

PANORAMA DA PRODUÇÃO  
CIENTÍFICA SOBRE  
MULHERES E TRABALHO:  
UM RETRATO  
BIBLIOMÉTRICO A PARTIR  
DAS METODOLOGIAS DE  
ANÁLISE DE CONTEÚDO

OVERVIEW OF SCIENTIFIC PRODUCTION ON WOMEN AND WORK: A  
BIBLIOMETRIC PORTRAIT BASED ON CONTENT ANALYSIS  
METHODOLOGIES

Ciências Humanas, Ciências Sociais Aplicadas, Linguística & Letras e

Artes

• 06/06/2026

REGISTRO DOI: [10.70773/revistatopicos/780635194](https://doi.org/10.70773/revistatopicos/780635194)

---

Francisco Alves de Freitas Neto<sup>1</sup>

Carlos Henrique Medeiro de Souza<sup>2</sup>

Josilaine Oliveira de Souza<sup>3</sup>

Maycon Gual Ribeiro<sup>4</sup>

Fabio Machado de Oliveira<sup>5</sup>

Raquel França Freitas<sup>6</sup>

---

## RESUMO

A relação entre o universo feminino e o trabalho é um campo de estudo consolidado. Contudo, é importante compreender como este tema é abordado metodologicamente. Este artigo mapeia a produção científica que utiliza especificamente a "Análise de Conteúdo" para investigar essa relação. O objetivo é traçar o panorama bibliométrico do campo, identificando os periódicos centrais, os autores mais produtivos e os núcleos temáticos. Trata-se de um estudo bibliométrico exploratório, de abordagem mista (quantitativa e qualitativa). A coleta de dados foi realizada na base Web of Science (WoS) com a string ("Content Analysis" AND "women" AND "work") e termos análogos, resultando em 1.669 publicações. Após filtros de relevância (1.000 artigos), temporal (pós-2002) e higienização (remoção de duplicatas e registros nulos), o corpus foi estabilizado em 867 artigos. A Lei de Bradford foi então aplicada para definir um "núcleo de elite" de 32 periódicos, que continha 218 publicações. Este corpus de elite serviu de base para as análises subsequentes. Os resultados demonstram que o campo é dominado por periódicos das áreas de Enfermagem (*Nursing*) e Saúde Pública. A análise de Lotka (autores) revelou uma rede de colaboração fragmentada, com 97,6% dos autores possuindo apenas uma publicação. A análise de Zipf (frequência de palavras), aplicada em Títulos, Resumos e Palavras-chave, confirmou que o campo é basicamente focado em estudos qualitativos (termos *qualitative, interviews*), sobre as experiências de saúde da mulher (termos *health, care, midwives*). Conclui-se que a Análise de Conteúdo é uma ferramenta metodológica chave para estudos qualitativos sobre trabalho e saúde feminina.

**Palavras-chave:** Bibliometria; Análise de Conteúdo; Mulheres e Trabalho; Lei de Bradford; Web of.

## **ABSTRACT**

The relationship between women and work is a consolidated field of study. However, it is important to understand how this theme is approached methodologically. This article maps the scientific production that specifically uses "Content Analysis" to investigate this relationship. The objective is to outline the bibliometric panorama of the field, identifying the central journals, the most productive authors, and the thematic cores. This is an exploratory bibliometric study using a mixed-methods approach (quantitative and qualitative). Data collection was conducted in the Web of Science (WoS) database using the string ("Content Analysis" AND "women" AND "work") and analogous terms, resulting in 1,669 publications. After applying filters for relevance (1,000 articles), time frame (post-2002), and data sanitization (removal of duplicates and null records), the corpus was stabilized at 867 articles. Bradford's Law was then applied to define an "elite core" of 32 journals, which contained 218 publications. This elite corpus served as the basis for subsequent analyses. The results demonstrate that the field is dominated by journals in the areas of Nursing and Public Health. Lotka's analysis (authors) revealed a fragmented collaboration network, with 97.6% of the authors having only one publication. Zipf's analysis (word frequency), applied to Titles, Abstracts, and Keywords, confirmed that the field is primarily focused on qualitative studies (terms: qualitative, interviews) regarding women's health experiences (terms: health, care, midwives). It is concluded that Content Analysis is a key methodological tool for qualitative studies on women's work and health.

**Keywords:** Bibliometrics; Content Analysis; Women and Work; Bradford's Law; Web of Science.

## **1. INTRODUÇÃO**

A relação entre o universo feminino e o trabalho é um campo de estudo que, embora amplo e consolidado, ainda necessita de pesquisas, pois essa relação ganha contornos de acentuada invisibilidade e desvalorização em setores tradicionalmente dominados por homens, como o da pesca artesanal no litoral fluminense (Timóteo, 2019). Neste estudo, Timóteo observa que, embora as mulheres sejam importantes em toda a cadeia produtiva, desde o beneficiamento do pescado até a gestão dos recursos familiares, seu papel é frequentemente subestimado ou classificado como "ajuda".

O estudo de Timóteo (2019) é um dos principais resultados do Projeto de Educação Ambiental (PEA) Pescarte. Este projeto surgiu como uma medida de mitigação socioambiental obrigatória, exigida pelo licenciamento ambiental federal conduzido pelo IBAMA, decorrente das atividades de exploração de petróleo e gás da Petrobras na Bacia de Campos. O PEA Pescarte concentra sua atuação nas comunidades de pescadores artesanais de sete municípios do litoral fluminense (Arraial do Cabo, Cabo Frio, Macaé, Quissamã, Campos dos Goytacazes, São João da Barra e São Francisco de Itabapoana), escolhidos por sua vulnerabilidade aos impactos da indústria petrolífera. A importância do projeto materializa-se na geração de um banco de dados inédito sobre as condições de vida destes trabalhadores, que, apesar de sua relevância para a produção de alimentos, vivem em bairros pobres, com precariedades latentes que beiram ao risco social e enfrentam dificuldades crônicas de acesso a serviços urbanos e políticas públicas.

A invisibilidade feminina nessa relação com o trabalho, explorada em profundidade no Capítulo 6 ("Trabalho e Relações de Gênero na

Pesca Artesanal") do referido trabalho, serve como a principal motivação para esta investigação. A superação desse silenciamento histórico exige o uso de ferramentas metodológicas que permitam analisar discursos, práticas e representações sociais de forma rigorosa. Entre essas ferramentas, a Análise de Conteúdo (AC), sistematizada por Bardin, destaca-se como um método chave para a pesquisa qualitativa neste campo (Bardin, 2016).

A presente pesquisa insere-se em um projeto de investigação mais amplo que visa, futuramente, aplicar a AC em um corpus de entrevistas com pescadoras artesanais, focando em suas percepções sobre o trabalho. No entanto, para fundamentar e contextualizar metodologicamente essa análise específica, torna-se fundamental primeiro responder a uma pergunta mais abrangente: Como a comunidade científica global tem utilizado a Análise de Conteúdo para investigar o tema "mulher e trabalho"?

Partindo deste questionamento, o presente artigo tem como objetivo central traçar o panorama bibliométrico da produção científica que articula, simultaneamente, os temas "Análise de Conteúdo", "mulher" e "trabalho". Utilizando os dados da Web of Science e as leis clássicas da bibliometria, buscou-se identificar os periódicos centrais (Lei de Bradford), os autores mais produtivos e as redes de colaboração (Lei de Lotka), e os núcleos temáticos emergentes ou consolidados (Lei de Zipf) (Grácio et al., 2020). Este mapeamento fornecerá o estado da arte e o alicerce metodológico que guiará futuras investigações qualitativas sobre o tema.

## **2. FUNDAMENTAÇÃO TEÓRICA**

Para fundamentar teoricamente esta pesquisa, dada sua natureza interdisciplinar, esta seção foi dividida em duas partes, cada qual abordando diferentes áreas do conhecimento. A primeira subseção explora os conceitos fundamentais da Análise de Conteúdo, com foco na análise de corpus discursivos e entrevistas. Já a segunda subseção discute os princípios da bibliometria, incluindo a definição e aplicação das Leis de Bradford, Lotka e Zipf, previamente mencionadas na Introdução. Por fim, busca-se estabelecer a conexão entre esses tópicos, demonstrando como a abordagem bibliométrica pode auxiliar na análise dos impactos das relações trabalhistas das mulheres, bem como na forma como a academia tem investigado essa temática.

## **2.1. Análise de Conteúdo**

Bardin (2016) define a Análise do Conteúdo (AC) como um conjunto de técnicas de análise das comunicações que utiliza procedimentos sistemáticos e objetivos de descrição do conteúdo das mensagens. O objetivo não é apenas ler o texto, mas ultrapassar a incerteza e enriquecer a leitura, extraíndo inferências sobre o emissor da mensagem ou seu contexto.

A AC, conforme sistematizada por Bardin, não é um ato único, mas um processo metodológico organizado em três fases cronológicas principais:

- Pré-análise: É a fase de organização e sistematização das ideias iniciais. Nela, realiza-se a "leitura flutuante" do *corpus* (no seu caso, os futuros resumos e entrevistas) para uma primeira impressão. Esta fase envolve a escolha dos documentos, a

formulação de hipóteses e a definição dos indicadores que guiarão a análise.

- Exploração do Material (Codificação): Esta é a fase central da análise. Consiste em transformar os dados brutos do texto em unidades de registro (palavras, temas, frases) através de um processo de codificação. As unidades de registro são então agrupadas por semelhança, dando origem às categorias de análise. É o processo de categorização que permite classificar e sintetizar os elementos do *corpus*.
- Tratamento dos Resultados, Inferência e Interpretação: Na fase final, os dados codificados e categorizados são tratados para que se tornem significativos. O analista busca padrões, realiza inferências e propõe interpretações sobre o material, cruzando os resultados encontrados com o referencial teórico ou as hipóteses iniciais. É o momento em que o "não-dito" emerge do "dito".

O presente estudo bibliométrico foca, portanto, em como a academia tem aplicado esta metodologia, mapeando os temas (frequência de palavras) que emergem destas análises.

## **2.2. Conceitos de Bibliometria**

Conforme destacado por Grácio et al. (2020), a bibliometria estabeleceu-se como um campo de estudo essencial para compreender a natureza e avaliar a comunicação científica. Este autor também aponta que a bibliometria, através de métodos quantitativos, pode apoiar processos decisórios e qualificar a pesquisa, criando um material de consulta bibliográfica de alta qualidade. Assim, a utilização da bibliometria na elaboração de uma

base de dados de publicações sobre o tema deste trabalho oferece um método objetivo para identificar deficiências informacionais, como a obsolescência de ideias, autores e termos indexados. Outro ponto a destacar neste processo é que a bibliometria permite identificar áreas emergentes ou em declínio.

A bibliometria se alicerça em três leis fundamentais: a Lei de Bradford, que discute a produtividade dos periódicos; a Lei de Lotka, que examina a produtividade dos autores; e a Lei de Zipf, que investiga a frequência de ocorrência de palavras (Ferreira, 2010). Considerando que este estudo busca estabelecer uma base de referências para a consulta de publicações acadêmicas, é importante identificar as fontes mais influentes de artigos científicos, como periódicos e anais de conferências. Dentro desse contexto, a Lei de Bradford assume um papel fundamental na orientação da pesquisa bibliométrica.

A Lei de Bradford, também conhecida como Lei de Dispersão, é uma ferramenta que quantifica a produtividade dos periódicos e identifica tanto o núcleo quanto as áreas de dispersão de um tema específico em um conjunto de publicações. De acordo com essa lei, os primeiros trabalhos sobre um novo tema costumam ser submetidos a um número limitado de revistas especializadas. Quando esses periódicos aceitam os estudos, eles atraem um volume crescente de publicações sobre o tema ao longo do tempo. Simultaneamente, outras revistas começam a publicar seus primeiros artigos sobre o assunto, criando novas áreas de dispersão. Se o campo continuar a progredir, forma-se um núcleo de periódicos de alta produtividade, que concentram a maior parte das pesquisas sobre o tema (Lousada et al., 2012).

Além de identificar fontes confiáveis para as referências de pesquisa, é igualmente fundamental compreender a relevância dos autores nas produções acadêmicas, considerando fatores como o número de artigos publicados sobre o tema, produtividade, citações, entre outros. Nesse contexto, a Lei de Lotka torna-se essencial para essa investigação. A seguir, é apresentada a descrição dessa lei:

A Lei de Lotka, também conhecida como Lei do Quadrado Inverso, descreve a distribuição da produtividade dos pesquisadores, utilizando um modelo de distribuição tamanho-frequência dos pesquisadores em um conjunto de artigos. Em termos práticos, a produtividade de um autor, medida pelo número de artigos publicados, reflete sua contribuição para o avanço do conhecimento científico. Estudada por Lotka (1926), essa lei estabelece que o número de pesquisadores que fazem  $n$  contribuições em um determinado campo do conhecimento científico é aproximadamente  $1/n^2$  do número de pesquisadores que fazem apenas uma contribuição. Além disso, a proporção de pesquisadores que contribuem com apenas um artigo gira em torno de 60% (Melo Ribeiro, 2017).

Este estudo bibliométrico tem como um dos objetivos específicos guiar os pesquisadores sobre as palavras mais comuns em publicações; por isso, a Lei de Zipf será discutida neste trabalho. A Lei de Zipf, introduzida em 1949 e também conhecida como a "lei do mínimo esforço", examina a frequência com que as palavras aparecem em um texto. Para esse fim, é gerada uma lista das palavras presentes no texto, organizadas em ordem decrescente de acordo com sua frequência de ocorrência. A posição de cada palavra nesta lista é chamada de ordem de série ou *rank*. Zipf expandiu sua proposta original em duas direções: a primeira, aplicada

exclusivamente a palavras de alta frequência; e a segunda, ajustada por Andrew Boot, sugere que múltiplas palavras com baixa frequência de ocorrência compartilham a mesma posição em determinado texto (Grácio et al., 2020).

### **3. METODOLOGIA E DESENVOLVIMENTO**

A presente pesquisa configura-se como um estudo exploratório de abordagem mista: quantitativa e qualitativa (Gil, 2002). A dimensão quantitativa baseia-se nos princípios da bibliometria, utilizando indicadores de produção para analisar um grande volume de dados. A dimensão qualitativa se manifesta na interpretação dos agrupamentos temáticos (*clusters*) identificados. Este mapeamento, portanto, serve como uma etapa fundamental para uma investigação futura de natureza puramente qualitativa: o objetivo é utilizar os temas e lacunas aqui identificados para informar a construção de categorias para uma análise de conteúdo, nos moldes de Bardin, de uma entrevista com uma pescadora artesanal sobre sua relação com o trabalho (Bardin, 2016).

Para a condução deste estudo, os procedimentos metodológicos foram divididos em etapas sequenciais, conforme ilustrado na Figura 1. O ponto de partida foi a Consulta na Web of Science (WoS), onde a aplicação da *string* de busca definida resultou no *corpus* base de 1669 publicações. Em seguida, realizou-se uma Análise Preliminar descritiva para verificar o volume total e a evolução temporal da produção científica. Na fase subsequente, o *corpus* foi analisado sob a ótica das leis bibliométricas clássicas, fundamentadas na seção anterior: aplicou-se a Lei de Bradford para identificar os periódicos mais relevantes; a Lei de Lotka para mensurar a produtividade e as relações entre os autores; e a Lei de Zipf para investigar a frequência

e a coocorrência dos principais temas e palavras-chave (Grácio et al., 2020).

**Figura 1:** Metodologia de Pesquisa Aplicada



Fonte: Autoria Própria

### 3.1. Consulta Ao Web Of Science

A primeira etapa operacional da metodologia (conforme Figura 1) foi a coleta de dados. A base de dados selecionada foi a Web of Science (WoS) Core Collection, reconhecida internacionalmente pela sua abrangência e indexação de periódicos de alto impacto.

A busca foi realizada em 06/10/2025, sem restrição de período, no campo "Tópico"(TS), que engloba pesquisas por título, resumo e palavras-chave. A *string* exata utilizada foi:

TS = (("Análise de Conteúdo" AND ("mulher" OR "mulheres")) AND "trabalho")

OR

("Content Analysis" AND ("woman" OR "women") AND "work"))

A aplicação destes critérios retornou um total de 1.669 publicações. No entanto, devido às limitações de exportação da plataforma WoS, que permite o *download* de um máximo de 1.000 registros por vez, adotou-se um critério de corte. Optou-se por exportar o conjunto dos 1.000 artigos mais relevantes (conforme a ordenação padrão da plataforma), considerando esta amostra suficiente para os objetivos exploratórios deste estudo.

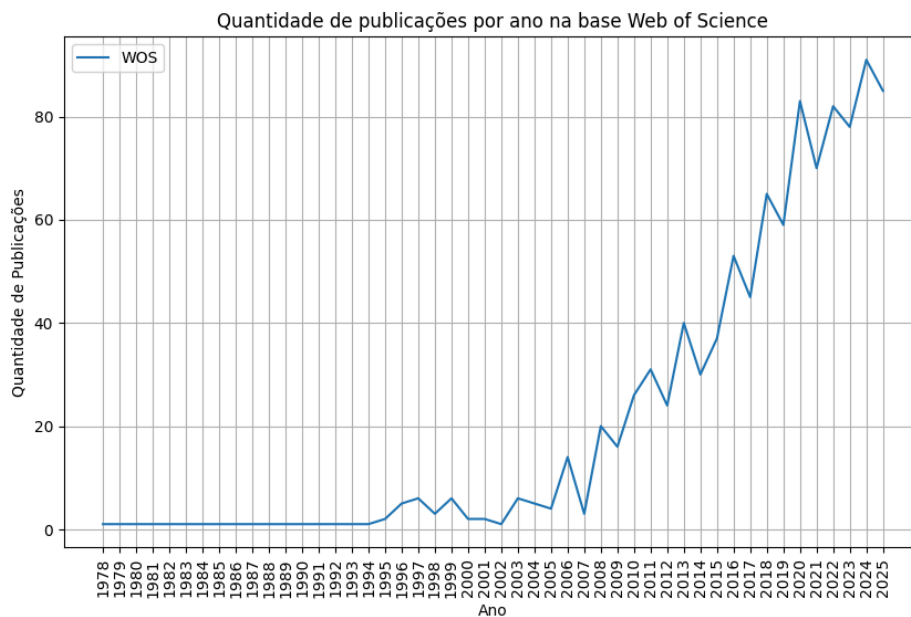
O *corpus* final de 1.000 publicações foi exportado no formato "Excel", servindo como base para a etapa seguinte de tratamento e análise.

### **3.2. Tratamento e Análise Preliminar dos Dados**

O arquivo de 1.000 publicações foi processado utilizando a linguagem de programação Python, especificamente com o auxílio da biblioteca Pandas. O primeiro passo foi o reconhecimento estrutural dos metadados, identificando as colunas relevantes para cada lei bibliométrica, como 'Publication Year', 'Source Title', 'Author Full Names' e 'Author Keywords'.

Em seguida, procedeu-se à Análise Preliminar (conforme Figura 1), iniciando pela distribuição da produção científica ao longo do tempo. Um gráfico da série temporal completa (1978-2025) foi gerado, como mostra a Figura 2.

**Figura 2:** Quantidade de publicações por ano (1978-2025)



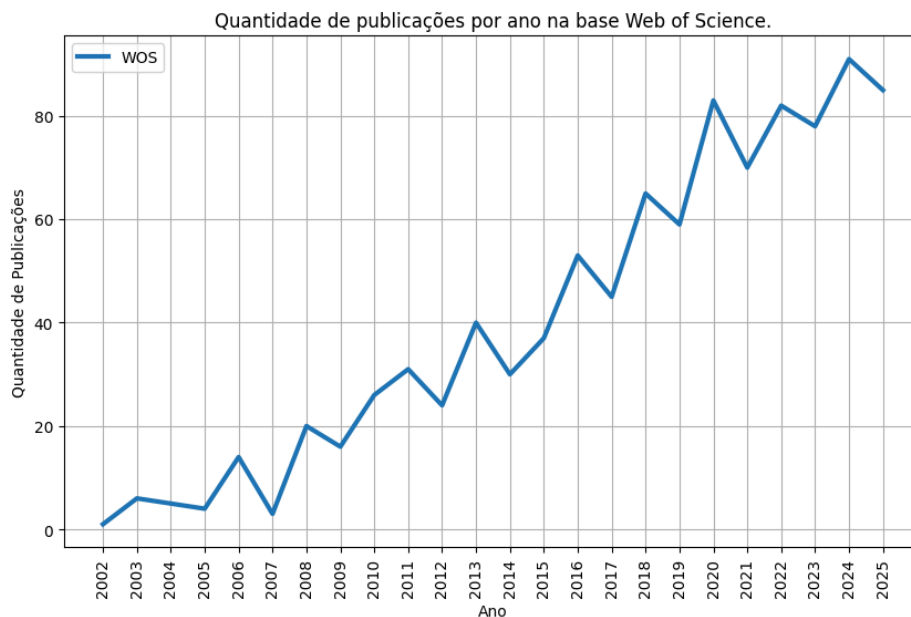
**Fonte:** Autoria Própria

A análise da Figura 2 tornou evidente que a produção científica sobre o tema foi praticamente inexistente ou inexpressiva até o ano de 2001. A vasta maioria das publicações concentra-se a partir de 2002.

Diante desta constatação, e com o objetivo de focar a análise no período de maior atividade e relevância científica, tomou-se a decisão metodológica de aplicar um filtro temporal. Foram descartados todos os artigos publicados antes de 2002. Esta filtragem refinou o *corpus*, resultando em uma amostra final de 968 publicações.

A Figura 3 ilustra a nova distribuição temporal do *corpus* filtrado, que servirá de base para todas as análises subsequentes.

**Figura 3:** Quantidade de publicações por ano no corpus filtrado (2002-2025)



**Fonte:** Autoria Própria.

A nova série temporal (Figura 3) demonstra claramente uma tendência de crescimento no interesse pelo tema, com uma aceleração notável a partir de 2008 e picos de produção nos últimos anos, validando a relevância contemporânea da pesquisa.

Após o filtro temporal, o *corpus* de 968 publicações passou por uma etapa final de higienização e tratamento de dados, visando garantir a integridade e a qualidade da amostra. Utilizando o Pandas, foram aplicados os seguintes critérios de exclusão:

- Remoção de registros sem 'Article Title' (Título do Artigo);
- Remoção de registros sem 'Author Keywords' (Palavras-chave do Autor), metadado essencial para a análise de temas (Lei de Zipf);
- Remoção de registros sem 'Abstract' (Resumo), campo fundamental para a futura análise qualitativa (Bardin).

Por fim, para assegurar a unicidade dos dados, os títulos dos artigos foram padronizados para letras minúsculas e as publicações

duplicadas foram removidas.

Este processo de tratamento resultou no *corpus* final de 867 publicações, que servirá de base para a aplicação das leis bibliométricas nas seções seguintes.

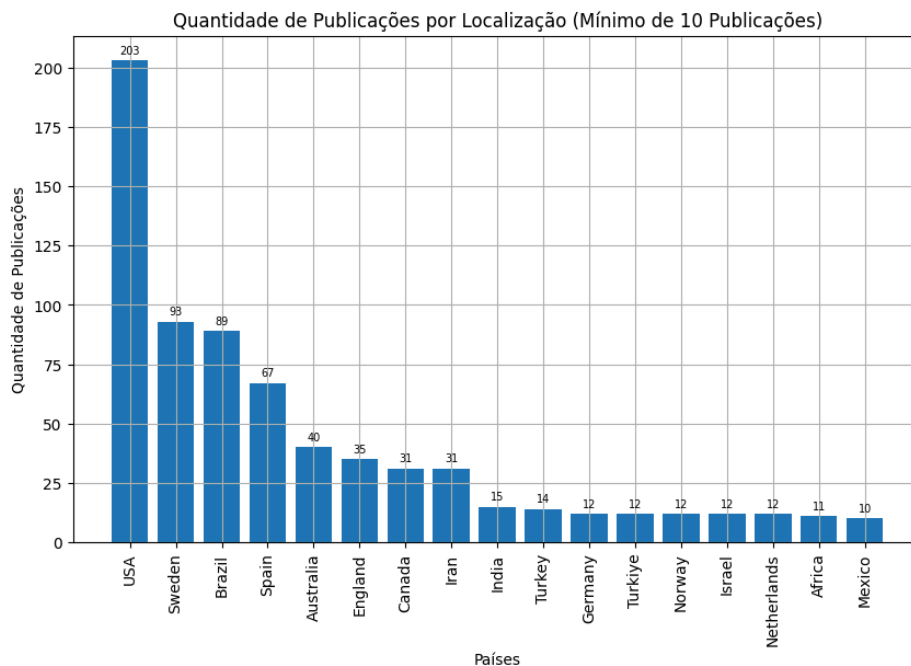
### **3.3. Análise de Perfil da Amostra: Distribuição Geográfica**

Após a definição do *corpus* final de 867 publicações, a primeira análise de resultados investigou o perfil da amostra, começando pela distribuição geográfica da produção científica.

Para esta análise, foi realizada uma tentativa de extração do país de afiliação principal a partir do metadado 'Addresses'. Foi aplicada uma heurística de processamento de texto (extração da última palavra da *string* de endereço) para identificar o país. É importante ressaltar que este método apresenta limitações, dada a falta de padronização do campo 'Addresses'. Como resultado, podem ocorrer agregações imprecisas (ex: "South Africa" sendo contabilizada como "Africa") ou inconsistências de idioma (ex: "Turkey" e "Turkiye" como entradas distintas), como pode ser observado na análise.

Apesar dessas limitações, a análise oferece um panorama da liderança na pesquisa. A Figura 4 ilustra a contagem de publicações por país, filtrada para exibir apenas locais com 10 ou mais contribuições.

**Figura 4:** Quantidade de Publicações por Localização (Mínimo de 10 Publicações)



**Fonte:** Autoria Própria

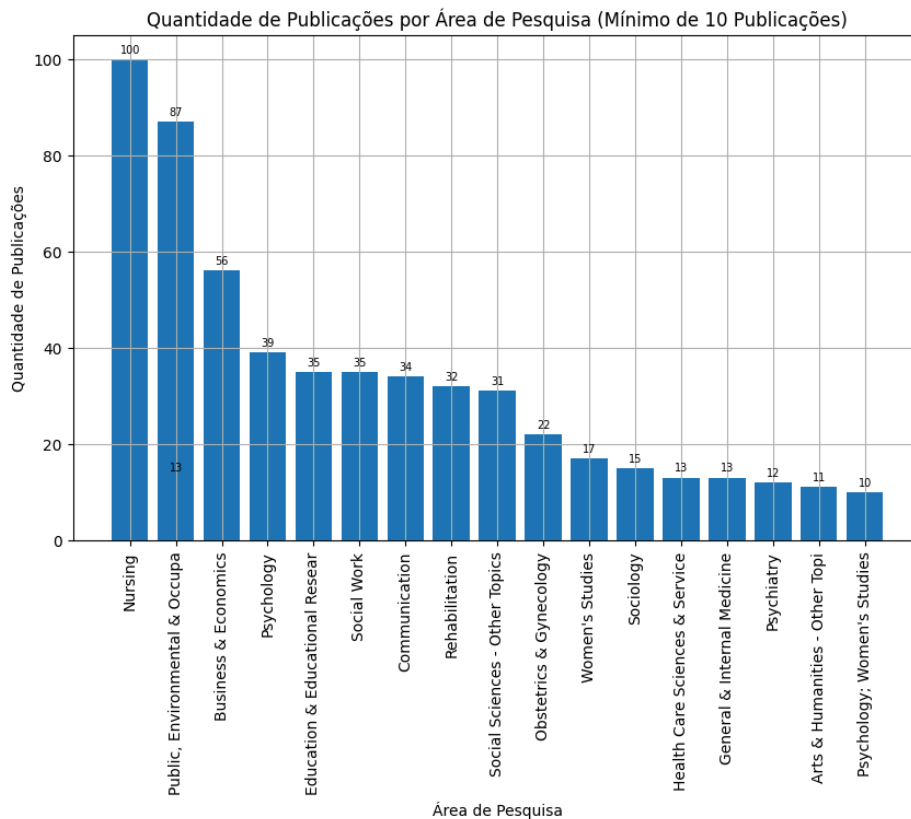
Observa-se uma liderança expressiva dos Estados Unidos (USA), com 203 publicações, o que representa aproximadamente 23,4% do corpus total. Na sequência, destacam-se a Suécia (Sweden), com 98 publicações, e o Brasil, com 89 publicações. Espanha (67), Austrália (40) e Inglaterra (35) também compõem o grupo de países com maior produção na área.

### 3.4. Distribuição por Áreas de Pesquisa

Complementando a análise de perfil da amostra de 867 publicações, investigou-se a distribuição do corpus pelas áreas de pesquisa. Para isso, foi utilizado o metadado 'Research Areas' da WoS, que permite identificar as disciplinas científicas que mais publicam sobre o tema.

A Figura 5 apresenta a distribuição dos artigos, filtrada para exibir apenas as áreas com 10 ou mais publicações.

**Figura 5:** Quantidade de Publicações por Área de Pesquisa (Mínimo de 10 Publicações)



**Fonte:** Autoria Própria

A análise revela que o tema é marcadamente multidisciplinar, mas com uma concentração notável em campos ligados às ciências da saúde e ciências sociais aplicadas.

Os resultados da Figura 5 mostram que a área de *Nursing* (Enfermagem) lidera a produção, com 100 publicações. Em seguida, destacam-se *Public, Environmental & Occupational Health* (Saúde Pública, Ambiental e Ocupacional), com 87 publicações, e *Business & Economics* (Negócios e Economia), com 56 publicações.

Campos tradicionalmente associados ao tema, como *Psychology* (Psicologia) (39 publicações), *Education & Educational Research* (Educação) (35), *Social Work* (Serviço Social) (35) e *Women's Studies* (Estudos da Mulher) (17), também figuram com relevância.

Esse panorama sugere que, metodologicamente, a Análise de Conteúdo está sendo amplamente empregada para investigar a

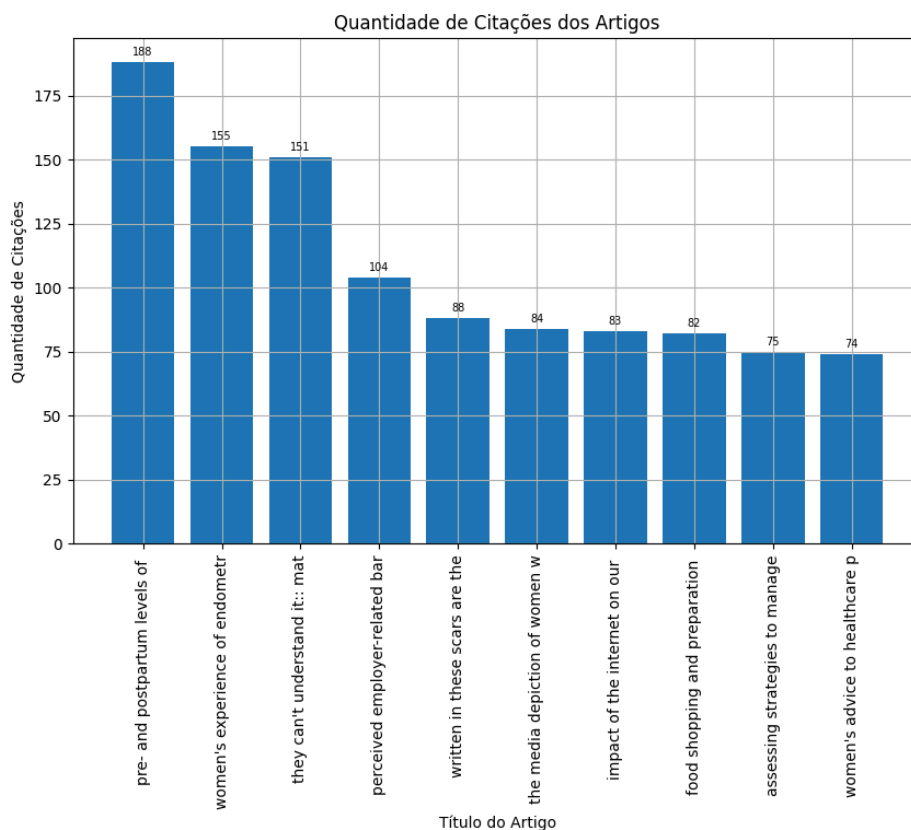
relação "mulheres e trabalho" sob a ótica da saúde ocupacional, da gestão de carreira e das dinâmicas econômicas, e não apenas por uma perspectiva puramente sociológica ou de gênero.

### 3.5. Análise de Citação: Obras de Maior Impacto

Além de analisar a produtividade e a distribuição da pesquisa, buscou-se identificar as obras de maior impacto dentro do corpus final de 867 artigos. O impacto de um artigo é medido pelo número de vezes que ele é citado por publicações subsequentes, indicando sua relevância para o campo.

Para esta análise, os artigos foram ordenados utilizando o metadado 'Times Cited, WoS Core' (Vezez Citado, Coleção Principal WoS). A Figura 6 apresenta os 10 artigos mais influentes do *corpus*.

**Figura 6:** Os 10 Artigos Mais Citados do Corpus (2002-2025)



Fonte: Autoria Própria

Observa-se um achado de grande relevância: os artigos de maior impacto no tema não estão focados em gestão, carreira ou economia (como a análise de áreas de pesquisa poderia sugerir), mas sim fortemente concentrados na área da saúde.

Títulos como "*preand postpartum levels of...*" (Níveis pré e pós-parto de...), "*women's experience of endometr...*" (Experiência de mulheres com endometriose...) e "*they can't understand it:: mat...*" (Eles não conseguem entender:: [saúde] materna...) dominam o topo da lista.

Este resultado corrobora a análise de distribuição por áreas (Figura 5), que apontou "*Nursing*" (Enfermagem) e "*Public, Environmental & Occupational Health*" (Saúde Pública) como os campos mais produtivos. Fica evidente, portanto, que a Análise de Conteúdo é uma metodologia crucial para entender as experiências de saúde das mulheres (especialmente materna e ginecológica) e sua intersecção com o trabalho ou o sistema de saúde.

### **3.6. Aplicação da Lei de Bradford: Produtividade dos Periódicos**

Conforme fundamentado na seção teórica, a Lei de Bradford foi aplicada para identificar os periódicos mais relevantes e analisar a dispersão da literatura sobre o tema. Para isso, foi realizada a contagem de frequência do metadado 'Source Title' (Título da Fonte) sobre o *corpus* final de 867 artigos.

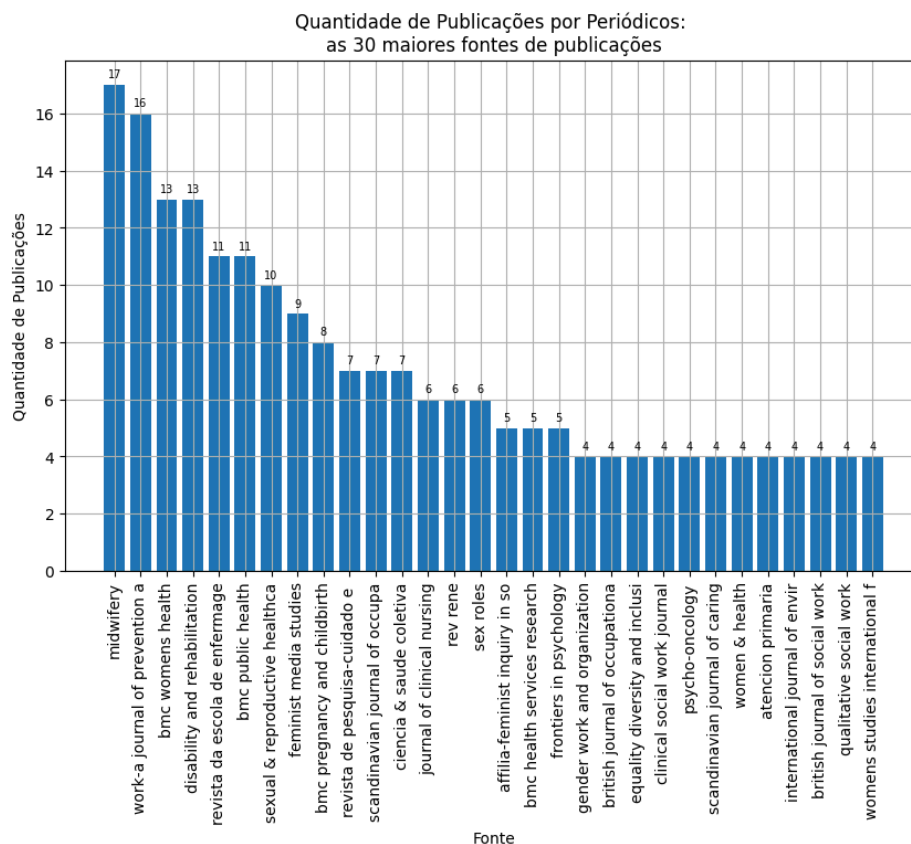
Os resultados revelaram que as 867 publicações estão distribuídas por 556 periódicos distintos. Este número demonstra uma alta pulverização da literatura, um padrão clássico descrito pela Lei de Bradford, onde um pequeno "núcleo" de periódicos concentra uma parte significativa das publicações, enquanto um grande número de

periódicos publica apenas um ou dois artigos sobre o tema (Lousada et al., 2012).

A Figura 7 apresenta uma visualização inicial deste fenômeno, listando os 30 periódicos mais produtivos do *corpus*.

A análise dos periódicos de maior produtividade corrobora, mais uma vez, os achados das seções anteriores. O núcleo de periódicos é dominado por publicações da área da Saúde, com destaque para Midwifery (17 publicações), Work: A Journal of Prevention... (16), BMC Women's Health (13) e Revista da Escola de Enfermagem (13).

**Figura 7:** Quantidade de Publicações por Periódicos (Top 30 Fontes)



Fonte: Autoria Própria

Este gráfico ilustra os periódicos candidatos a compor o Núcleo de Bradford. A etapa seguinte consiste em calcular as zonas de produtividade para formalizar esta classificação.

### 3.7. Aplicação da Lei de Bradford: Zonas de Produtividade

Após a identificação dos 30 periódicos mais produtivos (Figura 7), foi realizado o cálculo formal das zonas de Bradford para o corpus de 867 artigos distribuídos em 556 periódicos.

O objetivo foi dividir o corpus em três zonas contendo um número similar de publicações (o "ideal" de  $k \approx 289$  artigos por zona). Devido à natureza indivisível dos dados (um periódico não pode ser quebrado), os cortes foram feitos nos pontos mais próximos da distribuição.

A Tabela 1 demonstra o agrupamento de artigos e periódicos que serviu de base para o cálculo.

**Tabela 1:** Cálculo Cumulativo de Periódicos e Publicações

Qtd.Pubs	Qtd.Fontes	Total Fontes	Total Pubs	Log Total Fontes
17	1	1	17	0.00
16	1	2	33	0.30
13	2	4	59	0.60
...	...	...	...	...
4	14	32	218	1.50
3	26	58	296	1.76
2	73	131	442	2.11
1	425	556	867	2.74

**Fonte:** Autoria Própria

A aplicação do método resultou na divisão apresentada na Tabela 2.

**Tabela 2:** Distribuição de Zonas de Bradford

Zona	Qtd. Periódicos	Qtd. Publicações	Composição
Zona 1 (Núcleo)	32	218	Periódicos com 4+ artigos
Zona 2 (Intermediária)	99	224	Periódicos com 2 ou 3 artigos
Zona 3 (Dispersão)	425	425	Periódicos com 1 artigo
<b>Total</b>	<b>556</b>	<b>867</b>	

**Fonte:** Autoria Própria

A Lei de Bradford estabelece que a proporção de periódicos entre as zonas deve seguir uma progressão geométrica ( $n : n \times b : n \times b^2$ ) (Guedes; Borschiver, 2005). Calculando o multiplicador de Bradford ( $b$ ) a partir dos nossos dados:  $b = (\text{Periódicos Zona 2} / \text{Periódicos Zona 1}) = 99/32 \approx 3.09$ . Verificando a progressão para a Zona 3:  $(\text{Periódicos Zona 2}) \times b = 99 \times 3.09 \approx 305.9$ .

O valor real da Zona 3 foi de 425 periódicos. Embora o valor de 306 não seja idêntico a 425, a proporção geral (32 : 99 : 425) demonstra claramente o princípio da lei: um núcleo muito pequeno de 32 periódicos é responsável por uma grande fatia do conhecimento, enquanto uma grande sequência de 425 periódicos (76% do total) publicou apenas um artigo sobre o tema.

Seguindo o cálculo das zonas, uma decisão metodológica foi tomada para refinar o escopo das análises subsequentes. Com o objetivo de focar a investigação nos periódicos de maior produtividade e relevância, o *corpus* foi filtrado para conter apenas as publicações pertencentes ao Núcleo (Zona 1). Conforme a Tabela 2, este núcleo é composto pelos 32 periódicos que publicaram quatro ou mais artigos sobre o tema.

Foi aplicado um filtro ao conjunto de dados para reter apenas os artigos publicados nesses 32 periódicos de elite. Este procedimento refinou o *corpus* de 867 publicações para um conjunto final de 218 publicações. Este novo *corpus* de 218 artigos, que representa o "núcleo de elite" da pesquisa, servirá como base para as análises

seguintes de produtividade de autores (Lei de Lotka) e de coocorrência de temas (Lei de Zipf).

### **3.8. Aplicação da Lei de Lotka: Produtividade dos Autores**

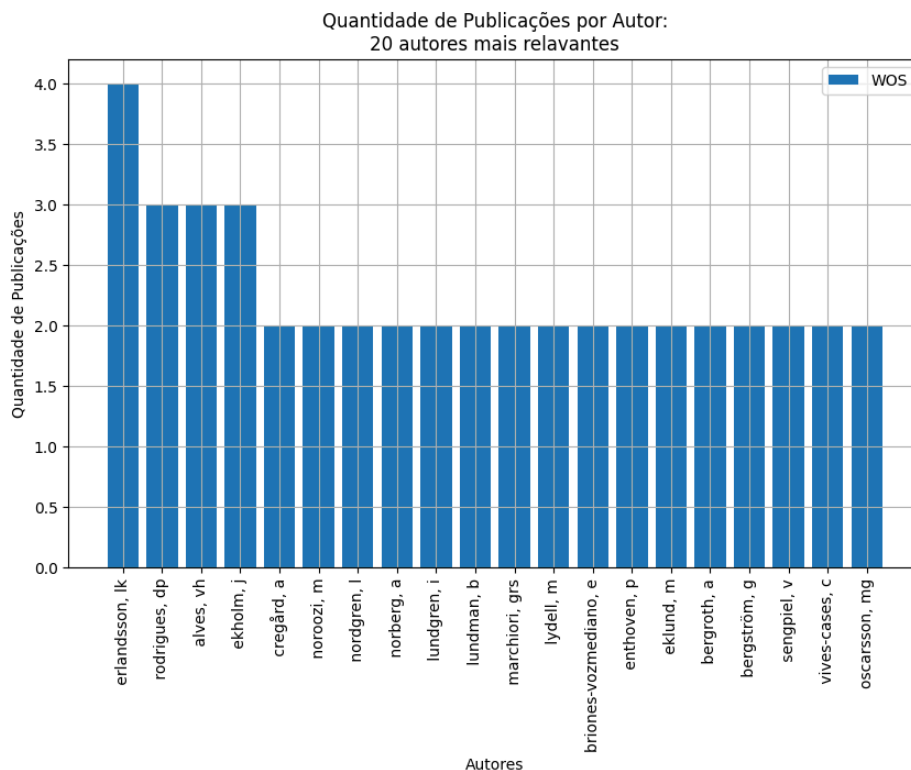
Dando sequência à análise do corpus de elite de 218 publicações, foi aplicada a Lei de Lotka para investigar a produtividade e o perfil dos autores. Conforme a metodologia proposta, o campo 'Authors' foi processado: as listas de autores de cada artigo foram desagregadas (separadas pelo caractere ';') e normalizadas (convertidas para minúsculas). Em seguida, foi realizada uma contagem de frequência para identificar o número de publicações de cada autor único.

A análise revelou que os 218 artigos foram produzidos por um total de 839 autores únicos. A Figura 8 apresenta o ranking dos 20 autores mais produtivos, que compõem o pequeno núcleo de pesquisadores com múltiplas publicações no tema.

Os resultados do gráfico ilustram um padrão clássico da Lei de Lotka, com uma alta pulverização da autoria. O autor mais produtivo do *corpus*, 'erlandsson, lk', possui 4 publicações, seguido por um pequeno grupo de três autores com 3 publicações cada ('rodrigues, dp', 'alves, vh', 'ekholm, j').

No entanto, o padrão observado é uma versão extrema da lei. Conforme citado na fundamentação teórica (Melo Ribeiro, 2017), a lei prevê que cerca de 60% dos autores publicam apenas um artigo. No presente corpus, o fenômeno é muito mais acentuado: dos 839 autores únicos, 819 publicaram apenas um artigo. Isso significa que 97,6% dos pesquisadores são autores transitórios, que contribuíram apenas uma vez para o núcleo de elite deste tema.

**Figura 8:** Quantidade de Publicações por Autor (Top 20)



**Fonte:** Autoria Própria

Essa discrepância acentuada (97,6% contra os 60% teóricos) é esperada e pode ser explicada por dois fatores:

- Limpeza de Dados (Desambiguação): A contagem de frequência de strings não realiza a desambiguação de nomes. Autores como 'alves, v' e 'alves, vh' seriam contados como duas pessoas distintas, inflando o número de autores únicos. O mesmo ocorre com pequenas variações de espaçamento ou abreviação.
- Natureza do *Corpus*: A Lei de Lotka é mais estável em campos científicos muito vastos. O *corpus* em questão, embora seja o "núcleo de elite" de Bradford, é relativamente pequeno (218 artigos) e focado, o que naturalmente eleva a proporção de autores de passagem.



karin'. Isso confirma o resultado da Lei de Lotka (Figura 8), onde esta autora foi identificada como a mais produtiva, com 4 publicações. Isso demonstra que a autora não apenas publica muito, mas também é central em um grupo de pesquisa.

- Rede Fragmentada: A rede de colaboração não é um único componente coeso, mas sim pulverizada em diversos *clusters* de cores distintas (vermelho, verde, azul, roxo, etc.) que não possuem conexões entre si.
- Grupos de Pesquisa: Os *clusters* representam grupos de pesquisa que publicam entre si, mas não colaboram com os outros grupos. O "*cluster verde*" (liderado por autores como 'rodrigues, vleso pereira' e 'alves, valerdy herdy', também identificados como de alta produtividade na Lei de Lotka) é um exemplo de grupo coeso, mas isolado do "*cluster vermelho*" de 'Erlandsson'.
- Autores Periféricos/Isolados: A grande quantidade de espaço em branco e de pequenos nós isolados (não visíveis neste nível de ampliação) corresponde aos grupos menos produtivos de Lotka: autores com menos publicações, que em sua maioria não colaboram com os núcleos de pesquisa estabelecidos.

### **3.9. Aplicação da Lei de Zipf: Análise de Frequência de Palavras**

A terceira lei bibliométrica clássica, a Lei de Zipf, foi aplicada para investigar a frequência de ocorrência de palavras e, assim, identificar os conceitos e temas centrais do debate. Esta análise é fundamental para complementar o que foi visto nas leis de Bradford (os periódicos) e Lotka (os autores).

Para obter uma visão mais granular e precisa, optou-se por aplicar a análise de frequência de forma separada em três campos de metadados distintos: os Títulos dos Artigos, as Palavras-Chave dos Autores e os Resumos (*Abstracts*). Esta seção inicia a análise pelo campo 'Article Title' (Títulos dos Artigos).

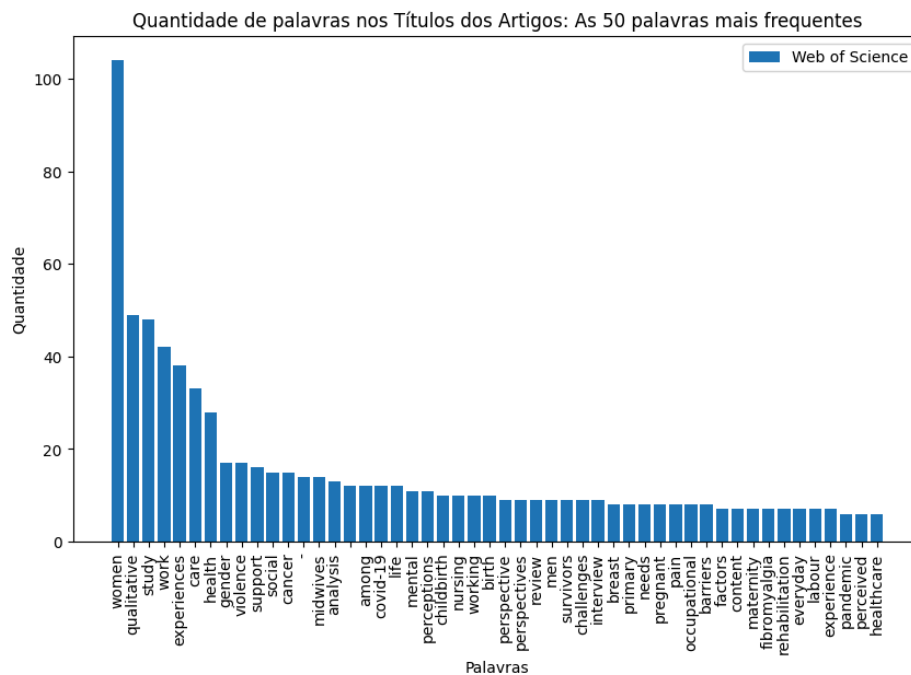
### **3.9.1. Análise dos Títulos (Article Title)**

Os títulos dos 218 artigos do corpus de elite foram processados (convertidos para minúsculas e remoção de toda pontuação) e, em seguida, organizados em *tokens* (divididos em palavras individuais). Foi aplicada uma extensa lista de *stop words* em inglês para remover termos de ligação (artigos, preposições, verbos auxiliares), garantindo que a contagem de frequência refletisse apenas palavras de conteúdo semântico.

O processo resultou em 1.021 palavras de conteúdo único. A Figura 10 apresenta as 50 palavras mais frequentes encontradas nos títulos.

A análise da Figura 10 fornece dados importantes sobre a natureza do corpus:

**Figura 10:** As 50 palavras mais frequentes nos Títulos dos Artigos



Fonte: Autoria Própria

- Validação do *Corpus*: As palavras mais frequentes são *women* (mulheres) (104 ocorrências) e *work* (trabalho) (42). Isso era esperado, pois são termos-chave da busca original, validando que o corpus de elite permanece focado no tema.
- O Achado Metodológico: O resultado mais significativo é a alta frequência de termos metodológicos. As palavras *qualitative* (49), *study* (48) e *experiences* (38) estão entre as cinco primeiras. Isso sugere fortemente que o "núcleo de elite" desta pesquisa é dominado por estudos qualitativos focados em capturar as experiências das mulheres, em vez de estudos quantitativos ou teóricos.
- Confirmação do Foco na Saúde: Os resultados corroboram fortemente as análises de Bradford e de citação. Termos como: *health* (saúde), *gender* (gênero), *violence* (violência), *support* (apoio), *cancer* (câncer), *midwives* (parteiras), covid-19, mental, *childbirth* (parto) e *nursing* (enfermagem) têm alta frequência. As presenças destas palavras nos títulos se alinham

perfeitamente com a identificação de periódicos como "Midwifery" e "Revista da Escola de Enfermagem" como o núcleo de Bradford.

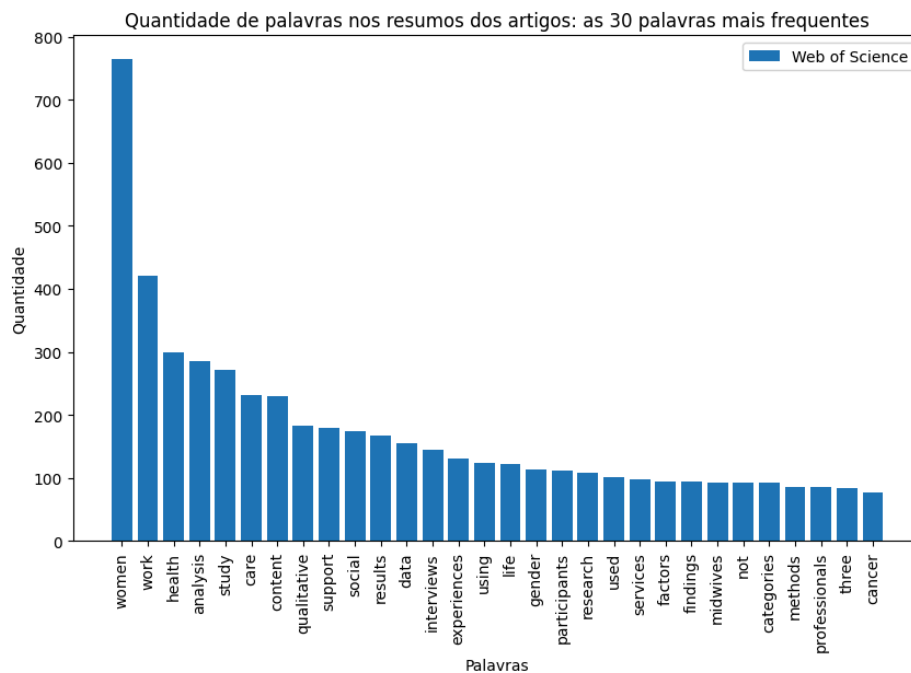
Em suma, a análise dos títulos revela que a pesquisa de elite sobre "mulheres e trabalho" (filtrada pela "análise de conteúdo") é, primariamente, um campo de estudo qualitativo focado na saúde da mulher.

### **3.9.2. Análise dos Resumos (Abstracts)**

Seguindo a análise de frequência, o mesmo procedimento foi aplicado ao campo 'Abstract' (Resumo) dos 218 artigos. Os resumos, por conterem a síntese da pesquisa (objetivos, método, resultados), oferecem um panorama mais detalhado dos conceitos que permeiam o *corpus*.

Após o processo de limpeza, "*tokenização*" e remoção de *stop words*, foi gerada a contagem de frequência das 5.757 palavras de conteúdo único. A Figura 11 apresenta as 30 palavras mais frequentes.

**Figura 11:** As 30 palavras mais frequentes nos Resumos dos Artigos



Fonte: Autoria Própria

A análise dos resumos (Figura 11) não apenas reforça os achados da análise de títulos, como também os expande com maior clareza:

- Validação dos Três Pilares: Os três termos mais frequentes validam perfeitamente o escopo da pesquisa: *women* (mulheres, 764 ocorrências), *work* (trabalho, 420) e *health* (saúde, 299). Isso solidifica a conclusão de que o núcleo da pesquisa sobre o tema está intrinsecamente ligado à saúde.
- Confirmação Metodológica Explícita: O achado mais importante é a confirmação da metodologia. Os termos metodológicos dominam o topo da lista:
  - *analysis* (análise, 285);
  - *study* (estudo, 272);
  - *content* (conteúdo, 231);
  - *qualitative* (qualitativa, 230);

- *results* (resultados, 182);
- *data* (dados, 172);
- *interviews* (entrevistas, 154);
- *participants* (participantes, 122);
- *methods* (métodos, 89);
- *categories* (categorias, 93).
- O surgimento da palavra *content* (conteúdo) em alta frequência é uma validação direta da *string* de busca original (*Content Analysis*). Isso prova que os artigos não apenas usam a Análise de Conteúdo, mas discutem ela como método central em seus resumos. Fica claro que o método predominante é a análise de conteúdo qualitativa, aplicada sobre dados coletados majoritariamente por entrevistas com participantes, com o fim de gerar resultados e categorias.
- Confirmação do Tópico (Saúde e Cuidado): Os termos temáticos corroboram o foco no cuidado: *care* (cuidado, 290), *support* (apoio, 181), *social* (social, 176), *midwives* (parteiras, 96), *professionals* (profissionais, 87) e *cancer* (câncer, 80).

Em síntese, a análise dos resumos confirma que o núcleo de elite (Bradford) desta pesquisa é formado por estudos que aplicam a análise de conteúdo qualitativa, baseada em entrevistas, para investigar as experiências de mulheres sobre trabalho no contexto do cuidado em saúde. Estes resultados tem uma implicação direta para a segunda fase desta pesquisa: a análise de conteúdo da

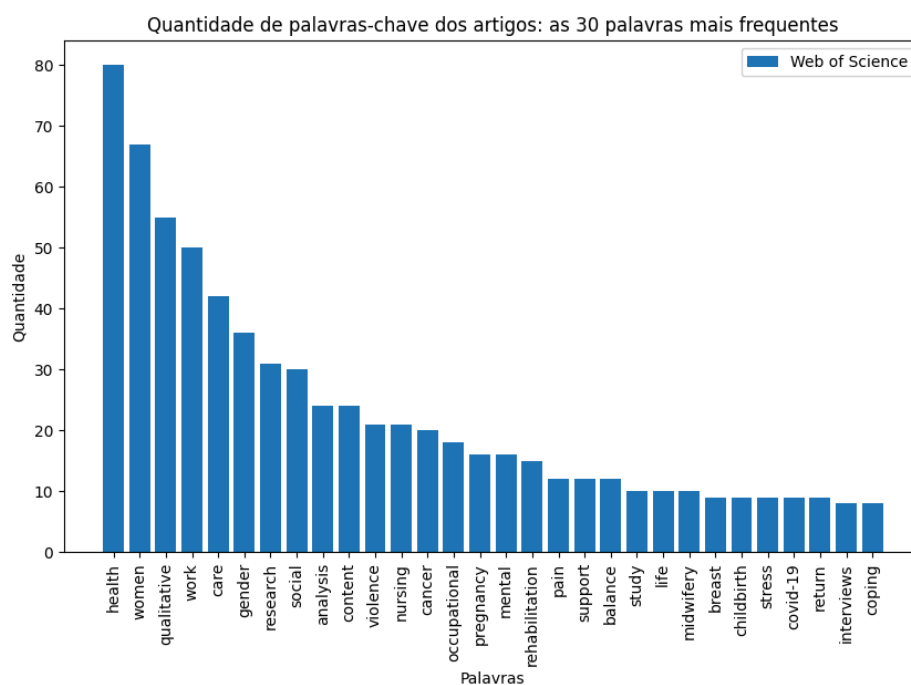
entrevista com a pescadora artesanal (como mencionado na metodologia), que se alinha perfeitamente com a prática científica dominante identificada no *corpus*.

### 3.9.3. Análise das Palavras-chave (Author Keywords)

Finalmente, a análise de frequência de palavras (Lei de Zipf) foi aplicada ao campo 'Author Keywords' (Palavras-Chave do Autor). Este campo representa os termos que os próprios autores selecionaram para categorizar seus trabalhos, indicando uma intenção temática explícita.

O mesmo processo de limpeza, "tokenização" e remoção de *stop words* foi aplicado às palavras-chave dos 218 artigos do corpus de elite, resultando em 746 palavras de conteúdo únicas. A Figura 12 apresenta as 30 palavras mais frequentes.

**Figura 12:** As 30 palavras mais frequentes nas Palavras-Chave dos Autores.



Fonte: Autoria própria

A análise das palavras-chave (Figura 12) indicam tendências anteriores, consolidando os resultados:

- **Consolidação dos Temas Dominante:** Três áreas temáticas do corpus são inegáveis. Os termos *health* (saúde, 80 ocorrências), *women* (mulheres, 67) e *work* (trabalho, 50) são os mais frequentes, com *health* surgindo como o tópico de indexação mais utilizado pelos autores.
- **O Foco Metodológico:** Assim como nos Títulos e Resumos, o método é um tema central. *Qualitative* (qualitativa, 55) é a terceira palavra-chave mais comum. Isso prova que o campo não apenas usa o método, mas se identifica como um campo de pesquisa qualitativa.
- **O Agrupamento de Saúde e Cuidado:** O restante da lista de alta frequência é uma confirmação dos sub-temas de saúde e cuidado: *care* (cuidado, 42), *gender* (gênero, 36), *nursing* (enfermagem, 21), *cancer* (câncer, 18), *occupational* (ocupacional, 16), *pregnancy* (gravidez, 16), *mental* (mental, 15) e *midwifery* (obstetrícia, 10).

Em suma, as três análises da Lei de Zipf (Títulos, Resumos e Palavras-chave) convergem para a mesma conclusão: o núcleo de elite da pesquisa que utiliza Análise de Conteúdo para estudar mulheres e trabalho é predominantemente um campo de estudo qualitativo focado na saúde da mulher, especialmente no contexto do cuidado (enfermagem, obstetrícia) e experiências de saúde (câncer, saúde mental, parto).

#### **3.9.4. Síntese da Frequência de Palavras: Nuvem Temática**



*interviews* (entrevistas). Por fim, os tópicos mais específicos, como *midwives* (parteiras), *violence* (violência) e *cancer* (câncer), completam o léxico principal do campo.

#### **4. CONCLUSÃO**

Este estudo teve como objetivo central traçar o panorama bibliométrico da produção científica que utiliza a Análise de Conteúdo para investigar a complexa relação entre mulheres e trabalho. A análise, partindo de um universo de 1.669 publicações e refinada para um corpus final de 867 artigos, permitiu identificar os contornos, os protagonistas e as tendências de um campo de pesquisa vasto e multidisciplinar.

Os resultados das leis bibliométricas clássicas foram reveladores. A Lei de Bradford, conforme analisado na Seção 3.6, demonstrou que o "núcleo de elite" da pesquisa não está em periódicos de sociologia do trabalho. Como detalhado na análise de áreas de pesquisa (Figura 5) e no ranking de periódicos (Figura 7), este núcleo está concentrado massivamente em publicações das áreas de Enfermagem (*Nursing*) e Saúde Pública.

A Lei de Lotka (Seção 3.8) expôs uma comunidade científica altamente fragmentada. Como detalhado na Tabela 2 e na análise subsequente, 97,6% dos 839 autores únicos do corpus publicaram apenas um artigo. Este achado, complementado pelo mapa de coautoria (Figura 9), revela uma baixa colaboração e a ausência de um núcleo coeso de pesquisadores.

Contudo, foi a análise da Lei de Zipf (Seção 3.9) que forneceu a evidência mais contundente. Os termos mais frequentes identificados nos títulos (Figura 10), resumos (Figura 11) e palavras-

chave (Figura 12) – como *qualitative, health, care, experiences, interviews* e *midwives* – confirmam que o foco acadêmico é majoritariamente ligado a saúde da mulher (materna, ginecológica, ocupacional) e suas experiências em contextos de cuidado ou violência, como sintetizado na nuvem de palavras (Figura 13).

Estes dados, vistos em conjunto, nos levam de volta à motivação inicial desta pesquisa, inspirada nos estudos do PEA Pescarte. O que a bibliometria revela não é apenas a marginalização comum da mulher em relações de trabalho, mas uma invisibilidade dupla e agravada quando o contexto é o da pesca artesanal. A aparente carência de artigos que tratem da mulher pescadora não é uma coincidência; é um sintoma de uma lacuna científica. Enquanto a academia dedica centenas de estudos para aplicar a Análise de Conteúdo em contextos de saúde e corporativos, o trabalho extrativista e artesanal feminino, como o observado no litoral fluminense, permanece metodologicamente ignorado.

Portanto, este estudo bibliométrico cumpre sua função ao fornecer o "estado da arte", mas sua principal contribuição é demonstrar onde a pesquisa não está. Ele valida a urgência da próxima fase desta investigação: a aplicação da Análise de Conteúdo (Bardin) nas entrevistas com pescadoras artesanais. Tal estudo não será apenas mais uma contribuição ao campo, mas um esforço necessário para preencher um vácuo científico e começar a corrigir uma invisibilidade histórica.

## **REFERÊNCIAS BIBLIOGRÁFICAS**

BARDIN, L. **Análise de conteúdo**. 1. ed. São Paulo: Edições 70, 2016. Tradução de Luís Antero Reto e Augusto Pinheiro. 3ª reimpressão da

1ª edição. ISBN 978-85-62938-04-7.

FERREIRA, A. G. C. Bibliometria na avaliação de periódicos científicos. **DataGramaZero-Revista de Ciência da Informação**, v. 11, n. 3, p. 1–9, 2010.

GIL, A. C. **Como elaborar projetos de pesquisa**. [S. l.]: Atlas São Paulo, 2002. v. 4.

GRÁCIO, M. C. C.; MARTÍNEZ-ÁVILA, D.; OLIVEIRA, E. F. T. d.; ROSAS, F. S. **Tópicos da bibliometria para bibliotecas universitárias**. [S. l.]: Editora UNESP, 2020.

GUEDES, V. L.; BORSCHIVER, S. Bibliometria: uma ferramenta estatística para a gestão da informação e do conhecimento, em sistemas de informação, de comunicação e de avaliação científica e tecnológica. **Encontro nacional de ciência da informação**, v. 6, n. 1, p. 18, 2005.

LOUSADA, M.; GARCIA, C. L. S.; WOIDA, L. M.; DAL'EVEDOVE, P.; GARCIA, R.; VALENTIM, M. L. P. Produção científica sobre gestão do conhecimento e gestão da informação no âmbito da ciência da informação: uma aplicação da Lei de Bradford. In: FACULTAD DE COMUNICACIÓN Y DOCUMENTACIÓN Y SERVICIO DE PUBLICACIONES DE LA . . ., 2. ANALES de documentación. [S. l.: s. n.], 2012. v. 15.

MELO RIBEIRO, H. C. Bibliometria: quinze anos de análise da produção acadêmica em periódicos brasileiros. **Biblios**, Los autores, n. 69, p. 1–20, 2017.

TIMÓTEO, G. M. (ed.). **Trabalho e Pesca no Litoral Fluminense: Reflexões a partir do Censo do PEA Pescarte**. 1. ed. Campos dos Goytacazes, RJ: EdUENF, 2019. ISBN 978-85-89479-56-1.

---

<sup>1</sup> Discente do Curso de Doutorado em Cognição e Linguagem – Universidade Estadual do Norte Fluminense (UENF). E-mail: [acesse o artigo original para visualizar o e-mail](#)

<sup>2</sup> Docente do Curso de Doutorado em Cognição e Linguagem – Universidade Estadual do Norte Fluminense (UENF).

<sup>3</sup> Discente do Curso de Mestrado em Cognição e Linguagem – Universidade Estadual do Norte Fluminense (UENF).

<sup>4</sup> Discente do Curso de Mestrado em Cognição e Linguagem – Universidade Estadual do Norte Fluminense (UENF).

<sup>5</sup> Docente do Curso de Doutorado em Cognição e Linguagem – Universidade Estadual do Norte Fluminense (UENF).

<sup>6</sup> Discente do Curso de Pós-Doutorado em Cognição e Linguagem – Universidade Estadual do Norte Fluminense (UENF).