European Journal of Cardiovascular Nursing (2023) 22, 547–551
https://doi.org/10.1093/eurjcn/zvad021

Methods Corner

New frontiers for qualitative textual data analysis: a multimethod statistical approach

Mariachiara Figura [1]*, Mary Fraire[2], Angela Durante [3], Angela Cuoco [1], Paola Arcadi [1], Rosaria Alvaro [1], Ercole Vellone [1], and Loredana Piervisani [1]

[1]Department of Biomedicine and Prevention, Biomedicine and Prevention Department, University of Rome 'Tor Vergata', via Montpellier 1, 00133, Rome, Italy; [2]Department of Social Research and Sociological Methodology (Ri.S.Me.S.), University of Rome 'La Sapienza', C.so Italia 38A, 00198 Rome, Italy; and [3]Predeparmental Nursing Unit, University of La Rioja, Calle Duquesa de la Victoria 88, 26004, 10 Logroño, Spain

Received 14 December 2022; revised 1 February 2023; accepted 2 February 2023; published 7 February 2023

In recent years, the increase in textual data production has meant that researchers require faster text analysis techniques and software to reliably produce knowledge for the scientific–nursing community. Automatic text data analysis opens the frontiers to a new research area combining the depth of analysis typical of qualitative research and the stability of measurements required for quantitative studies. Thanks to the statistical–computational approach, it proposes to study more or less extensive written texts produced in natural language to reveal lexical and linguistic worlds and extract useful and meaningful information for researchers. This article aims to provide an overview of this methodology, which has been rarely used in the nursing community to date.

Keywords Qualitative research • Automatic textual analysis • Multimethod approach • Multidimensional qualitative method • Rigour

Learning objectives

- Describe a methodological approach to textual data that overcomes common problems in traditional qualitative analysis
- Describe a multimethod technique to enhance methodological rigour, reliability, and rapidity in textual data analysis
- Develop a strategy to overcome issues related to the treatment of complex qualitative data such as natural language and extensive collections of texts (i.e. Big Data)

Introduction

Qualitative research is widely used in nursing and caring sciences to generate knowledge about human phenomena by considering in-depth, context-driven details.[4,5] The main potential of qualitative research lies in exploring the complex meanings of social phenomena experienced by individuals in their natural context.[6] Supported by standard scientific criteria, checklists, and guidelines, such as the commonly used COREQ checklist, amongst others,[7–11] qualitative research methods involve the systematic collection, organization, and interpretation of text derived from documents, discourses, or observations.[12] Data are analysed using step-by-step processes, following precise procedural rules to ensure reliability.[9,13] It is considered a real investment in studying complex phenomena or constructing a new theory.

Although qualitative analysis is excellent in comparing different perspectives of the same phenomenon by adding meaning to the quantitative value,[6] some authors report that researchers have often experienced time-consuming and costly data analysis techniques.[3,14,15] For examples, researchers usually deal with many pages of qualitative data offering unique stories and perspectives; in contrast to faster quantitative methods, they have to read texts several times to extract codes and get close enough to the meaning units and themes, trying to retain the integrity of each respondent's story.[11,16,17] Patterns, themes, and categories do not emerge on their own but demand intellectual work, and it becomes more challenging when dealing with profuse and unstructured textual data, such as Big Data or natural language (NL).[18,19]

Actually, some of the best-known software for qualitative analysis (e.g. ATLAS, NVivo) provide some features that facilitate the researcher in data analysis but they are only available for a semi-automatic analysis. It means that they assist the researcher in coding and extracting meaning units and themes, but they typically do not analyse the data automatically; instead, they make them more manageable and easier

* Corresponding author. Tel: +39 3396222975, Email: chiarafigura9@gmail.com

Box 1 Table of definitions[1]

Multidimensional analysis	Multidimensional analysis of textual data is a set of statistical techniques for analysing large amounts of data from different points of view (dimensions) to interpret complex phenomena. It is characterized by the joint observation of *k* variables (v.) over *n* statistical units (v.). The multidimensional analysis includes three groups of statistical methods: (1) classificatory (cluster analyses); (2) factorial for two-way tables (principal component analysis, simple and multiple correspondence analysis, multidimensional scaling, etc.); and (3) analysis for multiple tables (three-way and multi-way data analyses). These are analyses with a solid computational basis and, therefore, only possible with computers and the appropriate advanced software.[1,2]
Lexical corpora/ corpus	Collection of text (i.e. interviews, journal articles, book chapters) considered consistent and relevant to be studied from some point of view or property.[2]
Latent dimensions/ content	Latent content is not directly observable and consists of the interpretation of the meaning underlying the text. Differently from the manifest content, which easily emerges from the text (it is defined as 'close to the text'), the latent content (defined as 'distant from the text') refers to something like the 'red thread' between the lines. During analysis, the researcher often begins by sorting the manifest content coded into categories and continues to look for latent content and formulates it as themes at various levels. In this way, the researcher takes different scientific positions depending on the study's objective.[3]

to handle.[13] Moreover, managing a large amount of data runs the risk of mistakes in analysis and interpretation, and except for some features, analysis and interpretation of qualitative data are consequentially still laborious for the researcher.[11] Finally, there is ample evidence of the difficulty for researchers to extract latent content (Box 1), except by adopting laborious analysis techniques.[3] Although there are a lot of great strengths to conventional approaches to qualitative data, it is still little is known about the practical use of newer data analysis techniques and their strengths and pitfalls in the analysis of complex qualitative data, and there are few published studies.[20,21]

Adopting an innovative multimethod approach (Box 1) connecting qualitative data and quantitative analysis helped by sophisticated software that automatically analyses data (such as IRaMuTeQ, Lexico, and T-Lab) allows qualitative researchers to take advantage of various strategies and ensure rigour, rapidity, and originality of in-depth qualitative data analysis.[5,22] Moreover, it can help overtake issues that led to the researcher's influence on data interpretation and help analyse data collected from larger and more representative samples. This is a need that also emerges in cardiovascular nursing, an area in which many qualitative studies have been conducted. Examples include the lived experience of living with cardiovascular disease,[23] the processes of heart failure trajectory,[24] and cultural orientation in cardiovascular disease recovery.[21,25]

Concerning technological development and the increase of electronic sources, automatic analysis of textual data (AATD) could represent a substantial innovation in qualitative research. Maintaining the hermeneutic character and the typical characteristics of qualitative research, it is able to increase qualitative analysis' credibility and trustworthiness compared with traditional methods,[22,26,27] considering latent dimension extractions and context analysis in data interpretation.[28] Little is known about the practical use of AADT and its strengths and pitfalls in supporting the complexity of qualitative data analysis in nursing science. This article aims to shed light on this new and exciting area of research and to give an overview of the usefulness of AATD in nursing studies, which is still little known to nurse researchers.[29–31]

Overview of the methodology

First developed by Reinert,[32] 'AATD proposes a qualitative analysis strongly integrated with the quantitative one to ensure the stability of the measures'.[2] The process aims to extract the underlying real-world lexical corpora (Box 1) of entire documents (e.g. interviews, monologues, debates),[33] by applying statistics on textual data from an exploratory–descriptive perspective[1] and using software that can analyse texts automatically.[2]

Statistical measurements are applied following Fraire's Exploratory Multidimensional Data Analysis model (EMDA),[1] in which multivariate variable-driven statistics allow the interpretation of complex phenomena (such as disease-related phenomena) related to the context. EMDA includes several types of statistical techniques, such as factorial analysis (e.g. Principal Component Analysis and simple and multiple Correspondence Analysis), and classifier methods, such as Classificatory Hierarchical Dendrograms.[2] Thanks to factorial analysis it is possible to extract words' or classes of words's proximity through their projection on a factorial plane, allowing for the exploration of lexical profiles and latent semantic dimensions.[34–36] By the use of clustering, that is an unsupervised process based on algorythms, it is possible to classify texts with similar vocabulary.[2,37] Text Mining (TM) and Latent Semantic Analysis (LSA) provide for the information extraction and the meanings attribution. Specifically, TM encodes unstructured textual data, automatically associates information, and extracts relevant meanings.[18,38–40] Latent semantic analysis is an advanced TM approach[41,42] providing for not only the text's explicit information for extraction, but also the semantic structures that are partially hidden by the randomness of word placement (latent information), promoting the recognition of more relevant meanings.[43,44]

Finally, thanks to the use of sophisticated software, it is possible to generate graphs that, with immediate impact, describe word proximities, similarities, distances, contrasts, and thematic patterns that emerge by applying statistics. It allows a more significant amount of distinct analyses based on the same corpus of data.[37,46]

Amongst the advantages offered by AATD, firstly, it is possible to overcome issues related to the different interpretations of the same texts.[47] Secondly, it allows for the reliable analysis of extensive collections of texts and complex qualitative data (i.e. Big Data and Natural Language) without prior reading. Finally, this approach makes it possible to compare single parts of the same text and different texts, a procedure that traditional methods cannot perform.[2] It is essential to emphasize that software automatically analyses data without eliminating or replacing the researcher's role.[27] Instead, the researcher plays a central role in making data robust for both analysis and interpretation that, guided by theoretical frameworks, allow researchers to develop themes and identify multiple subjectivities without preconceived ideas.[28,48–50] However, scientific rigour and complete efficiency in managing and retrieving qualitative data will depend on the researcher's knowledge of the software and its functionality, their mastery of computer technology, and their ability to analyse organized data.[22,27]

Concerning the disadvantages of the method, decontextualization of words could occur if context analysis is not done correctly, as researchers work first on words and then on concepts. As a result, researchers might find it challenging to catch linguistic ambiguities if they are not adequately trained. Finally, the possible excess of automaticity and standardization of processes could be questioned. While it strengthens the rigour and trustworthiness of research, traditional qualitative researchers might find it difficult to accept the quantification of qualitative phenomena and concepts and the translation of texts and words into numbers and indices.

However, it is also true that we can affirm the two-dimensionality of research data. In a sense, all data are qualitative: all the data we collect correspond to subjective, psychic, and cultural phenomena that we try to translate into words to form empirical bases helpful in explaining social phenomena and controlling our theories. But we could also say that all data are quantitative because it is always possible to convert the language of words into the language of numbers through a process of encoding and then, in turn, lead the numbers (or rather the measures) and the relationships identified amongst the numbers back to interpretations and explanations that can be nothing more than ordered sequences of words endowed with meaning.[28]

Step-by-step approach

According to Fraire,[1] there are seven essential steps to carrying out EMDA (*Central illustration*). The first four steps are the preliminary phases in which data must be organized for analysis.

Step one: in this phase, the investigator defines the object and purpose of the research, the material, and data collection to address all the further analysis.

Step two is called 'a priori coding'. In this phase, researchers start working on data. All texts collected (i.e. the set of interviews or articles) are organized into a single document, called 'textual corpus', corresponding to the initial data matrix. In this phase, some pre-processing and processing operations on text (normalization, lexicalization, and lemmatization) are required to permit the software to recognize words and work on them. Normalization aims to remove typos and standardize spaces, apostrophes, and accents; through lemmatization, it is possible to recognize the grammatical categories and lead the words to their basic form, useful, for example, for word counting; lexicalization is needed to identify composed words and transform them into one by placing an underscore between them. Then, according to the aim, the corpus can be subdivided into smaller units by putting metadata lines between them. Smaller units could be single texts (i.e. an interview or a single chapter of a book), fragments (i.e. the single answer to a question or a paragraph), and elementary sense units (single words). Metadata lines (i.e. **** *n_1, **** n_2 up to **** *n_21) are command lines created by external variables, according to the aim. For example, if your corpus is composed of interviews or focus group discussions, you can run the analysis based on the questions, themes treated, or interviewers' characteristics. Instead, if you have a corpus of journal articles or book chapters, your command line can present variables related to 'journal' or 'book' references. These command lines will permit the selection of the variable guiding the multivariate analysis described later.

Step three: this is called 'a posteriori coding'. It is about lexicometric (or lexical) analysis, which means a first statistical description called 'lexical balance'. It is based on the primary standard criteria (e.g. frequency, co-occurrences, and proximity of the words) of the initial data matrix and provides a descriptive overview of your text. From this analysis, you can obtain three primary data: the total number of occurrences (i.e. frequency of words) determining the corpus size (N), the largeness of vocabulary (V) (i.e. the number of unique words in the text), and the number of Hapax (H) (words occurring only once in the text or rare forms). According to Bolasco,[50] to proceed to multidimensional analysis, some prerequisites lead to consistency, and statistical reliability must be satisfied. For these reasons, some indices, such as linguistic richness (V/N) providing information about the language richness, percentage of Hapax, and the number of total occurrences in the text (at least 25 000 needed), must be calculated.

Step four: the initial data matrix (based on segment frequencies) is coded into contingency tables compatible with multivariate analysis on which statistical measures and TM will be applied.

Once the preliminary steps have been completed, the researcher proceeds with the multidimensional analysis with the following three steps.

Step five consists of the 'choice of the measure' to apply to the just-produced contingency tables. It is very similar to quantitative, multidimensional analysis but adapted to the type of data table (contingency in this case). The most used statistical measures the researcher can choose are the scalar product, cosine (standardized measure), and chi-square (χ^2). The choice of measure depends on the aim of the analysis and the type of matrix (lexical or textual) the researcher wants to work on. For further information, see 'The Automatic Analysis of Texts. Doing Research with Text Mining' by Bolasco.[2]

Step six: this step enables the development of multidimensional statistics (i.e. clustering and factorial plans).

Step seven: This step returns the summary outputs of the results, both numerical (eigenvalues, factorial weights, factorial scores, trajectories, etc.) and graphical (graphs of factorial plans, correlation circles, dendrograms, etc.).[1]

Software

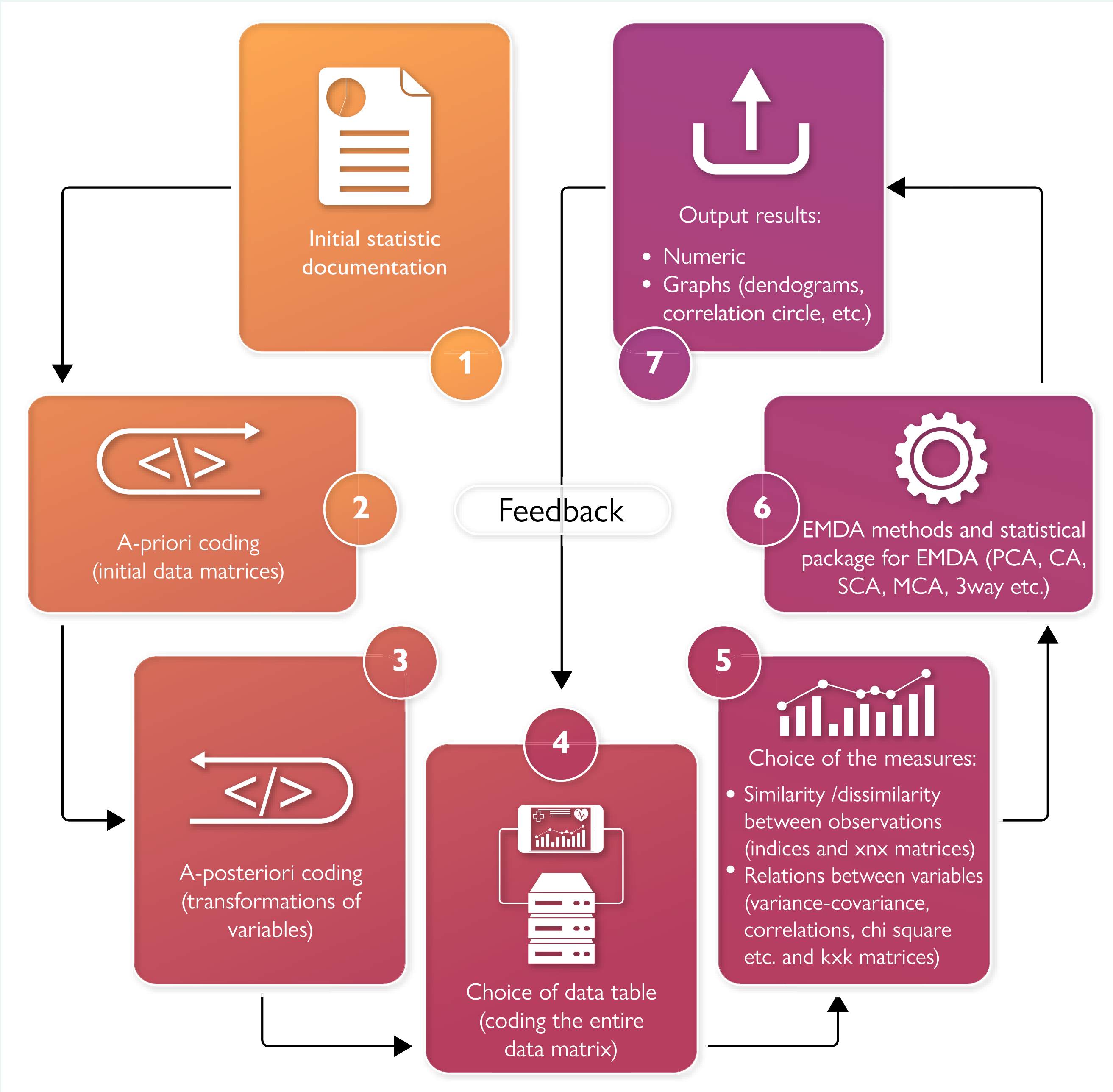
There are several software packages for data analysis. The most commonly used by researchers are IRaMuTeQ 0.7 alpha 2,[51] Taltac,[52] Lexico,[53] and T-Lab.[54] Given the complex and technical discussion of single software descriptions and their selection criteria (not feasible in the article), the authors refer to Bolasco's[2] papers or software websites for a more in-depth discussion.

Example of exploratory multidimensional data analysis in the cardiovascular field

Although EMDA is just beginning to make its way into the nursing field, to our knowledge, only one study has been published in the cardiovascular field.[21] The authors applied multidimensional statistics on interviewees to investigate caregivers' needs and the challenges of individuals with heart failure related to their sociodemographic characteristics. Applying EMDA, it was possible to obtain findings associated with sociodemographic characteristics, such as country of origin, age, gender, and the kind of informal caregiving relationship with the patient, highlighting that they are continually trying to cope with their social isolation and deteriorating health.

Reporting

Since the AADT is a multimethod approach that originates from qualitative data and provides structured information, a scientific paper needs to primarily report the descriptive indices of the analysis related to the lexical balance. Moving on to the description of the graphs, the



Central illustration

classificatory hierarchical dendrogram must be discussed in terms of class relationships, the theme that emerged, and the percentage of variance covered by each class. Factorial plans, showing, respectively, words' and classes of words' proximity projections, need to be described from both lexical and textual points of view, and the relationships that emerge from the content of the graphs must be interpreted in terms of meaning.[34,45] It is essential to report the factors extracted and their percentage of variance covered.

Visualization

For further clarification of the practical use of the entire process of EMDA on textual data, readers can refer to the article 'The nurse in the mirror: image of the female nurse during the Italian fascist period',[20] where also graphs and their respective interpretations are clearly shown.

Conclusion

Nursing is recognized as a human science because it understands experiences as humans live them. Automatic analysis of textual data is a versatile, person-centred strategy that allows us to study relationships between previously unobserved questions and subgroups. Although AADT exploits computational strategies and statistical measurements with methodological rigour increasing the reliability of the analysis, it should not be conducted without full consideration of theory, previous research, and the clinical relevance of the results.[5,27,29]

Author contributions

Rosaria Alvaro (Professor), Paola Arcadi (PhD student), Loredana Piervisani (Research Fellow), Ercole Vellone (Associate Professor), Mary Fraire (Professor of Statistics), Mariachiara Figura (MSN), Angela Cuoco (PhD student), and Angela Durante (Associate Professor).

Acknowledgements

The authors would like to thank the people who participated in this study.

Funding

This paper received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Conflict of interest: None declared.

Data availability

The authors confirm that the data supporting the findings of this study are available within the article (and/or its supplementary materials).

References

1. Fraire M. Statistical methods for exploratory multidimensional data analysis on time use. *Statistica* 2009; **69**: 317–341.
2. Bolasco S, De Mauro T. *L'analisi automatica dei testi: fare ricerca con il text mining*. Carocci editore; 2013.
3. Graneheim UH, Lindgren BM, Lundman B. Methodological challenges in qualitative content analysis: a discussion paper. *Nurse Educ Today* 2017;**56**:29–34.
4. Carr LT. The strengths and weaknesses of quantitative and qualitative research: what method for nursing? *J Adv Nurs* 1994;**20**:716–721.
5. Góes FGB, Santos ASTD, Campos BL, Silva ACSSD, Silva LFD, França LCM. Utilização do software IRAMUTEQ em pesquisa de abordagem qualitativa: relato de experiência. *Revista de Enfermagem da UFSM* 2021;**11**:e63.
6. Froggatt KA. The analysis of qualitative data: processes and pitfalls. *Palliat Med* 2001;**15**: 433–438.
7. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. *Int J Qual Health Care* 2007;**19**:349–357.
8. Lincoln YS. Emerging criteria for quality in qualitative and interpretive research. *Qualitative Inquiry* 1995;**1**:275–289.
9. Lincoln YS, Guba EG. But is it rigorous? Trustworthiness and authenticity in naturalistic evaluation. *New Directions for Program Evaluation* 1986;**1986**:73–84.
10. Cypress BS. Rigor or reliability and validity in qualitative research: perspectives, strategies, reconceptualization, and recommendations. *Dimens Crit Care Nurs* 2017;**36**:253–263.
11. Dierckx de Casterlé B, De Vliegher K, Gastmans C, Mertens E. Complex qualitative data analysis: lessons learned from the experiences with the qualitative analysis guide of Leuven. *Qual Health Res* 2021;**31**:1083–1093.
12. Khankeh H, Ranjbar M, Khorasani-Zavareh D, Zargham-Boroujeni A, Johansson E. Challenges in conducting qualitative research in health: a conceptual paper. *Iran J Nurs Midwifery Res* 2015;**20**:635–641.
13. Renz SM, Carrington JM, Badger TA. Two strategies for qualitative content analysis: an intramethod approach to triangulation. *Qual Health Res* 2018;**28**:824–831.
14. Malterud K. Qualitative research: standards, challenges, and guidelines. *Lancet* 2001; **358**:483–488.
15. Boddy CR. Sample size for qualitative research. *Qual Market Res Int J* 2016;**19**:426–432.
16. Sasso L, Bagnasco A, Ghirotto L. La ricerca qualitativa. Una risorsa per i professionisti della salute; 2015.
17. Jormfeldt H. The scientific position of qualitative studies-comprehensive understanding of health and well-being. *Int J Qual Stud Health Well-being* 2019;**14**:1667661.
18. Bolasco S, Bisceglia B, Baiocchi F. Estrazione automatica d'informazione dai testi. 2004; **3**.
19. Fraire M, Spagnuolo S, Stasi S. L'utilizzo dei big social data per la ricerca sociale: il caso della cittadinanza attiva in difesa del territorio. doi: 10.3280/SR2016-109014. Published online ahead of print.
20. Piervisani L, Palombo A, Albanesi B, Rocco G, Stasi S, Vellone E, *et al.* The nurse in the mirror: image of the female nurse during the Italian fascist period. *J Adv Nurs* 2021;**77**: 957–972.
21. Durante A, Cuoco A, Boyne J, Brawner B, Juarez-Vela R, Stasi S, *et al.* Needs and problems related to sociodemographic factors of informal caregiving of people with heart failure: a mixed methods study in three European countries. *J Adv Nurs* 2022;**78**:3034–3047.
22. Acauan LV, Abrantes CV, Stipp MAC, Trotte LAC, Paes GO, Queiroz ABA. Use of the IraMuteq® software for quantitative data analysis in nursing: a reflective essay. *Reme Revista Mineira de Enfermagem* 2020;**24**:1–5.
23. Simeone S, Pucciarelli G, Perrone M, Dell'Angelo G, Rea T, Guillari A, *et al.* The lived experiences of the parents of children admitted to a paediatric cardiac intensive care unit. *Heart Lung* 2018;**47**:631–637.
24. Riegel B, Dickson VV, Faulkner KM. The situation-specific theory of heart failure self-care: revised and updated. *J Cardiovasc Nurs* 2016;**31**:226–235.
25. Dickson VV, McCarthy MM, Howe A, Schipper J, Katz SM. Sociocultural influences on heart failure self-care among an ethnic minority black population. *J Cardiovasc Nurs* 2013; **28**:111–118.
26. Soares SSS, Souza N, Carvalho EC, Queiroz ABA, Costa C, Souto J. COVID-19 pandemic and nursing week: analysis from software Iramuteq. *Rev Bras Enferm* 2021;**75**: e20200690.
27. Soares SSS, Costa CCP, Carvalho EC, Queiroz ABA, Peres PLP, Souza NVDO. Teaching Iramuteq for use in qualitative research according to YouTube videos: an exploratory-descriptive study. *Rev Esc Enferm USP* 2022;**56**:e20210396.
28. Giuliano L LRG. L'analisi automatica e semi-automatica dei dati testuali: Strategie di ricerca e applicazioni. 2 ed; 2010.
29. Jusoh S, Alfawareh HM. Techniques, applications and challenging issue in text mining. 2012.
30. Kami MTM, Larocca LM, Chaves MMN, Lowen IMV, Souza VMPD, Goto DYN. Working in the street clinic: use of IRAMUTEQ software on the support of qualitative research. *Escola Anna Nery—Revista de Enfermagem* 2016;**50**:1–5.
31. Souza MAR, Wall ML, Thuler A, Lowen IMV, Peres AM. The use of IRAMUTEQ software for data analysis in qualitative research. *Rev Esc Enferm USP* 2018;**52**:e03353.
32. Reinert M. Alceste une methodologie d'analyse des donnees textuelles et une application: aurelia de gerard de nerval. *BMS Bull Sociol Methodol/Bull Méthodol Sociol* 1990;**26**:24–54.
33. Robieux L, Karsenti L, Pocard M, Flahault C. Let's talk about empathy!. *Patient Educ Couns* 2018;**101**:59–66.
34. Neta AAC. The use of the IRAMUTEQ software in data analysis in qualitative or quali-quanti research. *Cenas Educacionais* 2021;**4**:1–17.
35. Leblanc J-M. Proposition de protocole pour l'analyse des données textuelles: pour une démarche expérimentale en lexicométrie. *Nouvelles Perspectives en Sciences Sociales* 2015;**11**:25–63.
36. Veraszto EV, Camargo EPD, Camargo JTFD, Simon FO, Miranda NAD. Evaluation of concepts regarding the construction of scientific knowledge by the congenitally blind: an approach using the correspondence analysis method. *Ciência and Educação (Bauru)* 2018;**24**:837–857.
37. Monteiro L, Melo RD, Braga B, Sá JD, Monteiro L, Cunha M, Canuto A. ALCESTE X IRAMUTEQ: comparative analysis of the use of CAQDAS in qualitative research. *World Conf Qual Res*;**2021**:67–79.
38. Ananiadou S, Chruszcz J, Keane J, McNaught J, Watry P. The National Centre for Text Mining: aims and objectives. *Ariadne* 2005;**42**.
39. Mingo I, Nocenzi M. The dimensions of gender in the last twenty years: an analysis of the International Review of Sociology. *Int Rev Sociol* 2020;**30**:6–25.
40. Talib R, Kashif M, Ayesha S, Fatima F. Text mining: techniques, applications and issues. *Int J Adv Comput Sci Appl* 2016;**7**:414–418.
41. Aryal A, Gallivan MJ, Tao Y. Using latent semantic analysis to identify themes in IS healthcare research. *AMCIS* 2015.
42. Maletic JI, Marcus A. Using latent semantic analysis to identify similarities in source code to support program understanding. *IEEE Comput Soc* 2000.
43. Landauer TK, Foltz PW, Laham D. An introduction to latent semantic analysis. *Discourse Process* 1998;**25**:259–284.
44. Evangelopoulos NE. Latent semantic analysis. *WIREs Cognitive Sci* 2013;**4**:683–692.
45. Six C. Analyse lexicale appliquée à une question ouverte à l'aide d'IRaMuTeQ. *Sciences du Vivant [q-bio]* 2019:1–20.
46. Canuto A, Braga B, Monteiro L, Melo R. Aspectos críticos do uso de caqdas na pesquisa qualitativa: uma comparação empírica das ferramentas digitais alceste e iramuteq. *New Trends Qual Res* 2020;**3**:199–211.
47. Peyrat-Guillard D, Miltgen C, Welcomer S. Analysing conversational data with computer-aided content analysis: the importance of data partitioning; 2014.
48. Ramos MG, do Rosário Lima VM, Amaral-Rosa MP. IRAMUTEQ software and discursive textual analysis: interpretive possibilities (ed.), *Computer supported qualitative research.* Brazil: Springer International Publishing; 2019. p58–72.
49. Nascimento Martins K, Sarro Gomes LP, Corrêa de Paula M. Software IRaMuTeQ: uma ferramenta auxiliar na análise textual discursiva. *Paradigma* 2022;**43**:205–227.
50. Bolasco S. Introduction to the automatic analysis of textual data via a case study. *Stat Appl Ital J Appl Stat* 2012;**22**:5–19.
51. IRaMuTeQ 0.7 alpha 2. In.
52. Taltac. In.
53. Lexico. In.
54. T-Lab. In.