

USO DO SOFTWARE IRAMUTEQ NAS PESQUISAS BRASILEIRAS DA ÁREA DA SAÚDE: UMA SCOPING REVIEW

Use of IRAMUTEQ software in the Brazilian healthcare research: a scoping review

Pétala Tuani Candido de Oliveira Salvador

Universidade Federal do Rio Grande do Norte - UFRN - Natal (RN) - Brasil

Andréa Tayse de Lima Gomes

Universidade Federal do Rio Grande do Norte - UFRN - Natal (RN) - Brasil

Cláudia Cristiane Filgueira Martins Rodrigues

Universidade Federal do Rio Grande do Norte - UFRN - Natal (RN) - Brasil

Flávia Barreto Tavares Chiavone

Universidade Federal do Rio Grande do Norte - UFRN - Natal (RN) - Brasil

Kisna Yasmin Andrade Alves

Universidade Federal do Rio Grande do Norte - UFRN - Natal (RN) - Brasil

Manacés dos Santos Bezerril

Universidade Federal do Rio Grande do Norte - UFRN - Natal (RN) - Brasil

Viviane Euzébia Pereira Santos

Universidade Federal do Rio Grande do Norte - UFRN - Natal (RN) - Brasil

RESUMO

Objetivo: Identificar o uso do *Interface de R pour les Analyses Multidimensionnelles de Textes et de Questionnaires* (IRAMUTEQ), nas pesquisas qualitativas produzidas no âmbito dos programas de pós-graduação da área da saúde no Brasil. **Métodos:** Os métodos empregues implicaram a realização de uma *scoping review*, em que se analisaram dissertações e teses produzidas, no âmbito dos programas de pós-graduação da área da saúde e que se configuravam como pesquisas qualitativas, com uso do IRAMUTEQ. **Resultados:** Os dados foram analisados por estatística descritiva simples. Os resultados indicaram uma amostra final que consistiu em 54 pesquisas. Verificaram-se certas inconsistências no uso do IRAMUTEQ, tais como as seguintes: mencionar o do IRAMUTEQ como técnica de análise de dados; utilizar pouco ou limitado volume textual; e usar a Classificação Hierárquica Descendente, mesmo com a quantidade de retenção do texto inferior ao recomendado. Tais elementos precisam de ser adequadamente considerados pelo pesquisador, com o objetivo de se produzir análises de dados confiáveis e fundamentadas. **Conclusão:** Poder-se-á afirmar que o quantitativo de investigações que utilizaram o IRAMUTEQ foi crescente ao longo dos anos e, portanto, um tal *software* tem sido incorporado nas pesquisas qualitativas da área da saúde produzidas no âmbito dos programas de pós-graduação brasileiros.

Descritores: Software; Pesquisa Qualitativa; Programas de Pós-Graduação em Saúde.

ABSTRACT

Objective: To identify the use of *Interface de R pour les Analyses Multidimensionnelles de Textes et de Questionnaires* (IRAMUTEQ), in qualitative research produced within postgraduate programs in the health area in Brazil. **Methods:** The methods included a *scoping review* held from dissertations and theses configured as qualitative research and produced within the framework of postgraduate programs in the health area, were analysed using IRAMUTEQ. Data were analysed by simple descriptive statistics. **Results:** The results imply a final sample consisted of 54 researches. There were certain inconsistencies in the use of IRAMUTEQ, such as the following: mention of IRAMUTEQ as a data analysis technique; use little or limited textual volume; and use the Descending Hierarchical Classification, even with the amount of text retention lower than the recommended one. These elements need to be adequately considered by the researcher, in order to produce reliable and informed data analysis. **Conclusion:** It can be stated that the number of researches that have used the IRAMUTEQ has increased over the years and therefore, this software has been incorporated in the qualitative researches of the health area produced in the ambit of Brazilian postgraduate programs.

Descriptors: Software; Qualitative Research; Health Postgraduate Programs.



Este artigo está publicado em acesso aberto (Open Access) sob a licença Creative Commons, que permite uso, distribuição e reprodução em qualquer meio, sem restrições, desde que o trabalho seja corretamente citado.

Este artigo foi selecionado, corrigido e aprovado pelo 6º CIAIQ - Congresso Ibero-Americano em Investigação Qualitativa, seguindo suas normas e formatação.

INTRODUÇÃO

Os dados produzidos em pesquisas qualitativas são compostos por linguagens e, portanto, tornam-se relevantes aos estudos sobre pensamentos, crenças ou opiniões⁽¹⁾. A subjetividade dos dados produzidos, o grande volume textual de algumas pesquisas, somados à busca pelo maior rigor metodológico das investigações qualitativas tornam a fase de análise de dados crucial para o pesquisador qualitativo^(2,3).

Frente a tais desafios, o uso de *softwares* específicos para análise de dados textuais estão cada vez mais presente em estudos na área de Ciências Humanas, Sociais e da Saúde, especialmente naqueles estudos em que o *corpus* a ser analisado é bastante volumoso⁽¹⁾.

A análise de dados textuais ou análise lexical, apoiada em *softwares*, propõe que se supere a dicotomia clássica entre o domínio quantitativo e o domínio qualitativo, por possibilitar que se quantifique e se empregue cálculos estatísticos sobre variáveis essencialmente qualitativas - os textos⁽¹⁾.

Nesta perspectiva, desde a década de 1990, vários programas foram desenvolvidos para facilitar o gerenciamento de dados qualitativos. Estes tem evoluído ao longo dos anos, de modo que os *softwares* se tornaram mais especializados para o tipo específico de método de pesquisa qualitativa⁽²⁾.

Alguns programas só utilizam texto, enquanto outros podem importar imagens, dados de áudio e de vídeo, recortes de jornais e de livros. Também já estão disponibilizados aqueles que possibilitam definir e organizar informações, bem como analisar relacionamentos e temas nos dados⁽²⁾.

Dentre os inúmeros *softwares* para apoio à análise de dados qualitativos, destaca-se neste estudo o IRAMUTEQ (*Interface de R pour les Analyses Multidimensionnelles de Textes et de Questionnaires*), um programa gratuito (www.iramuteq.org) e com fonte aberta, desenvolvido pelo pesquisador francês Pierre Ratinaud, que utiliza o mesmo algorítmico do ALCESTE para análises estatísticas de textos⁽¹⁾.

Em 2009, Ratinaud desenvolveu-o na língua francesa, mas atualmente possui dicionários completos em outras línguas. No Brasil, seu uso iniciou-se a partir de 2013 e a área da saúde tem se apropriado dessa ferramenta⁽⁴⁻⁶⁾.

O IRAMUTEQ caracteriza-se como um método informatizado para análise de dados/textos, em que se busca apreender a estrutura e a organização do discurso, informando as relações entre os mundos lexicais mais frequentemente enunciados pelo sujeito⁽⁷⁾. Utiliza como base a estrutura do software R para a realização dos cálculos e da linguagem Python⁽⁸⁾. Assume que as palavras usadas em contexto similares estão associadas a um mesmo mundo lexical. Realiza, deste modo, análises quantitativas de dados textuais, pautadas em múltiplos contextos e segundo classes de conteúdo, com base na similaridade de vocabulário⁽⁹⁾.

Considera-se que o número de artigos que divulgam a utilização do IRAMUTEQ, no Brasil, na área da saúde, ainda é limitado⁽⁴⁾. Nesse sentido, se apresenta a questão de pesquisa do estudo: “como as pesquisas qualitativas, produzidas no âmbito dos programas de pós-graduação da área da saúde brasileiros, têm utilizado o IRAMUTEQ?”

MÉTODOS

Sendo uma *scoping review* construída a partir das recomendações do *JBI Institute Reviewer's Manual*⁽¹⁰⁾, procedimentos metodológicos detalhados encontram-se descritos em trabalho apresentado e publicado nos anais do 6º Congresso Ibero-Americano em Investigação Qualitativa (CIAIQ)⁽¹¹⁾.

O grupo de estudo teve por base dissertações e teses produzidas no âmbito dos programas de pós-graduação brasileiros, da área da saúde, que se configuravam como pesquisas qualitativas e que utilizaram o IRAMUTEQ.

Foi utilizada a estratégia PCC para a formulação da pergunta de pesquisa: P (*Population*) – Pesquisa qualitativa; C (*Concept*) – Software; C (*Context*) – Programas de pós-graduação brasileiros da área da saúde, o que resultou na seguinte questão: “como as pesquisas qualitativas produzidas no âmbito dos programas de pós-graduação brasileiros da área da saúde têm utilizado o IRAMUTEQ?”

A pesquisa foi realizada a partir do Banco de Teses e Dissertações da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), base de dados que reúne as dissertações e teses defendidas no contexto brasileiro. Elegeu-se esses documentos como fonte para a coleta de dados por se tratar de estudos que apresentam informações mais aprofundadas acerca da metodologia utilizada nas pesquisas.

A busca ocorreu por meio da combinação de dois descritores extraídos dos Descritores em Ciências da Saúde (DeCS), que corresponderam aos dois elementos iniciais do mneumônico PCC desta pesquisa – (P) Pesquisa Qualitativa AND (C) *Software*. Além disso, foi utilizado o filtro Área de Concentração disponível no próprio banco de

dados para selecionar as áreas correspondentes aos programas de pós-graduação da saúde, o que configurou o outro (C) do mneumônico utilizado.

Como critérios de inclusão, foram estabelecidos: dissertações e teses; produzidas no âmbito de programas de pós-graduação brasileiros da área da saúde; publicados na íntegra *online*, que se configuravam como pesquisas qualitativas com uso do IRAMUTEQ. Foram excluídas as dissertações e teses que utilizaram outro *software* para a análise de dados e que não tinham a versão completa disponível *online*.

Não foi estipulado limite temporal. Todavia, como o banco de dados utilizado passou por atualizações recentes, algumas limitações foram encontradas na busca e seleção dos estudos: 1) só estavam disponíveis resumos dos trabalhos publicados entre 2013 e 2016; e 2) a disponibilização dos resumos foi limitada aos primeiros 10.000 trabalhos.

A partir da estratégia de busca referida, foram identificados 48.124 trabalhos. Desse total, apenas 16.268 tinham os resumos disponíveis no *site*, os quais tiveram seu título e resumo avaliados, com base nos critérios de inclusão e exclusão estabelecidos. Todavia, isso só foi possível para os primeiros 10.000 trabalhos, por limitação estabelecida pelo próprio banco de dados.

Assim, do total de 6.268 pesquisas, foi identificado que 306 se configuravam como investigações qualitativas que tinham utilizado algum *software* de análise de dados. Deste último quantitativo, foram extraídas as pesquisas que usaram o IRAMUTEQ, o que totalizou a amostra final deste estudo: 54 trabalhos, cujas publicações foram selecionadas para leitura completa e procedeu-se com a extração dos dados (Tabela I).

Tabela I - *Softwares* utilizados nas pesquisas identificadas. Natal, Rio Grande do Norte, 2017.

Software	n	%*
Atlas.ti	69	22,5
Nvivo	58	19,0
Alceste	54	17,6
IRAMUTEQ	54	17,6
EVOC	33	10,8
Qualiquantisoft	18	5,9
WebQDA	8	2,6
Etnograph	7	2,3
Tri-deux-mots	6	2,0
MaxQDA	3	1,0
OpenLogos	3	1,0
oTranscribe	1	0,3
ELAN	1	0,3
AnCo-Redes	1	0,3
f4 Plus	1	0,3
Express Scribe	1	0,3

*A soma é maior que 100,0% pois alguns estudos utilizaram mais de um *software*

Os seguintes dados foram extraídos a partir de planilha construída no *Microsoft Excel 2010*: nível acadêmico; ano de publicação; Instituição de Ensino Superior; formação do autor (segundo Currículo *Lattes* do autor); tipo e abordagem de pesquisa; população estudada; etapa do uso do IRAMUTEQ; tipo de análise utilizada; benefícios e dificuldades enfrentadas no uso do *software*.

Os dados foram analisados a partir de estatística descritiva simples. Não foi necessária apreciação ética, por se trabalhar com estudos de domínio público.

RESULTADOS

Cinquenta e quatro pesquisas constituíram a amostra final deste estudo. Destas, 29 (53,7%) eram dissertações decorrentes de mestrado acadêmico, 15 (27,8%) de mestrado profissional e 10 (18,5%) eram teses de doutorado. O quantitativo de investigações que utilizaram o IRAMUTEQ foi crescente ao longo dos anos: 1 (1,9%) em 2013; 8 (14,8%) em 2014; 24 (44,4%) em 2015; e 21 (38,9%) em 2016.

Os trabalhos foram desenvolvidos no âmbito dos programas de pós-graduação de 16 Instituições de Ensino Superior. Tiveram destaque o Centro Universitário UNINOVAFAPI (14; 25,9%), a Universidade Federal do Paraná (11;

20,4%) e a Universidade Federal da Paraíba (8; 14,8%) (Figura 1). Ao agrupar as Instituições de Ensino Superior por sua localização geográfica, a região Nordeste brasileira teve o maior número de pesquisas analisadas (32; 59,2%).

Os autores das publicações avaliadas provinham de oito cursos da área da saúde. Predominaram os da Enfermagem (37; 68,3%), seguidos por: Nutrição (4; 7,4%); Fisioterapia (3; 5,6%); Medicina (3; 5,6%); Psicologia (3; 5,6%); Educação Física (2; 3,7%); Farmácia (1; 1,9%); e Odontologia (1; 1,9%).

Quanto ao tipo de pesquisa descrito pelos próprios autores, destacaram-se as investigações descritivas (23; 42,6%). As demais se caracterizaram como pesquisas descritivo-exploratórias (16; 29,6%), exploratórias (9; 16,7%) e avaliativas (2; 3,7%). Quatro (7,4%) trabalhos não explicitaram o tipo de pesquisa.

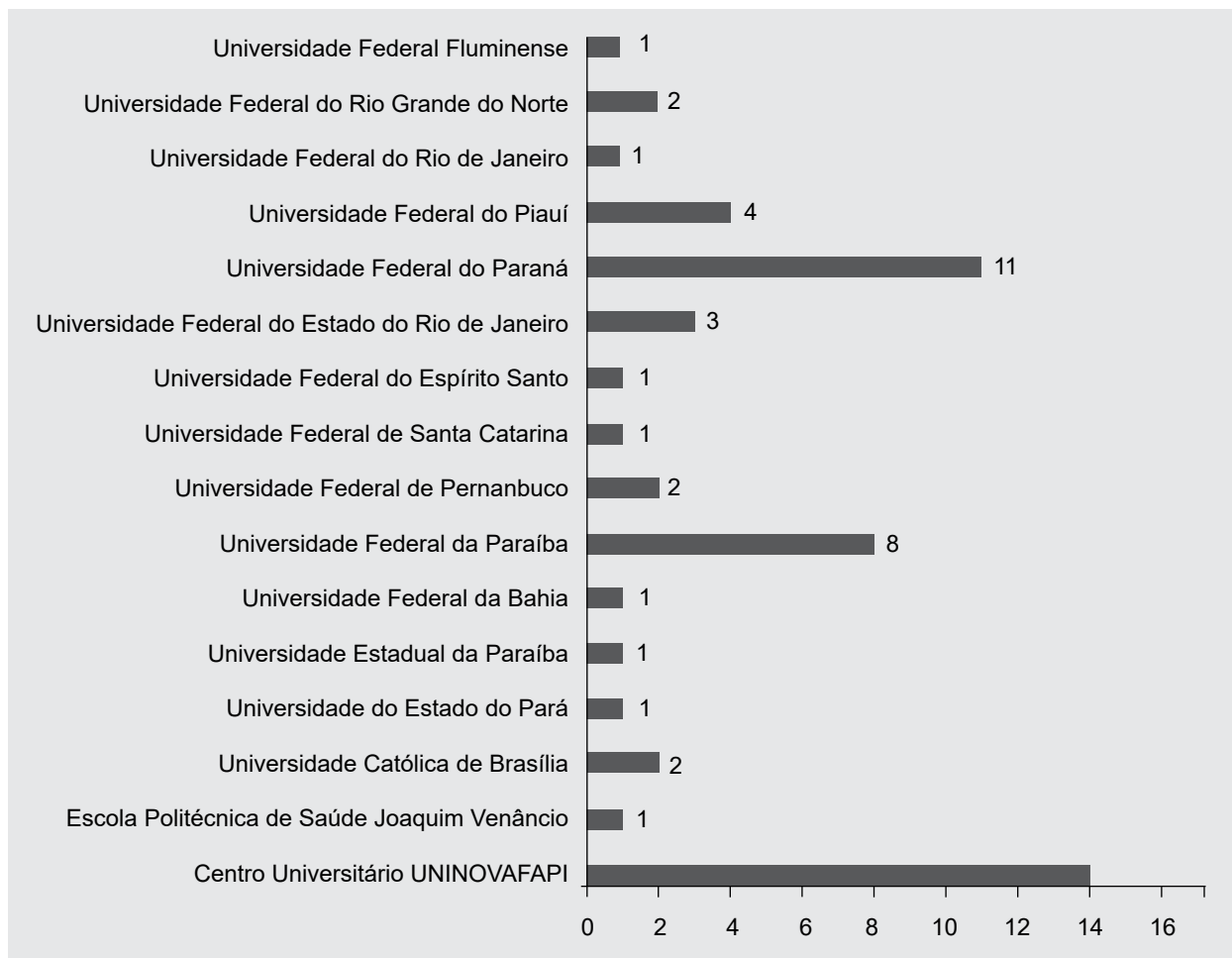


Figura 1 - Instituições de Ensino Superior em que as pesquisas foram desenvolvidas (em números absolutos). Natal, Rio Grande do Norte, 2017.

No que concerne à abordagem utilizada, 45 (83,3%) denotaram ser a perspectiva qualitativa, enquanto os outros 9 estudos (16,7%) revelaram ser a abordagem mista, com integração de análises quantitativas e qualitativas.

Houve predomínio de uso do IRAMUTEQ para tratamento de dados decorrentes de entrevistas, o que foi realizado em 53 (98,1%) pesquisas. Apenas 1 trabalho (1,9%) utilizou o *software* para analisar textos provenientes de documentos.

O número de participantes que compuseram a amostra dos estudos analisados variou de 6 a 406, de modo que 46 trabalhos (85,2%) utilizaram o IRAMUTEQ para sistematizar a análise de entrevistas de mais de 20 pessoas.

Quarenta e dois autores (77,8%) referiram o uso do *software* na etapa de tratamento de dados e os demais 12 autores (22,2%) utilizaram-no na etapa de análise de dados. Os tipos de análises de dados fornecidas pelo IRAMUTEQ e utilizadas pelos autores dos trabalhos analisados estão descritas na Tabela II.

Tabela II - Quantidade e tipos de análises utilizadas pelos autores (N=54). Natal, Rio Grande do Norte, 2017.

Quantidade de análises	n	%
Apenas 1	36	66,7
2 tipos de análises	10	18,5
3 tipos de análises	8	14,8
Tipos de análises	n	%*
Classificação Hierárquica Descendente	46	85,2
Análise de similitude	17	31,5
Nuvem de palavras	10	18,5
Análise Fatorial de Correspondência	7	13,0

*A soma é maior que 100,0% pois alguns estudos utilizaram mais de um tipo de análise.

Predominaram as pesquisas que utilizaram apenas um tipo de análise fornecida pelo IRAMUTEQ (36; 66,7%) e a Classificação Hierárquica Descendente (CHD) foi a mais usada (46; 85,2%). Desses estudos, 28 (61,0%) referiram uma retenção do texto analisado superior a 75,0%, 6 estudos (13,0%) um aproveitamento do *corpus* com valores entre 70 e 75%, 6 estudos (13,0%) tiveram retenção textual inferior a 70% e os outros 6 estudos (13,0%) não elucidaram a porcentagem de texto retido na análise do *software*.

Quanto aos benefícios de utilização do IRAMUTEQ, foram citados os seguintes: viabiliza diferentes tipos de análise (13; 24,1%); potencializa a organização dos dados da pesquisa (2; 3,7%); possibilita a análise das *falas* no contexto em que foram citadas (2; 3,7%); promove maior rigor metodológico à pesquisa qualitativa (2; 3,7%); e é gratuito (1; 1,9%). Destaca-se que 37 (68,5%) pesquisas não citaram os benefícios de uso do *software* analisado.

No que concerne às dificuldades enfrentadas no manuseio do IRAMUTEQ, apenas 2 autores (3,7%) denotaram a demanda de tempo para estudo do pesquisador como um elemento dificultador do uso do *software*. Cinquenta e dois autores (96,3%) não citaram as limitações experimentadas.

DISCUSSÃO

A predominância das dissertações resultantes de mestrados acadêmicos coaduna-se com as realidades exploradas na pós-graduação brasileira, em que seguindo uma tendência histórica é esse o tipo de pós-graduação que mais titula no país⁽¹²⁾.

Destaca-se, ainda, o número quantitativo expressivo de pesquisas desenvolvidas no âmbito de mestrados profissionais, modalidade de pós-graduação que tem crescido no Brasil⁽¹²⁾.

O ascendente de estudos que utilizaram o IRAMUTEQ ao longo dos anos faz parte do movimento crescente de uso de *softwares* de apoio a pesquisas qualitativas, em resposta às demandas atuais, em que se busca o rigor metodológico e a criatividade do pesquisador^(2-3,13).

Enfatiza-se que o IRAMUTEQ foi desenvolvido inicialmente em língua francesa e começou a ser utilizado no Brasil apenas em 2013, momento em que se compôs o dicionário em língua portuguesa por várias equipes: Laboratório de Psicologia Social da Comunicação e Cognição, da Universidade Federal de Santa Catarina (LACCOS/UFSC), em parceria com o Centro Internacional de Estudos em Representações Sociais e Subjetividade - Educação, da Fundação Carlos Chagas (CIERS-ed/FCC) e com o grupo de pesquisa Valores, Educação e Formação de Professores, da Universidade Estadual Paulista Júlio de Mesquita Filho (UNESP)⁽¹⁾.

Observou-se nesta pesquisa o destaque de uso do IRAMUTEQ nas Instituições do Nordeste brasileiro, com um significativo número de estudos desenvolvidos no estado do Piauí (18; 33,3%), o que revela uma apropriação dos pesquisadores desta região geográfica no uso do *software*.

Quanto à expressividade da Enfermagem, no desenvolvimento de pesquisas qualitativas com apoio de *softwares*, o resultado encontrado coaduna com outras pesquisas que já destacaram o maior número de investigações com abordagem qualitativa realizadas pelos pesquisadores da Enfermagem, o que decorre da possibilidade de estudos qualitativos alcançarem as respostas para inquietações particulares que envolvem o cuidado e a assistência de enfermagem, ou seja, pela adequação aos objetos de estudo das pesquisas dessa área de atuação, que também usa de *softwares* em suas investigações^(3,14,15).

Outro elemento revelado foi a predominância do uso do IRAMUTEQ para o método de análise de discurso. Denotou-se o uso do *software* para análise de textos de documentos em apenas um estudo.

Enfatiza-se que a análise textual realizada pelo IRAMUTEQ se configura a partir de material verbal transcrito, ou seja, de textos. Todavia, essa análise tem várias finalidades, sendo possível analisar além de textos, entrevistas, documentos, redações, etc. Pode-se a partir da análise textual e descrever um material produzido por um produtor, seja individual ou coletivo, como também utilizar a análise textual com a finalidade comparativa, relacional, comparando produções diferentes em função de variáveis específicas, que descrevem quem produziu o texto⁽¹⁶⁾.

Nesta perspectiva, destaca-se que o uso do IRAMUTEQ para analisar textos documentais já foi descrita em pesquisas realizadas na França⁽¹⁷⁾, no Brasil^(8-9,18) e no Reino Unido^(19,20). Este tipo de análise, portanto, pode ser melhor explorada nas pesquisas qualitativas da área da saúde brasileira.

Quanto ao número de entrevistas analisadas, no domínio das pesquisas identificadas neste estudo, denotou-se que 46 pesquisas (85,2%) realizaram uma análise de mais de 20 entrevistas. É válido destacar que o uso do IRAMUTEQ foi idealizado para grandes volumes textuais, de modo que há cuidados a serem considerados na composição do *corpus* de análise, para viabilizar um processo de tratamento de dados efetivo.

Nessa circunstância, no caso de entrevistas, a recomendação é a seguinte: havendo *falas* que produzem textos mais extensos, convém que o grupo seja homogêneo, sendo suficiente considerarem-se entre 20 e 30 textos. Se o delineamento é comparativo, sugere-se pelo menos 20 textos para cada grupo. Quando as respostas apresentarem uma média em torno de três ou quatro linhas, é necessário um número bem maior de respostas para a constituição de um *corpus* de análise⁽¹⁶⁾.

Quanto à etapa da pesquisa qualitativa em que o IRAMUTEQ foi utilizado, apesar de a maioria dos estudos (42; 77,8%) referirem seu uso para o tratamento dos dados, merece destaque que 22,2% dos trabalhos mencionarem o uso do *software* para análise dos dados.

Ressalta-se que o uso do IRAMUTEQ não é um método de análise de dados, mas uma ferramenta para processá-los. Portanto, não conclui essa análise, já que a interpretação é essencial e de responsabilidade do pesquisador⁽²¹⁾. Por conseguinte, esse material deve ser compreendido como um *software* qualitativo de análise de dados - léxicos - para auxiliar e complementar a análise de conteúdo⁽²²⁾.

Nenhuma das ações realizadas pelo IRAMUTEQ produz resultados. Em alternativa, o pesquisador utiliza todos esses materiais, juntamente com o texto original, para construir a compreensão dos discursos e ajudar a identificar as inferências mais plausíveis, a partir dos dados. Importa ainda dizer que esse não é um processo automático, mas uma etapa analítica em que o pesquisador é o elemento central⁽¹⁹⁾.

Pesquisadores já revelaram uma preocupação em torno do fenômeno que denominaram de “fascínio” exercido pelos *softwares* de análise textual, o que resulta em inúmeras publicações que citam o próprio *software* como se fosse a técnica de análise dos dados, e ainda, como se fosse o método da pesquisa⁽¹⁾.

Quanto às análises realizadas pelo IRAMUTEQ, destaca-se que este viabiliza diferentes tipos de análise de dados textuais, desde a lexicografia básica, que abrange sobretudo a lematização e o cálculo de frequência de palavras; até análises multivariadas, com destaque para a CHD, a Análise Fatorial de Correspondência (AFC) e a análise de similitude, além da nuvem de palavras⁽¹⁶⁾.

A CHD foi a mais utilizada nos trabalhos analisados. Neste processo de análise, o vocabulário do *corpus* é identificado e quantificado em relação à frequência e à sua posição no texto, ou seja, é submetido a cálculos estatísticos para posterior interpretação. Esta análise visa obter classes de segmentos de texto que, ao mesmo tempo, apresentam vocabulário semelhante entre si, bem como vocabulário diferente dos segmentos das outras classes^(1,16).

Assim considerada, a classe é compreendida como um conjunto de palavras que aparecem próximas umas das outras, com a formação de um segmento específico. Com base nessas palavras, pontos centrais do texto podem ser apontados, além de se compreender como os termos estão associados uns aos outros. Essas classes representam o ambiente de sentido das palavras e indicam representações sociais ou elementos de representações sociais sobre o objeto social em estudo⁽⁹⁾.

Com a CHD, o IRAMUTEQ organiza as palavras em um dendograma, que representa a quantidade e composição léxica de classes, a partir de um agrupamento de termos, do qual se obtém a frequência absoluta de cada um deles e o valor de qui-quadrado agregado⁽⁷⁾.

Todavia, nem sempre o uso da CHD é possível. Dentre os elementos centrais para o uso adequado desse tipo de análise está a percentagem de texto do *corpus* que o IRAMUTEQ considera para efetuar a CHD. É importante salientar que as análises do tipo CHD, para serem úteis à classificação de qualquer material textual, requerem uma retenção mínima de 75% dos segmentos de texto, sendo que certos autores reconhecem a possibilidade de se considerar o aproveitamento de 70% dos segmentos de textos⁽¹⁶⁾.

Caso a CHD ofereça uma classificação com retenção inferior a esta percentagem, este método não pode ser considerado para uma análise adequada aos dados coletados, pois fornece apenas uma classificação parcial. Nessas condições (quando a retenção de segmentos de texto for inferior a 75%), sugere-se que o método de CDH seja abandonado e a análise do material seja feita com outros recursos, como por exemplo, a análise de especificidades⁽¹⁶⁾.

Encontraram-se 6 estudos (13,0%) com retenção do texto com valores entre 70 e 75%, 6 estudos (13,0%) com aproveitamento do texto inferior a 70% e 6 estudos (13,0%) em que não foi referido o percentual de texto que foi retido. Estes dados são preocupantes, pois revelam análises inadequadas, a partir do IRAMUTEQ.

A análise de similitude foi a segunda forma de análise mais utilizada nas pesquisas analisadas. Este tipo de análise, também denominado “de semelhanças”, apoia-se na teoria dos grafos, em que se considera que um grafo constitui o modelo matemático ideal para o estudo das relações entre objetos discretos de qualquer tipo e possibilita identificar as coocorrências entre as palavras e o seu resultado⁽²³⁾.

Neste tipo de análise, o tamanho das palavras e a espessura dos traços que as unem têm significado a evidenciar, o que traz indicações da conexão entre as palavras e auxilia na identificação da estrutura de um *corpus* textual^(1,16).

A nuvem de palavras foi a terceira forma de análise mais utilizada. Agrupa e organiza as palavras graficamente em função da sua frequência. É a análise lexical mais simples, porém graficamente interessante, na medida em que possibilita a rápida identificação das palavras-chave de um *corpus*⁽²⁴⁾.

Neste tipo de análise, as palavras são posicionadas aleatoriamente de tal forma que as palavras mais frequentes aparecem maiores do que as outras, o que demonstra seu destaque no *corpus* de análise da pesquisa⁽⁴⁾.

Por fim, a AFC foi utilizada em 7 pesquisas analisadas (13,0%). A AFC representa-se num plano cartesiano das diferentes palavras associadas a cada uma das classes da CHD⁽²⁵⁾. As representações gráficas resultantes deste tipo de análise mostram, assim descritas, as proximidades, oposições e tendências dos segmentos do *corpus*⁽¹⁸⁾.

Quanto aos benefícios do uso do IRAMUTEQ, destacaram-se os diferentes tipos de análises possibilitadas por este *software*, com apoio em elementos gráficos e em análises multivariadas; a otimização da organização de dados; o resgate do contexto no processo de análise; o maior rigor metodológico do processo de análise; e a gratuidade do *software*.

Pesquisadores apontam que, por ser uma ferramenta agilizável, o IRAMUTEQ facilita a pesquisa e dá oportunidade a análises mais consistentes e confiáveis^(5,17,26). Além dessas vantagens, um tal *software* possibilita que se recuperem, no *corpus* original, os segmentos de texto associados a cada classe, momento em que se obtém o contexto das palavras estatisticamente significativas⁽²⁵⁾.

Dentre outros ganhos, enfatiza-se, ainda, que o IRAMUTEQ apresenta rigor estatístico e permite aos pesquisadores utilizarem diferentes recursos técnicos de análise lexical. Além dessa possibilidade, sua interface é simples e facilmente compreensível e, sobretudo, seu acesso é do tipo *open source*^(1,17,19).

Quanto às limitações no uso do *software*, de forma preocupante, destaca-se que 96,3% das pesquisas não citaram as dificuldades vivenciadas, aspecto essencial a realçar e que pode se traduzir em aprendizado, tanto para o pesquisador quanto para o leitor dos trabalhos publicados.

Os dois trabalhos que comentaram tais aspectos de dificuldades apontam o tempo necessário ao pesquisador para estudar o *software*, antes de utilizá-lo adequadamente. Nesse panorama, realça-se, mais uma vez, a essencialidade da capacidade analítica e crítica do pesquisador qualitativo em qualquer processo de pesquisa, apoiado ou não no uso de *softwares*.

Importante é ressaltar, por fim, que foram verificadas algumas inconsistências no uso do IRAMUTEQ, tais como as que se seguem: mencioná-lo como técnica de análise de dados; utilizar pouco volume textual; e usar-se a CHD com menor quantidade de retenção do texto, inferior ao recomendado. Tais elementos precisam ser adequadamente considerados pelo pesquisador, a fim de se produzirem análises de dados confiáveis e fundamentadas.

Realça-se, mais uma vez, a necessidade de se compreender o IRAMUTEQ, assim como outros *softwares*, como mecanismo de apoio e que auxiliam no processo de tratamento de dados, sem substituírem o papel central do pesquisador.

Além desses critérios, enfatiza-se a essencialidade de estudo do pesquisador, que deve entender em profundidade o *software* que será utilizado e, sobretudo, alcançar competência analítica e crítica para efetuar a análise de dados.

Enquanto limitações desta pesquisa, destacam-se as dificuldades encontradas no manuseio do banco de dados: instabilidade; limitação temporal de disponibilização dos trabalhos (2013 a 2016); e limitação de acesso aos primeiros 10.000 trabalhos disponíveis. Desse modo, é possível que a amostra analisada nesta pesquisa esteja subestimada e, portanto, deve ser analisada sob esta perspectiva.

Diante de tais aspectos, sugere-se a replicação desta pesquisa a fim de integrar todos os estudos, independente de dimensão temporal. Além disso, destaca-se a importância de realizar tal pesquisa englobando dissertações e teses de outros países, a fim de se efetuar uma comparação do uso do IRAMUTEQ nas pesquisas qualitativas da área da saúde, nos diferentes continentes.

Espera-se contribuir para a reflexão dos pesquisadores qualitativos quanto ao uso consciente e fundamentado de *softwares*, bem como incentivar o ensino do uso de *softwares* qualitativos nos programas de pós-graduação.

CONCLUSÃO

Concluiu-se que o domínio quantitativo de investigações que utilizaram o IRAMUTEQ foi crescente ao longo dos anos e, portanto, tal *software* tem sido incorporado nas pesquisas qualitativas da área da saúde no Brasil, produzidas no âmbito dos programas de pós-graduação brasileiros.

REFERÊNCIAS

1. Camargo BV, Justo AM. IRAMUTEQ: Um software gratuito para análise de dados textuais. *Temas Psicol.* 2013;21:513-8.
2. Cope DG. Computer-Assisted Qualitative Data Analysis Software. *Oncol Nurs Forum.* 2014;41(3):322-3.
3. Paula MC, Viali L, Guimarães GTD. Uso do software Sphinx durante a realização de análise textual discursiva: outros percursos. *Revista Eletrônica DECT.* 2015;5(3):173-92.
4. 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 Rev Enferm.* 2016;20(3):e20160069
5. Lowen IMV, Peres AM, Crozeta K, Bernardino E, Beck CLC. Competências gerenciais dos enfermeiros na ampliação da Estratégia Saúde da Família. *Rev Esc Enferm USP.* 2015;49(6):967-73.
6. Camargo BV, Bousfield ABS, Giacomozzi AI, Koelzer LP. Representações sociais e adesão ao tratamento antirretroviral. *Liberabit.* 2014;20(2):229-38.
7. Almico T, Faro A. Enfrentamento de cuidadores de crianças com câncer em processo de quimioterapia. *Psicol Saúde Doenças.* 2014;15(3):723-37.
8. Silva JP, Bousfield ABS, Cardoso LH. A hipertensão arterial na mídia impressa: análise da revista *Veja*. *Psicol Saber Soc.* 2013;2(2):191-203.
9. Andrade Júnior EO, Andrade EO. Lexical analysis of the Code of Medical Ethics of the Federal Council of Medicine. *AMB Rev Assoc Med Bras.* 2016;62(2):123-30.
10. JBI. The Joanna Briggs Institute. Reviewers' Manual 2015: Methodology for JBI Scoping Reviews. South Australia: The Joanna Briggs Institute; 2015.
11. Santos V, Salvador P, Gomes A, Rodrigues C, Tavares F, Alves K, Bezerril M. IRAMUTEQ nas pesquisas qualitativas brasileiras da área da saúde: scoping review. In: 6º Congresso Ibero-Americano em Investigação Qualitativa; 2017 Jul 12-14; Salamanca, Espanha; 2017.
12. Geocapes. Sistema de Informações Georreferenciadas CAPES: Distribuição de programas de pós-graduação no Brasil [online]. 2015 [acesso em 2017 Nov 15]. Disponível em: <http://geocapes.capes.gov.br/geocapes2/#>.
13. Leitch J, Oktay J, Meehan B. A dual instructional model for computer-assisted qualitative data analysis software integrating faculty member and specialized instructor: Implementation, reflections, and recommendations. *Qualitative Social Work.* 2016;15(3):392-406.
14. Prado C, Casteli CPM, Lopes TO, Kobayashi RM, Peres HHC, Leite MMJ. Espaço virtual de um grupo de pesquisa: o olhar dos tutores. *Rev Esc Enferm USP.* 2012;46(1):246-51.
15. Erdmann AL, Pagliuca LMF. O conhecimento em enfermagem: da Representação de Área ao Comitê assessor de enfermagem no CNPq. *Rev Bras Enferm.* 2013;66(esp):51-9.
16. Camargo BV, Justo AM. Tutorial para uso do software IRAMUTEQ [online]. 2016. Disponível em: http://www.iramuteq.org/documentation/fichiers/Tutorial%20IRaMuTeQ%20em%20portugues_17.03.2016.pdf.

17. Plumecocq G. The second generation of ecological economics: How far has the apple fallen from the tree? *Ecol Econ*. 2014;107:457-68.
18. Queiroz Filho AP. As definições de assentamentos precários e favelas e suas implicações nos dados populacionais: abordagem da análise de conteúdo. *Rev Bras Gest Urbana*. 2015;7(3):340-53.
19. Smallman M. Public Understanding of Science in turbulent times III: Deficit to dialogue, champions to critics. *Public Underst Sci*. 2016;25(2):186-97.
20. Terama E, Smallman M, Lock SL, Johnson C, Austwick MZ. Beyond Academia ± Interrogating Research Impact in the Research Excellence Framework. *PLoS ONE*. 2016;11(12):e0168533.
21. Lahlou S. Text mining methods: An answer to Chartier and Meunier. *Papers on Social Representations*. 2012;20(38):1-7.
22. Bengough T, Bovet E, Bécherraz C, Schlegel S, Burnand B, Pidoux V. Swiss family physicians' perceptions and attitudes towards knowledge translation practices. *BMC Family Practice*. 2015;16:177.
23. Marchand P, Ratinaud P. L'analyse de similitude appliquée aux corpus textuelles: les primaires socialistes pour l'élection présidentielle française. In: 11^o Journées Internationales d'Analyse Statistique des Données Textuelles; 2012 Jun 13-15; Liège, Belgique; 2012.
24. Mendes FRP, Zangao 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.
25. Trigueiro DRSG, Almeida SA, Monroe AA, Costa GPO, Bezerra VP, Nogueira JA. AIDS and jail: social representations of women in freedom deprivation situations. *Rev Esc Enferm USP*. 2016;50(4):554-61.
26. Vaucher C, Bovet E, Bengough T, Pidoux V, Grossen M, Panese F, et al. Meeting physicians' needs: a bottom-up approach for improving the implementation of medical knowledge into practice. *Health Res Policy Syst*. 2016;14:49.

Endereço do primeiro autor:

Pétala Tuani Candido de Oliveira Salvador
 Universidade Federal do Rio Grande do Norte - UFRN
 Campus Universitário - Escola de Saúde
 Bairro: Lagoa Nova
 CEP: 59064-741- Natal - RN - Brasil
 E-mail: petalatuani@hotmail.com

Endereço para correspondência:

Viviane Euzébia Pereira Santos
 Universidade Federal do Rio Grande do Norte - UFRN
 Campus Universitário - Departamento de Enfermagem
 Bairro: Lagoa Nova
 CEP: 59064-741- Natal - RN - Brasil
 E-mail: vivianeepsantos@gmail.com