Review

A glance at the use of IRaMuTeQ® software in scientific research: a bibliometric study

Um olhar sobre o uso do software IRaMuTeQ® na pesquisa científica: estudo bibliométrico
Una mirada al uso del software IRaMuTeQ® en la investigación científica: estudio bibliométrico

Dayara de Nazaré Rosa de Carvalho¹
ORCID: 0000-0001-8569-3392
Viviane Ferraz Ferreira de Aguiar²
ORCID: 0000-0003-3025-1065
Daiane Baldo Apolinário³
ORCID: 0000-0002-3492-6114
Dandara de Fátima Ribeiro Bendelaque¹
ORCID: 0000-0002-5580-284X
Luana Cunha Galvão Pereira¹
ORCID: 0000-0002-9639-5945
Simone Aguiar da Silva Figueira¹
ORCID: 0000-0001-6368-6124
Fabiana de Souza Orlandi³
ORCID: 0000-0002-5714-6890

Abstract
Objective: To analyze and map the national and international scientific production on IRAMuTeQ® use software in scientific research in health in the historical series from 2012 to August 2022. **Methods**: This is a descriptive bibliometric and scientometric research and exploratory. Collection was carried out in August 2022 through institutional online access to the CAPES Journals Portal, in the Scopus database (Elsevier). To guide the research, the keyword "IRaMuTeQ" was used. After the selection process of included studies, data set was saved in a single file in CSV Excel format and later exported to the VOSviewer® software. **Results:** It was observed a greater number of articles on the topic in 2021 (127; 34.70%). In Brazil, there is the largest number of authors who publish using IRaMuTeQ® (271; 74.04%). Social sciences and nursing are the two professional areas that most use the software in their articles. Of the 15 authors who most publish with IRaMuTeQ®, 12 are Brazilian. **Conclusion:** IRaMuTeQ® use assessment in scientific research allows us to point out the expressive role of Brazil in using this tool in scientific research.

Descriptors: Software; Qualitative Research; Investigative Techniques; Data Analysis.

¹Universidade do Estado do Pará. Belém, Pará, Brasil.
²Universidade Federal do Pará. Belém, Pará, Brasil.
³Universidade Federal de São Carlos. São Carlos, Santa Catarina, Brasil.

Corresponding author:
Dandara de Fátima Ribeiro Bendelaque
E-mail: bendelaqued@gmail.com

Whats is already known on this?
IRAMuTeQ® is based on the statistical bases of the R software to support the performance of different forms of statistical analysis on text corpus and tables of individuals by words.

What this study adds?
It allowed emphasizing the expressive use of the tool by Brazilian authors in scientific research as well as in social sciences and nursing categories.

Resumo

Objetivo: *Analisar e mapear a produção científica nacional e internacional sobre o uso do software IRaMuTeQ® na pesquisa científica em saúde na série histórica de 2012 a agosto de 2022.* ***Métodos:*** *Trata-se de uma pesquisa bibliométrica e cientométrica, de caráter descritivo e exploratório. A coleta foi efetuada no mês de agosto de 2022 por meio do acesso institucional online ao Portal de Periódicos CAPES, na base de dados Scopus (Elsevier). Para guiar a pesquisa, foi utilizada a palavra-chave "IRaMuTeQ". Após o processo de seleção dos estudos incluídos, o conjunto de dados foi salvo em um único arquivo no formato CSV Excel e, posteriormente, exportado para o software VOSviewer®.* ***Resultados:*** *Foi observado maior número de publicações acerca da temática no ano de 2021 (127; 34,70%). No Brasil, encontra-se o maior número de autores que publicam com o uso do IRaMuTeQ® (271; 74,04%). As ciências sociais e a enfermagem são as duas áreas profissionais que mais utilizam o software em suas publicações. Dos 15 autores que mais publicam com o IRaMuTeQ®, 12 são brasileiros.* ***Conclusão:*** *A avaliação do uso do IRaMuTeQ® nas pesquisas científicas nos permite pontuar o expressivo protagonismo do Brasil na utilização dessa ferramenta nas pesquisas científicas.*

Descritores: *Software; Pesquisa Qualitativa; Técnicas de Pesquisa; Análise de Dados.*

Resumén

Objetivo: *Analizar y mapear la producción científica nacional e internacional sobre el uso del software IRaMuTeQ® en investigaciones científicas en salud en la serie histórica de 2012 a agosto de 2022.* ***Métodos:*** *Se trata de una investigación bibliométrica y cienciométrica descriptiva y exploratoria. La recolección se realizó en agosto de 2022 a través del acceso institucional en línea al Portal de Revistas Periódicas CAPES, en la base de datos Scopus (Elsevier). Para guiar la investigación se utilizó la palabra clave "IRaMuTeQ". Luego del proceso de selección de los estudios incluidos, el conjunto de datos se guardó en un solo archivo en formato CSV Excel y posteriormente se exportó al software VOSviewer®.* ***Resultados:*** *Se observó un mayor número de publicaciones sobre el tema en el año 2021 (127; 34,70%). En Brasil, existe el mayor número de autores que publican utilizando IRaMuTeQ® (271; 74,04%). Ciencias sociales y enfermería son las dos áreas profesionales que más utilizan el software en sus publicaciones. De los 15 autores que más publican con IRaMuTeQ®, 12 son brasileños.* ***Conclusión:*** *La evaluación del uso del IRaMuTeQ® en la investigación científica permite señalar el papel expresivo de Brasil en el uso de esta herramienta en la investigación científica.*

Descriptores*: Software; Investigación Cualitativa; Técnicas de Investigación; Análisis de Datos.*

INTRODUCTION

The increasing improvement of information technology represents an important ally for the development of research in the field of health, as it enables significant advances and temporal progress in its progress. Debates about computer program use for data management in research began in the 1980s.[1] Among the possible data to be managed are those of a textual nature, and these, in turn, are extensive.[2]

In the 1980s, there was an increase in Computer Aided Qualitative Data Analysis Software (CAQDAS) use, commonly used in qualitative studies.[3] CAQDAS make it possible to manage data with an audiovisual and textual nature, thus allowing a faster process of organizing and coding materials.[4]

Among the available software, those most frequently used in health research depend on the acquisition of a license to authorize its use; of these, ALCESTE, which was developed in France, represented an innovation in the face of demands for textual analysis at the end of the 1980s, due to the fact that it allows lexicographical analysis and together with a large contingent of statistical methods. As for its operation, it performs the Descending Hierarchical Classification (DHC): a mechanism that makes it possible to carry out the analysis and establishment of lexical classes, a fact that allows the formation of vocabularies and text segments, thus making the interpretation process more reliable to the as a process of data exploration and description is carried out. In short, the software allows a glance under the totality of the material under analysis, but it does not allow the understanding about the subject of the content in question.[4]

Among the software available with "open source" technology and free access, the R *Interface pour les Analyzes Multidimensionnelles de Textes de Questionnaires* (IRaMuTeQ®) stands out, based on statistical bases of the R software and the "phyton" programming language, developed by Pierre Ratinaud.[2] The R software allows performing different forms of statistical analysis on text *corpus* and tables of individuals by words.[5]

It has an innovative character, as it makes it possible to measure the frequency in which words appear in the text, in addition to different types of statistical analysis of textual data and matrices, such as classic textual statistics, correspondence factorial analysis, DHC, similarity analysis and word cloud.[1,2]

Thus, such characteristics allow IRAMuTeQ® to be used in the face of the analysis demands of the most varied types of texts. As for the program's analytical character, it is worth noting that it is guided by the lemmatization process, in which words are collected and classified by their root.[1]

IRAMuTeQ® promotes statistical rigor, enabling the analysis and categorization of large volumes of textual data. Allows objectivity as it facilitates material organization and text *corpus* statistical analysis,

with the aim of comparing and relating specific variables present in the text, expanding its vision to possible levels of categorization and decision making.[6]

In Brazil, research using IRAMuTeQ® began in 2013.[3] The highest prevalence in psychology studies in the field of social representations is highlighted, followed by studies carried out by nursing. These initial movements highlight the software's flexibility in the face of the different possibilities of the most diverse data processing, a fact that subsidizes text organization, and this provides more credibility to the constructed material.[4]

Finally, it should be noted that, with regard to ALCESTE, the software has advantages over specific demands for text analysis, and one of the barriers to its use is the fact that it is a paid software. IRAMuTeQ®, on the other hand, also allows performing DHC analysis together with other analyses, which result in graphic materials, and there is also the free factor, which makes it more frequently chosen for carrying out text *corpus* analysis. However, the understanding and mastery of IRAMuTeQ® is built gradually, and this results in more investment time to study the tool and, many times, this keeps researchers away from the software.[2]

Between 2020 and 2021, there is virtually a greater number of materials and courses available on IRAMuTeQ®, and such evidence may be a consequence of the COVID-19 pandemic, resulting social isolation and naturalization of the classroom in the virtual environment.[4]

Therefore, this study aims to analyze and map the national and international scientific production on IRAMuTeQ® use in scientific research in the historical series from 2012 to August 2022.

METHODS

This is a bibliometric or scientometric research, with a descriptive and exploratory character. Bibliometric research consists of a quantitative and statistical method that maps the scientific production and dissemination of a given subject, making it possible to analyze the productivity of certain topics.[7] It is also called scientometrics, as it includes tables and graphs to show descriptive statistics, using advanced text and network analysis to reveal hidden patterns in abstract content regarding relationships between terms.[8]

Bibliometric studies are based on basic laws, known as bibliometric laws, the three main ones being: Bradford's Law (characterized by journals with the highest publication on a given topic); Lotka's Law (which consists of the authors who most publish on a certain area of knowledge); and Zipf's Law (which addresses the relationship and frequency between words in the text, associating with the topic's main guiding approaches).[9]

This study had as a research question: how is the scientific production on IRAMuTeQ® use between 2012 and August 2022?

The collection was carried out in August 2022, through institutional online access to the Coordination for the Improvement of Higher Education Personnel (CAPES - *Coordenação de Aperfeiçoamento de Pessoal de Nível Superior*) Journals Portal, in the Scopus database (Elsevier). To guide the research, the keyword "IRaMuTeQ" was used. Scopus is considered the largest database of citations and abstracts of academic literature, covering scientific journals, books and conference papers.[10] The history series was chosen to expand the number of articles on the topic.

For data analysis, the following variables were used: 1) citation information (author; document title; year; citation count; document type; publication stage; and DOI); 2) bibliographic information (affiliations; language of original document); 3) abstracts; author keywords and indexed keywords; and 4) references of selected studies.

After the selection stage with the keyword, the following inclusion criteria were then established: full text available in the database; original and review articles (integrative, bibliometric and systematic); available in full in Portuguese, English and Spanish; articles published between 2012 and 2022 that addressed the issue in question in a manner relevant to the proposed object of the study.

As exclusion criteria, the following were used: conference documents; letters to the editor; theses; dissertations; or even works that did not deal specifically with the topic in question. After defining the pre-established criteria, the search began in the said database, which resulted in 387 articles. After this moment, the title and abstract were previously read. However, after applying the filters, the final sampling resulted in 366 articles, as shown in Figure 1 below:

Figure 1. PRISMA flowchart adapted for the article selection process. Belém, Pará, Brazil, 2022.
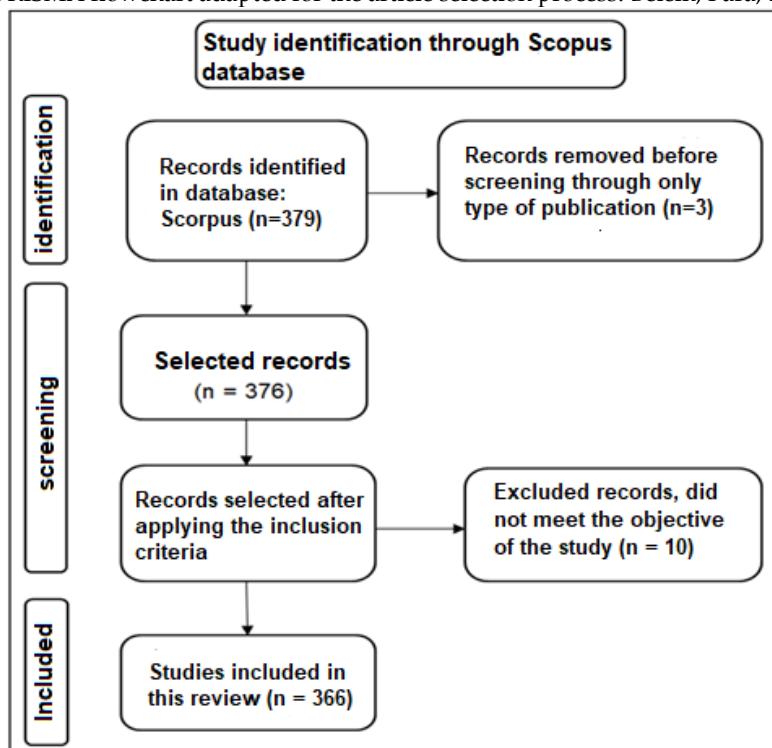
Source: research authors (2022).

After carrying out the entire selection process of included studies, the data set was saved in the cloud (online drive), in a single file in CSV Excel format and later exported to the VOSviewer® software (version 1.6.6). It is worth noting that VOSviewer® is a tool that creates co-citation networks from the analysis of documents and articles in a database, such as Web of Science, PubMed and Scopus [10].

This bibliometric mapping with a quantitative approach allowed the visualization of several bibliometric aspects of scientific publications, carried out in the form of different networks, whose emphasis is on the analysis and visualization of large sets of bibliographic data from a "distance based" approach. Therefore, the bibliographic network nodes are placed in such a way that the distance between them approximately indicates their relationship according to certain aggregation criteria, forming a map. [11]

The *corpora* set selected from articles and exported to VOSviewer® underwent analysis regarding co-authorship by country, number of citations by institution and overlay display, citations by institution based on distribution by geothermal density, visualization by network of co-occurrence of keywords and visualization by keyword density. Furthermore, Microsoft Excel was also used to carry out the descriptive statistics of results. Because it is a study that uses secondary data and is freely available in a data repository, there was no need for submission to the Research Ethics Committee (REC).

RESULTS

In the studies selected for this bibliometric review, it was observed that the largest number of articles on the topic was in 2021 (127; 34.70%), followed by 2020 (76; 20.77%) and 2019 (45;12.30%), as can be seen in Figure 2 below:

Figure 2. Distribution of selected articles by year in the Scopus Elsevier database. Belém, Pará, Brazil, 2022.
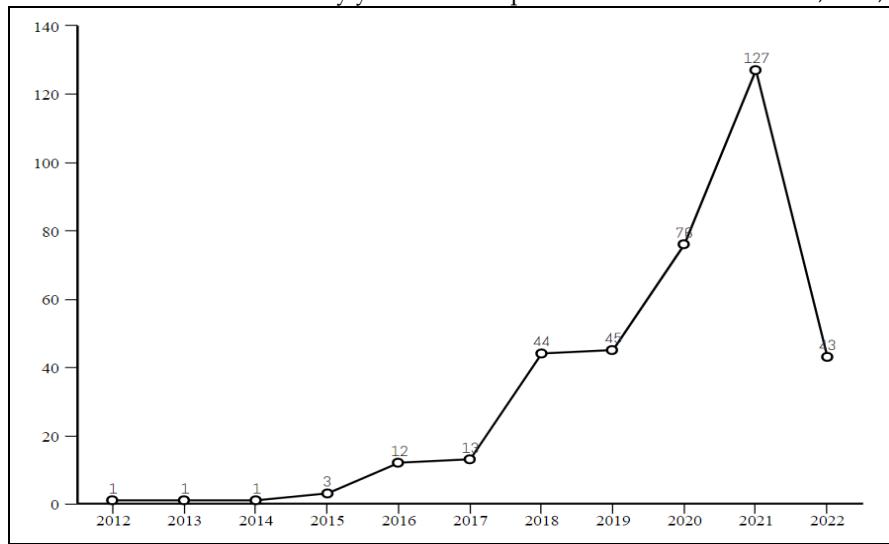
Source: research authors (2022).

As for the number of articles per country, it is observed that Brazil has the highest number of authors who publish massively using IRaMuTeQ® (271; 74.04%), ranking 1st with almost ¾ of all articles in Scopus; France (32; 8.74%) ranks 2nd; Portugal and Spain (30; 8.30%) rank 3rd; and Italy (11; 3.01%) ranks 4th. Thus, the five countries in the world that most publish with software data analysis are formed. For better observation, the chart below shows the 15 countries in the world that publish the most.

Figure 3. Distribution of the number of articles by country. Belém, Pará, Brazil, 2022
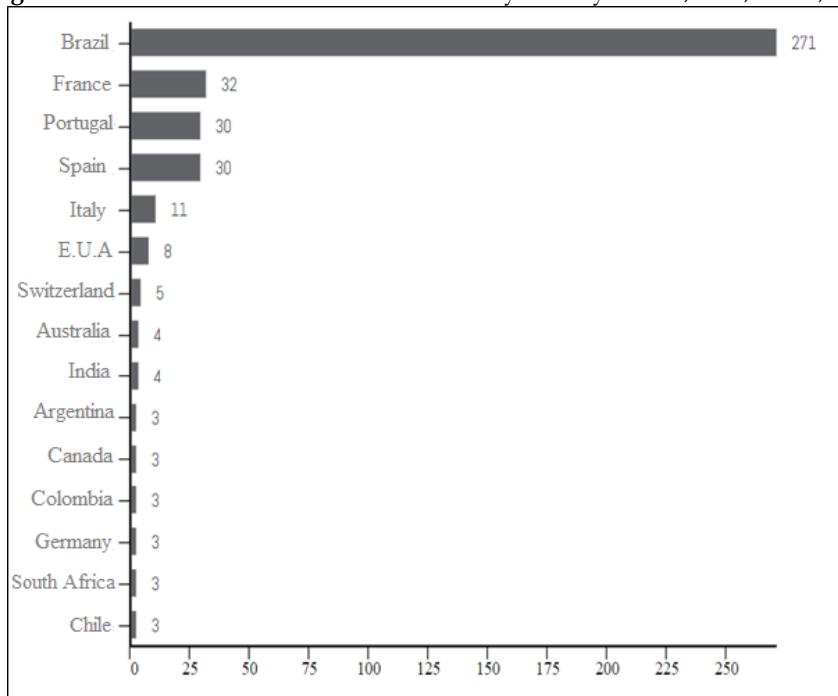
Source: research authors (2022).

As for the authors' publication area, the social sciences occupy the 1st position (121; 33.06%), followed by nursing (105; 28.69%). The sum of the two areas represents 226 (61.75%) articles out of a total of 366, thus showing the massive domain of both in IRAMuTeQ® use. Figure 4 shows the number of articles using the software by professional area.

Figure 4. Distribution of articles by professional area. Belém, Pará, Brazil, 2022.
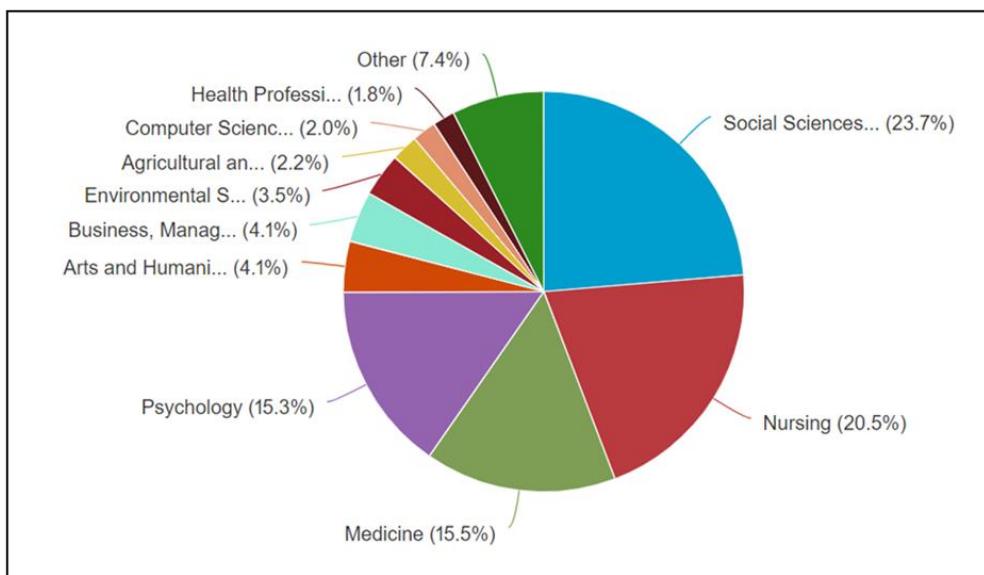
Source: research authors (2022).

With regard to Lotka's Law, it was possible to notice that, of the 15 authors who published the most on the topic, 12 are from Brazil, followed by authors from Spain, France and Portugal, each with one article, respectively. Regarding H-index, the author Peres AM has the highest number of articles (7;1.96%), with an H7 index; Idoiaga Mondragon N (6; 1.64%) has an H11 index; Ratinaud P (6; 1.64%) has an H5 index; Araújo LF (6;1.64%) has an H6 index; and Abreu PD (5; 1.37%) has an H3 index. It is noteworthy that the author Santos AF has fewer articles (4; 1.09%), however his index is high (H46), as seen in Table 1 below:

Table 1. Distribution of authors with the highest number of articles on the topic, frequency, H-index, institution of affiliation and country (Lotka's Law). Belém, Pará, Brazil, 2022.

nº	Authors	f*	H-index	Institution	Country
1	Peres, AM	7	h-7	*Universidade Federal do Paraná*	Brazil
2	Idoiaga Mondragon, N	6	h-11	*Universidad del País Vasco*	Spain
3	Ratinaud, P	6	h-5	*Université Fédérale Toulouse Midi-Pyrénées*	France
4	de Araújo, LF	6	h-6	*Universidade Federal do Piauí*	Brazil
5	Abreu, PD	5	h-3	*Universidade de São Paulo*	Brazil
6	Araújo, ECd	5	h-3	*Universidade Federal de Pernambuco*	Brazil
7	Lowen, IMV	5	h-3	*Universidade Federal do Paraná*	Brazil
8	Melo, CdF	5	h-2	*Universidade de Fortaleza*	Brazil
9	Salerno, MB	5	h-1	*Universidade Federal de Mato Grosso do Sul*	Brazil
10	Suto, CSS	5	h-2	*Universidade do Estado da Bahia*	Brazil
11	Almeida, CAPL	4	h-3	*Universidade Católica Portuguesa*	Portugal
12	Carvalho, EC	4	h-2	*Universidade do Estado do Rio de Janeiro*	Brazil
13	Santos, AF	4	h-46	*Universidade Federal de Sergipe*	Brazil
14	Fernandes, MA	4	h-4	*Universidade Federal do Piauí*	Brazil
15	Gomes, AMT	4	h-10	*Universidade do Estado do Rio de Janeiro*	Brazil

f*= frequency
Source: research authors, 2022.

With regard to the most cited studies, 4 (1.09%) articles obtained a value between 45 and 53 citations. The article with the highest number of citations is from 2020 (n=53 citations), followed by an article from 2018 (n=50 citations) and an article from 2019 (n=45 citations). The most cited article among the 366 articles selected for this study is found in Frontiers in Psychology, with impact factor (IF) of 4.232,

considered high, according to the literature. Table 2 shows the distribution of the most cited studies in terms of year, title, authors, journal, impact factor and number of citations.

Table 2. Distribution of the most cited studies in terms of year, title, authors, journal, impact factor and nº of citations. Belém, Pará, Brazil, 2022.

nº	Year	Title	Authors	Journal	nº of citations	Impact factor (IF)
1	2020	Exploring Children's Social and Emotional Representations of the COVID-19 Pandemic	Idoiaga N; Berasategi N; Eiguren A; Picaza M	Frontiers in Psychology	53	4.232
2	2018	The use of IRAMUTEQ software for data analysis in qualitative research	Souza MAR; Wall ML; Thuler ACMC; Lowen IMV; Peres AM	*Revista da Escola de Enfermagem da USP*	50	1.123
3	2019	Education for advancing the implementation of the Sustainable Development Goals: A systematic approach	Avelar ABA; Silva-Oliveira KDD; Pereira RDS	International Journal of Management Education	46	2.707
4	2021	Struggling to breathe: a qualitative study of children's wellbeing during lockdown in Spain	Idoiaga N; Berasategi N; Eiguren A; Picaza M	Psychology and Health	45	3.898
5	2014	The second generation of ecological economics: How far has the apple fallen from the tree?	Plumecocq G	Ecological Economics	32	5.389
6	2016	Public Understanding of Science in turbulent times III: Deficit to dialogue, champions to critics	Smallman M	Public Understanding of Science	23	4.254
7	2012	Improbable search of a homogeneous diversity: The debate on national identity	Ratinaud P; Marchand P	Langages	16	1.583
8	2018	The generativity and attractiveness of social representations theory from multiple paradigmatic approaches in various thematic domains: An empirical meta-theoretical analysis on big-data sources from the specialised repository "SoReCom 'A.S. de Rosa' @-library"	De Rosa AS; Bocci E; Dryjanska L	Papers on Social Representations	14	1.31
9	2019	Mapping pharmacy journals: A lexicographic analysis	Mendes AM; Tonin FS; Buzzi MF; Pontarolo R; Fernandez-Llimos F;	Research in Social and Administrative Pharmacy	13	2.84
10	2019	Extracurricular activities in higher education and the promotion of reflective learning for sustainability	Díaz-Iso A; Eizaguirre A; García-Olalla A	Sustainability (Switzerland)	13	3.889
11	2018	Knowledge about HIV/AIDS and implications of establishing partnerships among Hornet® users	Queiroz AAFLN; Sousa AFL; Matos MCB; Araújo TME; Reis RK; Moura MEB	*Revista Brasileira de Enfermagem*	13	0.705
12	2018	Good practices for patient safety in the operating room: nurses' recommendations	Gutierres LS; Santos JLG; Peiter CC; Menegon FHA; Sebold LF; Erdmann AL	*Revista Brasileira de Enfermagem*	13	0.705

13	2020	The rise of biohacking: Tracing the emergence and evolution of DIY biology through online discussions	Meyer M; Vergnaud F	Technological Forecasting and Social Change	11	1. 115
14	2020	Public understanding of science and common sense: Social representations of the human microbiome among the expert and non-expert public	Galli I; Fasanelli R	Health Psychology Open	11	2.40
15	2018	Conditioning factors for exclusive breastfeeding from the mother's perspective	Rocha GP; Oliveira MdCF; Ávila LBB; Longo GZ; Cotta RMM; Araújo RMA	*Cadernos de Saúde Publica*	11	1.277

Source: research authors, 2022.

According to Zipf's Law, the co-occurrence and relationships between the main keywords are observed. For this study, words that appeared at least seven times in the *corpus* were selected as a cutoff point, totaling n=379 types of word forms found in the articles. Data were operationalized by VOSviewer®. From then on, the 5 clusters were formed, shown below:

- Green: the main words were "Iramuteq Software" (n=210), "Interview"(n=97), "Social Representation" (n=78), "Representation" (n=54), "Woman" (n=49), "Life" (n=45);
- Red: the main words were "Process" (n=110), "Research" (n=107), "Development" (n=59), "Use" (n=53), "Perspective" (n=50);
- Yellow: the main words were "Education" (n=114), "Student" (n=100), "Lexical analysis" (n=82), "Discourse" (n=71)" and "Construction" (n=9);
- Blue: the main words were "Class" (n=122), "Practice" (n=111), "Care" (n=91), "Service" (n=78) "Nursing" (n=69) and "Brazil" (n=54);
- Lilac: the main words were "World" (n=88), "Crisis" (n=67), "Anger" (n=52) and "Emotion" (n=38).

Respecting Zipf's Law, Figure 5 describes the co-occurrence network of keywords, showing the words and their clusters. Based on geothermal analysis, the most frequent word zones are identified according to Kernel density.

Figure 5. Keywords co-occurrence network based on the title and abstract of selected studies, data operationalized in the VOSviewer® software (Zipf's Law). Belém, Pará, Brazil, 2022.
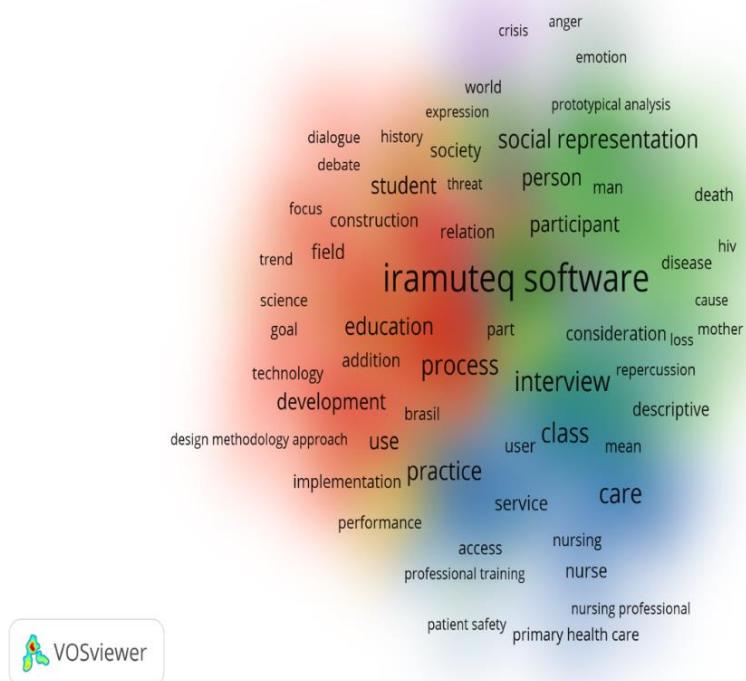
Source: research authors, 2022.

Based on the analysis of selected studies, the articles addressed 3 major thematic groups: 1) IRaMuTeQ® use in studies of social representations; 2) IRaMuTeQ® in technology development and construction studies; and 3) IRAMuTeQ® use in the field of nursing.

DISCUSSION

In this study, it was observed that the scientific production with IRAMuTeQ® use has been growing all over the world, with a peak in the number of articles in 2021, however, when taking into account the time frame of 10 years, the quantity is still small. This fact corroborates with studies that claim that there are numerous inconsistencies of publications with IRAMuTeQ® use.[3]

When it comes to the countries with the highest number of authors who use IRaMuTeQ® in research, Brazil stands out, with France, where the software was created, ranking 2nd. Software use by Brazilian researchers is expressive as a result of professional master's programs in the country, where researchers have chosen to use software to support qualitative research, in response to current needs in the search for greater methodological rigor. Therefore, there is still a predominance of software use in dissertations resulting from academic master's degrees, with a significant growth of research carried out within the scope of professional master's degrees and postgraduate courses.[12]

Also noteworthy is IRAMuTeQ® use in institutions in northeastern Brazil, with significant amounts of studies developed in the state of Piauí, denoting appropriation of researchers from this geographic region in software use.[12]

Qualitative studies that use IRaMuTeQ® tools, due to the statistical rigor in data processing and analysis, combined with instrument quality in relation to data collection and creation of a robust and systematic text *corpus*, allow researchers, through integration of statistical methods with qualitative analysis, graphical representations, presenting results with better processing and in order to help in discussion, supported by the methodological rigor of software use and the researcher's creativity.[5]

With regard to the publication area of the authors who publish the most using IRaMuTEQ®, social sciences and nursing prevail with a significant number of articles. The findings corroborate other studies that show IRAMuTeQ® use in data analysis from interviews in research in different areas, predominantly in the areas of psychology and nursing.[13]

It is noteworthy that the scientific revolution of the seventeenth century culminated in mathematical sciences use as a fundamental instrument for analysis. This fact guided research in social sciences to adopt scientific and academic legitimacy through the positive sciences. Software use is an alternative to this positivist, empirical and mathematizing view of quantitative research methods. Thus, numerous qualitative social researchers add software to their research.[14]

The expressiveness of Brazilian nursing in the development of qualitative research and analysis with IRaMuTeQ® may occur due to the need for qualitative studies to respond to the concerns of nurses with regard to nursing care and assistance, due to the adequacy to the objects of study of research in this area, who also use software in their investigations.[10]

With regard to Lotka's Law, it was observed that Brazilian authors were the leading actors in terms of publications on the topic, with a significant number of articles, with a single author with frequency 7 and index H7, being considered of great relevance in the area. H-index was proposed in 2005 by the physicist Jorge E. Hirsch, with the aim of assessing the individual scientific production of a researcher. Aiming to estimate the importance and impact of their accumulated contribution in research, this metric has three relevant characteristics, namely: it presents a value that never decreases over the course of the researcher's trajectory; its increase does not occur linearly; and its value depends on the nature of the researcher's area.[15]

Scholars have been unanimous in considering H-index a measure that tends to value the researcher's scientific effort, when considering the scientist's entire academic career cumulatively. This statement highlights the result obtained in the research in which the author Peres AM has the highest number of articles, with the H7 index, and Idoiaga MN, with an H11 index, however, Santos AF, despite having fewer articles, its index is higher than the of the others (H-46). It is noteworthy that Pierry Ratinaud, creator of the method, presented an H5 index.[16]

The survey results also indicate that the most cited article among those chosen for the study is found in the journal Frontiers in Psychology, which has IF of 4.232. It is considered the largest cited multidisciplinary psychology journal in the world, publishing rigorously peer-reviewed research in the psychological sciences. It is an open access journal that is at the forefront of disseminating and communicating scientific knowledge and discoveries, with over 8,000 articles published and around 37,000 citations in Scopus.[17]

IF, created in 1955 by Eugene Garfield to assess the impact of scientific journals and journals, should be understood as the average number of citations received per work in a given period.[18]

It is of great importance to highlight that the second article with the highest number of citations was published in the Brazilian journal *Revista da Escola de Enfermagem da USP*, which presents a continuous flow of publications, being peer-reviewed, and which aims to publish unpublished empirical articles that represent a breakthrough significant for professional practice or for the fundamentals of nursing. The journal presents IF of 1.123 and counts that renowned indexing sources, such as Web of Science, Scopus, CINAHL, among others.[19]

Both journals use peer review of scientific research. Peer review at the start of publication of scientific journals was done to enhance the editorial experience, rather than exercising more conventional modes of quality control. However, with the scientific development, it is currently considered one of the most accepted procedures worldwide by the scientific community in determining the quality of scientific works, more commonly conceived as an assessment of scientific articles carried out by researchers or specialists.[20]

With regard to the number of citations of the articles, we observed that only 1 article obtained n ≥ 53 citations among the 366 articles selected for this study. The number of citations of an article, a metric widely used in academia, plays an important role in science when there is an interest in seeing the consolidated influence of a particular scientist or institution on the progress of scientific knowledge over time.[21]

With regard to Zipf's Law, we observed that, for this study, a Kernel density map was constructed using VOSviewer®, thus making it possible to identify the most frequent words that determine the main topic addressed in the selected documents. Kernel density consists of quantifying the relationships of points within a radius (R) of influence based on a given statistical function, analyzing the patterns traced by a given set of specific data, estimating their density in the study area.[22]

Kernel density analysis generates a density surface for visual identification called "hotspots". Thus, it becomes possible to identify the occurrence of a hot area as a concentration of events that somehow indicates the agglomeration of a phenomenon in a spatial distribution.[23]

With regard to the thematic groups, it is observed that 3 major formed axes address IRAMuTeQ® use in scientific research, being a widely used tool in studies of social representations, development and construction of technologies in health and in the area of nursing. In the field of social representations, the choice of IRaMuTeQ® is justified by the fact that it offers a large number of tools for qualitative data analysis based on textual statistics, or lexicometry, in addition to the fact that it is a free software, using statistical calculations on the vocabulary of analyzed *corpus*, thus promoting greater statistical rigor to the set of analyzed texts.[23,24]

Moreover, the type of analysis most used by researchers is DHC, which enables the rescue of text segments, together with the association of each one, allowing statistically significant words to be grouped and qualitative data analysis to be performed. With regard to IRAMuTeQ® use in studies of development and construction of health technologies, the software is an excellent aid in the process of creating technologies, since the software performs textual analysis through anchoring to the statistical program R, generating data from texts (text *corpus*) and tables.[5]

In this way, the products of these analyzes show the position and structure of words in a text, their connections and other textual characteristics, which allow the detection of indicators and, thus, intuitively visualize the structure and environments of the text to be analyzed. Authors agree, pointing out that the software is relevant, as it allows using chi-square (χ^2) as a statistical test to determine the correlation and association between words, allowing the establishment of word class distribution through descriptive statistics use.[25]

In the field of nursing, the importance of producing and improving knowledge about research by nurses is highlighted, thus contributing to the quality of care and users. Therefore, the development of research in nursing needs to rely on robust methods that guarantee data systematization and quality for adequate analysis and interpretation.[26]

IRaMuTeQ® offers a range of options for using isolated or combined techniques, meeting the wishes of all qualitative researchers who intend to overcome the qualitative and quantitative dichotomy, recognizing the limits and advantages of complementarity in text data analysis. Great benefits are observed, such as statistical rigor, analysis of a large volume of data, objectivity, different possibilities of analysis, simple interface and gratuity, expanding the researcher's vision to possible levels of categorization and decision making.[27,28]

As limitations of this research, we highlight the data collection carried out solely by the Scopus database. Furthermore, doctoral theses and master's dissertations were not part of study analysis. Furthermore, as evidenced in the results, research using IRaMuTeQ® is still incipient, and this fact reflects the limitations of studies on this topic. Therefore, it is recommended to develop further studies that demonstrate the importance of IRaMuTeQ® in qualitative research.

CONCLUSION

Analysis and mapping of national and international scientific production on IRAMuTeQ® use in scientific health research show that most articles are current, that is, the software has had a greater proportion, mainly from 2019 to 2021. Brazil has been highlighted in the number of articles, with approximately 3/4 of all publications in Scopus, and the professional areas social sciences and nursing are the main areas that use.

It is noticed that the scientific production of a quantitative nature that uses IRaMuTeQ® is still not very expressive; however, it is possible to notice the significant use of the tool in academic and professional master's dissertations, mainly in the field of health. This fact is due to the advantages of the tools that IRaMuTeQ® has as the researcher inserts data and builds a text *corpus*. Analysis of program results in graphic representations and classification of variables that, together, help researchers developing the discussion of their topic, and also occurs the methodological strengthening of the material in question, allowing better reliability of inferential and interpretative processes.

CONTRIBUITIONS

Contributed to the conception or design of the study/research: Carvalho DNR. Contributed to data collection: Carvalho DNR, Aguiar VFF, Apolinário DB, Bendelaque DFR, Pereira LCG, Figueira SAS,

Orlandi FS. Contributed to the analysis and/or interpretation of data: Carvalho DNR, Aguiar VFF, Apolinário DB, Bendelaque DFR, Pereira LCG, Figueira SAS, Orlandi FS. Contributed to article writing or critical review: Carvalho DNR, Aguiar VFF, Apolinário DB, Bendelaque DFR, Pereira LCG, Figueira SAS, Orlandi FS. Final approval of the version to be published: Carvalho DNR, Aguiar VFF, Apolinário DB, Bendelaque DFR, Pereira LCG, Figueira SAS, Orlandi FS.

REFERENCES

1. 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. [Internet]. 2020;24: e1326. Available from: http://dx.doi.org/10.5935/1415-2762.20200063.

2. Camargo BV, Justo AM. IRAMUTEQ: um software gratuito para análise de dados textuais. Temas psicol. [Internet] 2013;21(2): 513-8. Available from: http://dx.doi.org/10.9788/TP2013.2-16.

3. Soares SSS, Costa CCP da, Carvalho EC, Queiroz ABA, Peres PLP, Souza NVD de O. Teaching Iramuteq for use in qualitative research according to YouTube videos: an exploratory-descriptive study. Rev esc enferm USP [Internet]. 2022;56:e20210396. Available from: https://doi.org/10.1590/1980-220X-REEUSP-2021-0396.

4. Canuto A, Braga B, Monteiro L, Melo R. Aspectos críticos do uso de caqdas em pesquisas qualitativas: uma comparação empírica de alcest e iramuteq como ferramentas digitais. NTQR [Internet] 2020; 3:199-211. Available from: https://doi.org/10.36367/ntqr.3.2020.199-211.

5. Souza MAR de, Wall ML, Thuler AC de MC, Lowen IMV, Peres AM. O uso do software IRAMUTEQ na análise de dados em pesquisas qualitativas. Rev esc enferm USP [Internet]. 2018;52:e03353. Available from: https://doi.org/10.1590/S1980-220X2017015003353.

6. Tinti D da S, Barbosa GC, Lopes CE. O software IRAMUTEQ e a Análise de Narrativas (Auto)biográficas no Campo da Educação Matemática. Bolema [Internet] 2021;35(69):479–96. Available from: https://doi.org/10.1590/1980-4415v35n69a22.

7. Amaral MCG do, Lima MC de, Rocha LS da, Carvalho D de NR de, Peixoto O da S, Orlandi F de S, Peixoto IVP. A saúde ocupacional de enfermeiros docentes: um estudo bibliométrico. Revista Recien [Internet] 2022;12(38):14-22. Available from: https://doi.org/10.24276/rrecien2022.12.38.14-22.

8. Damar HT, Bilik O, Ozdagoglu G, Ozdagoglu A, Damar M. Scientometric overview of nursing research on pain management. Rev Latino-Am Enfermagem [Internet]. 2018;26:e3051. Available from: https://doi.org/10.1590/1518-8345.2581.3051.

9. Carvalho DNR, Neves LNA, Sobrinho CRO, de Aguiar VFF, Ferreira LGS, Orlandi AAS, Pavarini SCL, Peixoto IVP, Orlandi FS. Avaliação da capacidade funcional da pessoa idosa: estudo bibliométrico. Casos e consultoria [Internet] 2022;13(1):e28857. Available from: https://periodicos.ufrn.br/casoseconsultoria/article/view/28857.

10. Santos A, Mangini ER, Urdan AT, Brasil Rossini FH. Avaliação bibliométrica em inovação em serviços. Perspectivas em Gestão & Conhecimento [Internet] 2017;7(1):212-31. Available from: http://dx.doi.org/10.21714/2236-417X2017v7n1p212.

11. Palludeto AWA, Felipini AR. Panorama da literatura sobre a financeirização (1992-2017): uma abordagem bibliométrica. Econ soc [Internet]. 2019;28(2):313–37. Available from: https://doi.org/10.1590/1982-3533.2019v28n2art02.

12. Salvador PTC de O, Gomes AT de L, Rodrigues CCFM, Chiavone FBT, Alves KYA, Bezerril M dos S, Santos VEP. Uso do software iramuteq nas pesquisas brasileiras da área da saúde: uma scoping review. Rev Bras Promoc Saúde [Internet] 2018;31. Available from: https://doi.org/10.5020/18061230.2018.8645.

13. Sousa YSO, Gondim SMG, Carias IA, Batista JS, Machado KCM. O uso do software Iramuteq na análise de dados de entrevistas. Pesquisas e Práticas Psicossociais 2020;15(2):1-19. Available from: http://pepsic.bvsalud.org/scielo.php?script=sci_arttext&pid=S1809-89082020000200015&lng=pt&nrm=iso.

14. Baccin Brizolla MM, Petry JF, Uchôa AGF, Ferreira HLB. Uma revisão sobre a pesquisa qualitativa em ciências sociais aplicadas. UFAMBR [Internet] 2020;2(3):103-30. Available from: https://doi.org/10.47357/ufambr.v2i3.8087.

15. Oliveira EFT, Gracio MCC. Indicadores bibliométricos em ciência da informação: análise dos pesquisadores mais produtivos no tema estudos métricos na base Scopus. Perspect ciênc inf [Internet]. 2011;16(4):16–28. Available from: https://doi.org/10.1590/S1413-99362011000400003.

16. Grupo Scimago. El índice h de Hirsch: aportaciones a un debate. El profesional de la información, 2007; 16(1): 47-9.

17. Góes FGB, Santos AST, Campos BL, Silva ACSS, Silva LF, França LCM. Utilização do software IRAMUTEQ em pesquisa de abordagem qualitativa: relato de experiência. Rev Enferm UFSM [Internet] 2021;11:e63. Available from: https://doi.org/10.5902/2179769264425.

18. Klant LM, Santos VS. The use of the IRAMUTEQ software in content analysis - a comparative study between the ProfEPT course completion works and the program references. RSD [Internet] 2021;10(4):e8210413786. Available from: https://doi.org/10.33448/rsd-v10i4.13786.

19. Monte WS, De-BortoliR, Feitosa ERM. Patentes de processos, gamification e proteção ao usuário a partir da análise do software Iramuteq. Rev. Tecnol. Soc. [Internet] 2021; 17(47): 172-82. Available from: http://dx.doi.org/10.3895/rts.v17n47.11924.

20. Slemian A, Couto KC, Vieira MV. O Processo de avaliação por pares e a Ciência Aberta: um nó górdio para os periódicos na área de História?. Rev Bras Hist [Internet]. 2022;42(90):7–15. Available from: https://doi.org/10.1590/1806-93472022v42n90-02.

21. Raboczkay T. Significado real do número de citações de um artigo. Jornal da USP [Internet] 2019. Available from: https://jornal.usp.br/artigos/significado-real-do-numero-de-citacoes-de-um-artigocientifico/.

22. Rizzatti M, Lampert Batista N, Cezar Spode PL, Bouvier Erthal D, Mauro de Faria R, Volpato Scotti AA, Trentin R, Petsch C, Turba Costa I, Quoos JH. Mapeamento da COVID-19 por meio da densidade de Kernel. RMA [Internet] 2020;3:44-53. Available from: https://doi.org/10.21166/metapre.v3i0.1312.

23. Souza Neto LT, Silva MTS, Vale WKS, Grigio AM. O uso do mapa de Kernel como subsídio para identificação da dispersão dos focos de queimadas no município de Mossoró (RN). Revista de Geografia, 2021; 38(2):244. Available from: https://doi.org/10.51359/2238-6211.2021.247425.

24. Sousa YSO. O Uso do Software Iramuteq : Fundamentos de Lexicometria para Pesquisas Qualitativas. Estudos e Pesquisas em Psicologia, 2021; 21(4):1541–60. Available from: https://pesquisa.bvsalud.org/portal/resource/pt/biblio-1359832.

25. Castro Neta AA, Cardoso BLC. O uso do *software* iramuteq na análise de dados em pesquisa qualitativa ou quali-quanti. Cenas Educ. [Internet] 2021;4:e11759. Available from: https://revistas.uneb.br/index.php/cenaseducacionais/article/view/11759.

26. Silva DD, de Almeida CC, Grácio MCC. Associação do Fator de Impacto e do Índice h para a avaliação de periódicos científicos: uma aplicação no campo da Ciência da Informação. EQ [Internet] 2018;24:132-51. Available from: https://doi.org/10.19132/1808-5245240.132-151.

27. Costa AP, Reis LP. Advantages and disadvantages of using software for qualitative data analysis. RISTI: Rev Ibér Sist Tecnol Inform. 2017;(23):ix-xii. Available from: http://www.scielo.mec.pt/scielo.php?script=sci_arttext&pid=S164698952017000300001&lng=pt&nrm=iso&tlng=pt.

28. Nunes JV, Woloszyn M, Gonçalves BS, Souza Pinto MD. Qualitative research supported by data analysis software: An investigation from examples. Rev Fronteiras Estudos Midiáticos. 2017;19(2):233-44. Available from: http://revistas.unisinos.br/index.php/fronteiras/article/view/fe.2017.192.08/6213.