

Analysis of the trend of middle-aged women's research using text network analysis and topic modeling

Do-Young Lee¹

¹ Associate Professor, College of Nursing, Changshin University, Changwon, Korea, shine@cs.ac.kr
Corresponding author: Do-Young Lee

ABSTRACT

This study aims to understand the research trends and key concepts related to education among middle-aged women in Korea. To analyze the research, we collected and analyzed a total of 200 papers by inputting the keywords 'middle-aged women' and 'education' from 2003 to 2022, excluding papers with unavailable abstracts, inaccessible downloads, and duplicates. We visualized the data through word clouds and sociograms. The search scope was limited to domestic academic papers, and we extracted bibliographic information for all papers retrieved using the keywords 'middle-aged women' and 'education' into yearly Excel files spanning 20 years from 2003 to 2022. After excluding duplicate papers, we identified 200 papers related to education for middle-aged women. We examined the trends and topics in health-related research on middle-aged women over the past 20 years and suggest that research on middle-aged women's health, reflecting future trends, should continue.

KEYWORDS:- Text Network Analysis, Topic modeling, A middle-aged woman, Research trends, Education

Date of Submission: 01-09-2024

Date of acceptance: 11-09-2024

I. INTRODUCTION

Middle age is a transitional period from aging adolescence to old age, and women become interested in themselves and begin to rediscover their lives as they become free from the child-rearing role[1]. Middle-aged women are experienced in various life events such as family care and childbirth, and they try to re-establish their life structure with their modified selves through individualization [2]. Therefore, the research topic on education only for middle-aged women is attracting attention [2]. In particular, middle-aged women are attracting attention as learners in countermeasures and life redesign for career-interrupted women [3]. The demand of the times leads to an increase in women's participation in vocational education, and women's education is gradually increasing [4].

In order to overcome difficulties and lead a better life in the middle-aged and old age, middle-aged women are participating in education, and they are leading a healthy family by restoring confidence and trust in themselves through "self-reflection," "self-realization," and "conversion of thinking" [5]. Advanced science and medical technology in modern society have led to an increase in the elderly population and prolonged old age, especially Korean women have longer life expectancy than men, making women's old age longer. Preparing to spend old age successfully is a very important task for middle-aged people. Therefore, it is necessary to understand trends in research topics such as how middle-aged women's education has been conducted for 20 years, which topics have become the core, or what researchers' interest was like over time. Understanding how the research topic has changed can provide the necessary meaning to try and advance new research in the future by synthesizing existing research achievements in the specialized field. Text network analysis is being used as a research method for analyzing these parts. It can collect and analyze big data and include all related data, and it can provide an objective basis by quantitatively and qualitatively interpreting the research results found by applying scientific rigor.

Therefore, this study attempted to explore the knowledge structure of the research topic of middle-aged women's health care by applying text network analysis to identify key topics for trend research on middle-aged women's education published over the past few years and visualizing the main semantic structure in sociograms. In addition, by analyzing changes in research topics over time, we would like to identify research topics that are attracting attention to middle-aged women's education.

The knowledge structure is a structure that grasps knowledge of a specific discipline as quantitative or relational data, and can grasp the characteristics that appear in the field. Methods of identifying the knowledge structure as a network include journal network analysis, author network analysis, literature network analysis, and

text network analysis [6]. Among them, text network analysis is a method that can interpret phenomena and structures by visualizing the relationship between concepts in a network [7]. It is a method of decomposing the words that make up the text of the paper to extract meaningful concepts, analyze the relationship between words in the form of a network, and identify the potential semantic structure inherent in the context [8]. Quantitative content analysis applying text network analysis methods identifies certain patterns in scattered data by quantitatively analyzing large amounts of papers and qualitatively interpreting the found knowledge structure [9]. Text network analysis is an analysis technique that extracts words from literature in a specific field, calculates the frequency at which each word appeared simultaneously, classifies the subject area of the field through various statistical techniques, identifies key main controls, or identifies the relationship between subject areas. Of course, the high frequency of simultaneous appearance presupposes that there is a high connection between the main words [10]. Freeman (1979)'s degree centrality, closeness centrality, and betweenness centrality are the three most basic indicators [10]. Connection centrality evaluates the degree to which it appears simultaneously with other keywords, i.e., activity, and keywords with high connection centrality are the core of the network, i.e., the topic of discussion [12][13][14]. Mediation centrality is an index that measures the degree to which a particular concept acts as a mediator when forming a network with other concepts, and if keywords with high mediation centrality are eliminated, it is impossible to connect from one topic to another. Proximity centrality calculates the connection distance between concepts to show how close a concept targeting a whole concept within a network is to another concept. When proximity centrality increases, it means that it is located at the center of the network on a distance, and it may be easier to secure and access influence and status within the network.

Topic modeling is an analysis method that can discover topics hidden in large amounts of text data and is a probability model algorithm [15]. It is possible to derive meaningful information and structures that could not be found before from unstructured text data that require too much effort and time to process by the hands of researchers [16]. The advantage of topic modeling is that it is possible to analyze the research topic more objectively and clearly through statistical reasoning to evaluate the research topic that can involve the subjectivity of the researcher [17].

Programs for network analysis include UCINET, Pajek, NetMiner, etc. [18]. The biggest advantage of NetMiner is that it can simultaneously input, convert, add, and analyze the direct relationship between nodes (1-mode, Adjacency), the relationship between nodes (2-mode, Affiliation) within a single data file, and the attribute of the nodes (attributes), and simultaneously visualize the network [18].

This study applied text network analysis to understand the education-related trends of middle-aged women in Korea. The specific purpose is to identify key topics, key semantic structures, and recent research topics. First, we identify the key words of middle-aged women's research.

Second, it confirms the connection centrality, which is the relationship between the main keywords of middle-aged women's research.

Third, the flow and trend of middle-aged women's research topics are identified.

II. Research methods

This study is a content analysis study that extracts keywords from domestic academic research that explore topics related to middle-aged women's education as text data and analyzes them by network analysis and topic modeling.

This study was conducted in the process of collecting academic research and extracting keywords using bibliographic information, developing a preprocessing and co-emerging network, and analyzing topic modeling.

Papers for use in the analysis of this study were collected from Research Information Sharing Service (RISS) and Korea Citation Index (KCI), which can extract bibliographic information from literature search databases. For the search area, domestic academic papers were selected, and the bibliographic information of all confirmed papers was extracted as an annual Excel file for 19 years from 2003 to 2022 by entering 'middle-aged women' and 'education' as keywords. Excluding the papers that were identified as duplicates, 200 papers on the education of middle-aged women were identified.

In this study, the keyword data refined through the preprocessing process were analyzed by centrality analysis and topic modeling through the 1-mode matrix conversion process using the NetMiner (version 4.4) program. In the analysis, a set of words or proper nouns expressing a meaning were set as designated words.

As an analysis method for keywords, keywords with a frequency of more than 10 times were identified through the frequency of appearance. Next, for keywords with high appearance frequency, the meaning was analyzed through connection centrality through the number of directly connected neighboring nodes. If the value of degree centrality is high, the keyword can easily spread to other keywords, which is influential, and it is recognized as important because it is considered a central topic of discussion at the core of the network. In addition, the type of study was analyzed through topic modeling using LSD techniques, and visualized word clouds and sociograms were presented.

Keyword analysis methods are as follows.

- 1) All extracted concepts are separated into one word based on spacing. However, 'middle age', which means middle age, is a keyword, and in order to see it as a single word, it was modified and expressed as 'middleage'.
- 2) Adjective keywords were modified into nouns.
- 3) If the meaning is similar or implied, all of them were modified into one word, and then the words changed in the abstract were modified and analyzed. For example, in the case of middle age, it was considered to be used in the same meaning in the context of the sentence and went through the step of revising it to 'middle age'-> 'middle age'.

The total frequency was analyzed for 690 confirmed words, and the network map transformed the co-occurrence frequency matrix (coocc.dbf) file extracted through Krttle into an Excel file during the KrKwic program, and then visualized the correlation as a graph using the NetMiner program. When an isolated node that is not related to other subject words appeared after the network map was completed, the map was created by constructing only related words except for the isolated node.

This study received an exemption from review from the Institutional Review Board for Life Ethics at C University (IRB No.: CSIRB-2023006), and research ethics procedures were conducted according to the exemption, resulting in the derivation of research outcomes.

III. Research results

A total of 200 studies have been published in Korea from 2003 to 2022, the main keywords for middle-aged women and education-related research. Since 2003, it has been steadily published in relation to the education of middle-aged women.

In the paper analyzed by species, 690 keywords were presented. Table 1 shows the key keywords of the frequency of appearance ranking up to 26th (frequency 10 or higher) in the paper analyzed in this study. The most common keywords appeared were middle age 159 times, followed by women 154 times, education 49 times, and life 48 times. Health (41 times), learning (39 times), program (31 times), lifelong (22), expiration (21), depression (19 times), and satisfaction (19 times).

[Table 1] Key Words in Research on Middle-Aged Women's Education

Rank	Keyword	Frequency
1	middleage	159
2	woman	154
3	education	49
4	life	48
5	health	41
6	learning	39
7	program	31
8	lifelong	22
9	exercise	21
10	depression	19
11	satisfaction	19
12	quality	17
13	female	15
14	behavior	14
15	experience	14
16	obesity	14
17	participation	14
18	stress	14
19	metabolic	13
20	research	13
21	dietary	12
22	nutrition	12
23	study	11

24	therapy	11
25	family	10
26	syndrome	10
27	support	10
28	career	10
29	death	9
30	learner	9

Relationship between Key Words in Middle-Aged Women's Education Research

The 30 highly influential keywords identified as a result of the analysis of the connection centrality of the keywords presented in the paper to be analyzed in this study are as shown in [Table 2]. In the connection centrality analysis, which checks the influence through the number of keywords connected around it, it was confirmed in the order of woman, middleage, education, life, and program. In the betweenness centrality analysis based on the shortest distance from other keywords, the highest ranking keywords were women, middleage, education, health, and life. Keywords with high betweenness centrality play a role in connecting expansion to other topics and become a key keyword to be considered when expanding related topics. [Table 2].

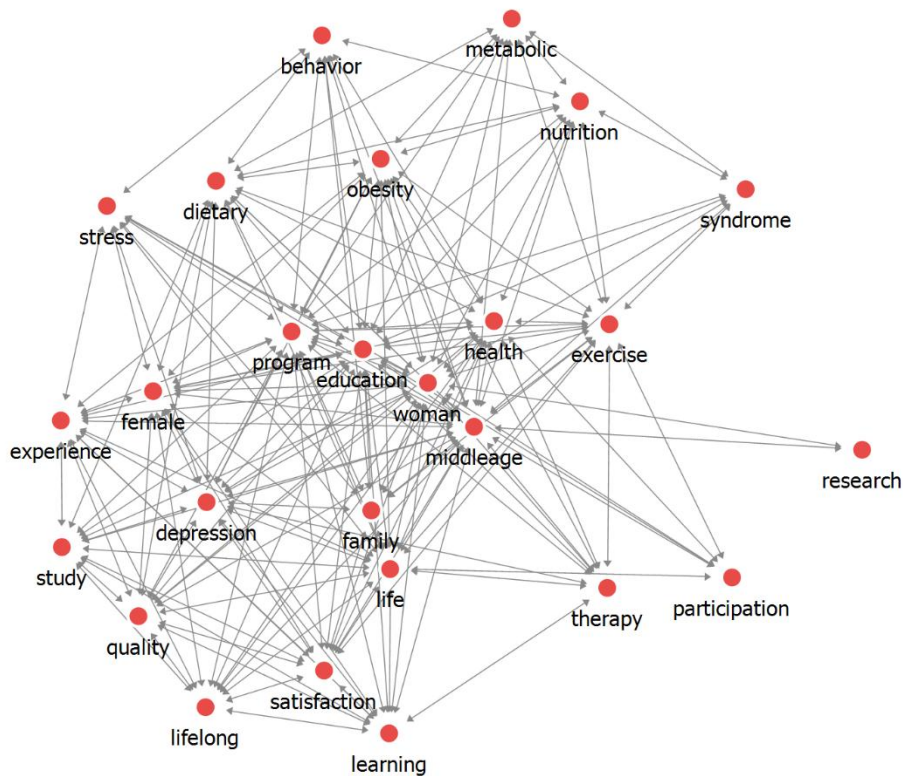
[Table 2] Relationships between key words

Rank	Keyword	Frequency	Keyword	Degree Centrality	Keyword	Between Centrality
1	middleage	159	woman	0.927	woman	0.123
2	woman	154	middleage	0.890	middleage	0.095
3	education	49	education	0.829	education	0.076
4	life	48	life	0.695	health	0.045
5	health	41	program	0.671	life	0.038
6	learning	39	health	0.610	program	0.036
7	program	31	age	0.512	learner	0.025
8	lifelong	22	depression	0.451	medical	0.021
9	exercise	21	meaning	0.427	support	0.015
10	depression	19	satisfaction	0.415	syndrome	0.014
11	satisfaction	19	family	0.390	depression	0.012
12	quality	17	learning	0.378	perception	0.010
13	female	15	female	0.378	therapy	0.010
14	behavior	14	exercise	0.354	satisfaction	0.009
15	experience	14	study	0.341	exercise	0.009
16	obesity	14	aging	0.317	female	0.008
17	participation	14	experience	0.305	career	0.007
18	stress	14	support	0.305	learning	0.006
19	metabolic	13	anxiety	0.305	transition	0.005
20	research	13	body	0.305	model	0.004
21	dietary	12	lifelong	0.293	quality	0.004
22	nutrition	12	quality	0.293	risk	0.004
23	study	11	counseling	0.293	death	0.004
24	therapy	11	job	0.293	obesity	0.004
25	family	10	dietary	0.280	dietary	0.004
26	syndrome	10	activity	0.280	theory	0.003

27	support	10	obesity	0.268	experience	0.003
28	career	10	status	0.268	gender	0.003
29	death	9	adult	0.268	menopause	0.003
30	learner	9	stress	0.232	lifelong	0.003

Flow and Trends of Research Topics Related to Middle-Aged Women's Education

The sociogram using the overall topic modeling of this study is shown in [Fig. 1], and the sociogram visualizing the relationship between the main keywords for the four topics is shown in [Fig. 2]. [Fig. 3] shows the word cloud for each period drawn based on the frequency of occurrence of keywords in the topic modeling analysis process.



[Fig. 1] Sociogram using topic modeling

Middle-aged women become self-directed learners with accumulated life experiences while experiencing changes in their roles and values during this period of life transition [21]. In other words, choosing new learning serves as a protective factor that reduces risk factors in middle age and helps achieve developmental tasks [22]. The education rate of middle-aged women is steadily increasing, and accordingly, interest in middle-aged female learners and educational content are also increasing. It is important to organize the learning process more systematically and identify the necessary parts to promote qualitative stability.

This work was supported by the National Research Foundation of Korea grant funded by the Korean government (No. 2022R1G1A1013123).

REFERENCE

- [1]. M. S. Jun, H. S. Lee, The Effects of Participation Motivation for Vocational Education Program of Middle-aged Female Learners' on Participation Continuation Intention and the Mediating Effects of Educational Satisfaction, *Journal of Lifelong Learning Society*, (2017), Vol 13, No.2, pp 109-133. DOI: 10.26857/JLLS.2017.05.13.2.109
- [2]. Y. J. Hwang, H. S. Lee, A Study on the Effect Analysis of the Re-Employment Program for Career-Interrupted Women Using a Program Logic Model : Focused on the Case of A Seoul S-Center, *The Journal of Vocational Education Research*, (2014), Vol.33, No.6, pp 55-79. <http://www.kosove.re.kr/?c=5/100>
- [3]. Report: Y. J. Choi, H. S. Chung, G. W. Ban, S. M. Kim, Competencies of Korean Women and Its Implications: Comparative Analysis Based on OECD PIAAC. Seoul: Korean Women's Development Institute. (2016). [https://www.kwdi.re.kr/flexer/view.jsp?FileDir=/CM001&SystemFileName=CM0004_1268_1&ftype=pdf&FileName=%EC%B7%A8%EC%97%85%20%EB%B0%8F%20%EB%B9%84%EC%B7%A8%EC%97%85%20%EC%97%AC%EC%84%B1%EC%9D%98%20%EC%97%AD%EB%9F%89%20%EC%8B%A4%ED%83%9C%EC%99%80%20%EA%B3%BC%EC%A0%9C%20-%20%EC%B5%9C%EC%9C%A4%EC%A0%95\(%EB%B3%B4%EC%9D%B4%EC%8A%A4%EC%95%84%EC%9D%B4\).pdf](https://www.kwdi.re.kr/flexer/view.jsp?FileDir=/CM001&SystemFileName=CM0004_1268_1&ftype=pdf&FileName=%EC%B7%A8%EC%97%85%20%EB%B0%8F%20%EB%B9%84%EC%B7%A8%EC%97%85%20%EC%97%AC%EC%84%B1%EC%9D%98%20%EC%97%AD%EB%9F%89%20%EC%8B%A4%ED%83%9C%EC%99%80%20%EA%B3%BC%EC%A0%9C%20-%20%EC%B5%9C%EC%9C%A4%EC%A0%95(%EB%B3%B4%EC%9D%B4%EC%8A%A4%EC%95%84%EC%9D%B4).pdf)
- [4]. Report: Ministry of Education, 2015 Lifelong Learning of Korean Adults, Seoul: Korean Education Development Institute, (2015), <https://kess.kedi.re.kr/publ/view?survSeq=2015&publSeq=24&menuSeq=0&itemCode=02&language=>
- [5]. M. H. Park, M. Jeong, The Study on the Effects of Middle-aged Woman's Participative Motivation and Lifelong Education Satisfaction on Subjection Well-being: Focused on the Medicated Effect of Self-efficacy. *The Journal of Humanities and Social Sciences 21*. (2016), Vol. 7 No. 6, 19-40. DOI <http://dx.doi.org/10.22143/HSS21.7.6.2>
- [6]. H. J. Seo, J. Y. Lee, Analysis of Future Education Trends Using Semantic Network Analysis. *The Journal of Education Information and Media*, (2018), Vol. 24, No. 4, 649-678. DOI 10.15833/KAFEIAM.24.4.649
- [7]. J. H. Lee, H. S. Lee, H. K. Lee, Research on Methods for Processing Nonstandard Korean Words on Social Network Service. *Journal of the Korea Industrial Information Systems Research*. (2016), Vol. 21 No.3, 35-46. DOI: [10.9723/jksjis.2016.21.3.035](https://doi.org/10.9723/jksjis.2016.21.3.035)
- [8]. C. S. Park, C. W. Chung, Text Network Analysis: Detecting Shared Meaning through Socio-cognitive Networks of Policy Stakeholders, *Institute of Governmental Studies*, (2013). Vol. 19, No.2, 73-108.
- [9]. C. S. Park, E. J. Park, Identification of Knowledge Structure of Pain Management Nursing Research Applying Text Network Analysis, *Journal of Korean academy of nursing*, (2019), Vol. 49, No. 5, pp. 538-549. DOI : [10.4040/jkan.2019.49.5.538](https://doi.org/10.4040/jkan.2019.49.5.538)
- [10]. For Book: S. S. Lee, *Social Network Analysis*, Seoul:Nonhyungbook, (2012)
- [11]. L. C. Freeman, Centrality in social networks conceptual clarification *Social Networks*, (1979), Vol. 1, No. 3, 215-239.
- [12]. Kim Yeon Hee, Moon Seong Mi, Kwon In Gak, Kim Kwang Sung, Jeong Geum Hee, Shin Eun Suk, Oh Hyang Soon, Kim Soo Hyun, Research Trends of Articles Published in the Journal of Korean Clinical Nursing Research from 2000 to 2017: Text Network Analysis of Keywords, *Journal of Korean clinical nursing research*, (2019), Vol. 25, No. 1, pp. 80-90. DOI: 10.22650/JKCNR.2010.16.2.95
- [13]. For Book: Y. H. Lee, Y. J. Kim, *Social Network Analysis: 4th*, Seoul: PYbook, (2016)
- [14]. Blei, D. M., Probabilistic topic models, *Communications of the ACM*, (2012), Vol.5, issue4, 77-84, 10.1145/2133806
- [15]. Blei, D. M., & Lafferty, J. D, Topic Models. In A. Srivastava, & M. Sahami (Eds.), *Open Journal of Social Sciences*, (2021), Vol.9, No.2, 71-93, 10.4236/jss.2021.92028
- [16]. B. E. Son, D. I. Jeon, Research Trends Analysis in Quality Monitoring of the General Education Using Text Network Analysis, *Korean Journal of General Education*, (2020). Vol. 14, No. 5, 97-108. DOI: 10.46392/kjge.2020.14.5.97
- [17]. Y. H. Lee, *Social Network Analysis*, Seoul: PYbook, (2011).
- [18]. M. H. Lee, (A) Study on the Social Network to Weather the Crisis of a Middle-aged Woman. hannam University. Master's thesis (2001).
- [19]. DaeMyung kim, The Structural Relationship among Middle-aged Women's Lifelong Education Participative Motivation, their Self-Directed Learning, Educational Service Quality and their Program Satisfaction, Department of Education The graduate School Chonbuk National University, (2013).
- [20]. S. H. Lee, J. H. Kim, The Structural Relationship among Learner Characteristics, Participatory Motivation, Learning Satisfaction and Lifelong Education Participation Performance of Middle-Aged Women. *Journal of Lifelong Education*, (2016). Vol. 22, No. 3, 149-178.. <https://dlps.nanet.go.kr/SearchDetailView.do?cn=KINX2016292782&sysid=nhn>
- [21]. Hyesuk Ha, Minji Park, The Mediating Effect of Middle-Aged Female Learner's Academic Engagement on the Relationship between Participative Motivation and Learning Satisfaction, *Journal of Lifelong Learning Society*, (2019), Institute of Future Distance Education, Vol. 15, No. 2, pp. 109-126. DOI: 10.26857/JLLS.2019.5.15.2.109