to be adopted, in addition to the reasons for their choices. Therefore, we found that the use of this software ensured the necessary statistical rigor in the face of the need to deal with extensive textual responses; the improvement of the analyzes, especially, by the integration of the quantitative and qualitative domains in the analytical process and of greater objectivity in the interpretation of data in researches of qualitative approach.[15] Also, we found that its interface is simple and easy to understand, as indicates by the literature.[10]

According to the Zipf diagram, several words were not very recurrent, and very rare words were repeated a lot, and the three most mentioned were relevant throughout the analytical process, remaining in central and prominent positions in the Word Cloud and the Similitude Analysis. The

first, even though it was a simpler lexical analysis, was useful and interesting, since, by enabling the rapid identification of the keywords in a corpus, [7,10] it even favored the identification of the list of materials to be used in the care of the baby.

The Similitude Analysis, based on a mathematical model to study the relationships between objects, allowed to identify the co-occurrences between the words, which indicated the links between them, helping to identify the structure of the textual corpus.[16] This analysis of similarities contributed to better visualization of how the participants related the various words to describe the process of caring for newborns at home, including what to use or not and the reasons for the options.

DCH has been the most used form of analysis in research using IRAMUTEQ.[7] A six-class dendrogram was generated in it, containing the different words that obtained the highest association index in each of the classes, by applying a statistical test widely known in studies with a quantitative approach. This is the Pearson's chi-square, that the larger, the more likely is the hypothesis of dependence between the active word and the class. This analysis indicated a high power of distinction of the software in the representativeness of the participants' sets of ideas.[15]

This resource enabled, based on the original corpus, the recovery of the text segments and the association of the words between them, which allowed the grouping of statistically significant words, with subsequent interpretative analysis by the researchers, favoring a more precise and assertive analysis of the data, [8,17] by distinguishing classes that are, at the same time, sufficiently distinct from the others and sufficiently similar to each other, in terms of vocabulary.

With the CFA, we could ascertain a clear difference in the distribution of words along the crossing of the X and Y axes in the factorial plane; however, in line with the distribution and interconnection between the classes provided by DCH. Such an interface favors the detailed evaluation of the links between the profiles of the individual responses, expressing convergences and oppositions and showing, graphically, the links existing between the different lexical worlds.[18]

In this directive, the semantic contents related to the baby's hygiene and nutrition were in distinctly different plans, which is an important differential in the use of IRAMUTEQ.

Because of the above, the importance of knowledge of the different text processing alternatives offered by the software is reinforced to use it correctly. Therefore, the relevance of the researcher in all phases of the research is emphasized, with due theoretical support for the preparation of the textual corpus.[19] The data analysis, based on IRAMUTEQ, cannot lose sight of the problem of research, which needs to guide the processes of data collection, treatment, analysis, and interpretation, as well as inferential processes must be supported by the theoretical references of the study and critical reflections on the validity and meaning of the findings.[20]

Some limitations in the application of the software need to be highlighted. As it is an aid tool in the analysis, the interpretation of what may be behind the lexical similarities and differences is a task that goes beyond the limits of the software. Therefore, it depends on the conceptual references and/or theoretical constructs adopted by the researchers. Also, coding errors in the command lines, in Portuguese and colloquial words, are not processed by the software, and it is up to the researcher to identify such errors and correct them, without changing the meaning of the lines. Therefore, it requires attention in transforming the transcribed content into a corpus that will be analyzed and, thus, obtains a good textual use.

Conclusion

The study reported the experience of using the IRAMUTEQ software as a computerized tool to support the analysis of textual data resulting from qualitative health research. The use of the software was useful, adequate, precise, and trustworthy, qualifying the analysis process and, therefore, the research results, enabling its interpretation with due scientific rigor. Therefore, the use of this tool enhanced the analysis of data from the dynamics of creativity and sensitivity.

The scarcity of research that discloses the use of this software, in Brazil, with the

description of the different types of analysis was a challenge found, which on the other hand, reinforced the importance of this study that described the analytical process considering all available interfaces. Therefore, the need for further studies that apply the IRAMUTEQ is highlighted, to expand the understanding and use of this tool in research with a qualitative approach in the area of health and nursing.

References

1. Souza DL, Zambalde AL, Mesquita DL, Souza TA, Silva NLC. A perspectiva dos pesquisadores sobre os desafios da pesquisa no Brasil. Educ Pesqui. 2020;46:e221628. doi: 10.1590/s1678-4634202046221628

2. Thorne S. Making our nursing research matter. Rev Bras Enferm. 2016;69(5):813-4. doi:10.1590/0034-7167-2015-0156

3. Erdmann AL, Peiter CC, Lanzoni GMM. Brazilian research groups in nursing: comparison of 2006 and 2016 profiles. Rev Gaúcha Enferm. 2017;38(2):e69051. doi: 10.1590/1983-1447.2017.02.69051

4. Patias ND, Hohendorff JV. Critérios de qualidade para artigos de pesquisa qualitativa. Psicol Estud. 2019; 24:e43536. doi: 10.4025/psicolestud.v24i0.43536

5. Salvador PTCO, Alves KYA, Rodrigues CCFL, Oliveira LV. Online data collection strategies used in qualitative research of the health field: a scoping review. Rev Gaúcha Enferm. 2020;41:e20190297. doi: 10.1590/1983-1447.2020.20190297

6. Nunes JV, Woloszyn M, Gonçalves BS, Pinto MDS. A pesquisa qualitativa apoiada por softwares de análise de dados: uma investigação a partir de exemplos. Rev Front. 2017;19(2):233-44. doi: 10.4013/fem.2017.192.08

7. Salvador PTCO, Gomes ATL, Rodrigues CCFM, Chiavone FBT, Alves KYA, Bezerril MS, et al. Uso do software IRAMUTEQ nas pesquisas brasileiras da área da saúde: uma scoping review. Rev Bras Promoç Saúde. 2018;31:1-9. doi: 10.5020/18061230.2018.8645

8. Souza MAR, Wall ML, Thuler ACMC, Lowen IMV, Peres AM. The use of IRAMUTEQ software for data analysis in qualitative research. Rev Esc Enferm USP. 2018;52:e03353. doi: 10.1590/S1980-220X2017015003353

9. Kami MTM, Larocca LM, Chaves MMN, Lowen IMV, Souza VMP, Goto DYN. Working in the street clinic: use of IRAMUTEQ software on the support of qualitative research. Esc Anna Nery. 2016;20(3):e20160069. doi: 10.5935/1414-8145.20160069

10. Camargo B, Justo AM. IRAMUTEQ: um software gratuito para análise de dados textuais. Temas Psicol. 2013; 21(2):513-18. doi: 10.9788/TP2013.2-16

11. Gomes AMT, Cabral IE. Ocultamento e silenciamento familiares no cuidado à criança em terapia antiretroviral. Rev Bras Enferm. 2010;63(5):719-26. doi: 10.1590/S0034-71672010000500005

12. Soratto J, Pires DEP, Cabral IE, Lazzari DD, Witt RR, Sipriano CAS. A maneira criativa e sensível de pesquisar. Rev Bras Enferm. 2014;67(6):994-9. doi: 10.1590/0034-7167.2014670619

13. Saunders B, Sim J, Kingstone T, Baker S, Waterfield J, Bartlam B, et al. Saturation in qualitative research: exploring its conceptualization and operationalization. Qual Quant. 2018;52(4):1893-907. doi: 10.1007/s11135-017-0574-8

14. Nascimento LCN, Souza TV, Oliveira ICS, Moraes JRMM, Aguiar RCB, Silva LF. Theoretical saturation in qualitative research: an experience report in interview with schoolchildren. Rev Bras Enferm. 2018;71(1):228-33. doi: 10.1590/0034-7167-2016-0616

15. Carvalho TS, Mota DM, Saab F. Utilização do software IRaMuTeQ na análise de contribuições da sociedade em processo regulatório conduzido pela Agência Nacional de Vigilância Sanitária. Vigil Sanit Debate. 2020;8(1):10-21. doi: 10.22239/2317-269x.01429

16. Mendes FRP, Zangão MOB, Gemito MLGP, Serra ICC. Social Representations of nursing students about hospital assistance and primary health care. Rev Bras Enferm. 2016;69(2):321-8. doi: 10.1590/0034-7167.2016690218i

17. Silva AMB, Enumo SRF. Description and analysis of psychological intervention with dancers by the Software IRAMUTEQ. Temas Psicol. 2017;25(2):577-93. doi: 10.9788/TP2017.2-11Pt

18. Trigueiro DRSG, Almeida AS, Monroe AA, Costa GO, Bezerra VP, Nogueira JA. AIDS and jail: social representations of women in freedom deprivation situations. Rev Esc Enferm USP. 2016; 50(4):554-61. doi: 10.1590/S0080-623420160000500003

19. Acauan LV, Abrantes CV, Stipp MAC, Trotte LAC, Paes GO, Queiroz ABA. Utilização do software Iramuteq® para análise de dados qualitativos na Enfermagem: um ensaio reflexivo. REME Rev Min Enferm. 2020;24:e1326. doi: 10.5935/1415-2762.20200063

20. Sousa YSO, Gondim SMG, Carias IA, Batista JS, Machado KCM. O uso do software Iramuteq na análise de dados de entrevistas. Pesqui Prát Psicossociais [Internet]. 2020 [acesso em 2020 nov 29];15(2):1-19. Disponível em: http://pepsic.bvsalud.org/scielo.php?script=sci_arttext&pid=S1809-89082020000200015&lng=pt&tlng=pt

Corresponding author
Fernanda Garcia Bezerra Góes
E-mail: ferbezerra@gmail.com
Address: Rua Recife, Lotes, 1-7 - Jardim Bela Vista, Rio das Ostras/RJ. ZIP-Code: 28895-532

Author´s contributions

1 – Fernanda Garcia Bezerra Góes
Contributions: She collaborated with the conception and design of the study, in the analysis and interpretation of data, and the final review with critical and intellectual participation in the manuscript.

2 – Andressa Silva Torres dos Santos
She collaborated with the conception and design of the study, in the analysis and interpretation of data, and the final review with critical and intellectual participation in the manuscript.

3 – Brenda Lucas Campos
She collaborated with the conception and design of the study, in the analysis and interpretation of data, and the final review with critical and intellectual participation in the manuscript.

4 – Aline Cerqueira Santos Santana da Silva
She collaborated in the analysis and interpretation of data and the final review with critical and intellectual participation in the manuscript.

5 – Liliane Faria da Silva
She collaborated in the analysis and interpretation of data and the final review with critical and intellectual participation in the manuscript.

6 – Luiz Carlos Moraes França
She collaborated in the analysis and interpretation of data and the final review with critical and intellectual participation in the manuscript.

How to cite this article