JTS

The Journal of Transdisciplinary Studies, 8(3): 263-271, 2024
https://doi.org/10.1001/jts.2024.8.3.263

Veterinary Research Trends Based on Semantic Network Analysis

Jae-Hee Roh

Professor, Dept. of Pet Health, Kwangju Women's University, Korea (snowida@kwu.ac.kr)

ARTICLE INFO

Article history:
Received 01 Nov 2024
Revised 25 Nov 2024
Accepted 26 Nov 2024

Keywords:
Journal of Veterinary Science,
Research trends,
Semantic network analysis,
Veterinary medicine

ABSTRACT

This study aimed to examine veterinary research trends by applying a semantic network analysis to the main keywords of articles. The analysis was conducted on 283 academic articles published in the Journal of Veterinary Science from 2020 to 2022, and the main words were refined and extracted from the English keywords presented in each article. When term frequency-inverse document frequency and term frequency were combined, the top 10 research keywords were canine, swine, feline, avian, and bovine for livestock species, and virus, cell, disease, fever, and analysis for non-livestock species. The results of an N-gram analysis also revealed that African swine fever and stem cells were the top research topics identified. The results of a centrality analysis showed that disease, analysis, and virus ranked highest in terms of research interest and impact for non-livestock species, whereas swine ranked highest in terms of those for livestock species. The results of a research trend analysis provided further useful information to guide future research and present directions on the topic. The results of this study are expected to provide evidence for identifying research trends in veterinary science and provide useful information for future research.

1. Introduction

The Korean Society of Veterinary Science (KSVS) was founded in July of 1957 and began publishing its specialized journal, the Korean Journal of Veterinary Research, in 1961. Now, KSVS publishes a total of 10 issues a year, including four issues of the Korean journal (Korean Journal of Veterinary Research) and six issues of the English journal (Journal of Veterinary Science, J Vet Sci). KSVS is also a representative Korean academic society in the field of veterinary science that holds spring and fall conferences twice a year and is responsible for active academic exchanges on science and technology in the field of veterinary science. Moreover, J Vet Sci is a leading domestic journal in the field of veterinary medicine, and

it has published more than 80 articles annually for the past five years (KCI, 2024.; KSVS, 2024). In addition, this journal is listed in Medline, PubMed, Index Medicus, Scopus, and SCI, and is ranked 44th in 194 general veterinary categories with a Scopus citescore of 3.1 in 2023 (Scopus Preview, 2024).

There is a need for information with which to identify research trends and design directions for further research. Therefore, this study aims to investigate veterinary research trends by analyzing research topics and keywords to provide useful information about the research that needs to be conducted in the future.

Previously, research trend analyses were conducted by having researchers read, code, and analyze documents, which could have resulted in subjective interventions by the researchers and limitations on validity (Jang et al., 2017; Park & Leydesdorff, 2004; Kim, 2018). However, in recent years, researchers have increasingly turned to text analytics, which is a big data analytics method that allows for the fast and effective analysis of large amounts of information. It enables new perspectives and predictions by identifying relationships and meaning in the collected data and provides objectivity to solve problems in existing research (Reardon, 2014; Kang et al., 2018).

Semantic network analysis (SNA) is a method of analysis that identifies connections based on co-occurrence relationships between words and examines how they change through the semantic structure and context of a document (Hwang & Hwang, 2018; Lee, 2014). It was initially called social network service analysis and has since been used in various ways, including co-occurrence network, language network, keyword network, and keyword linkage analyses (Jung & Lee, 2020; Hwang & Kim, 2019; Hwang & Park, 2019). This study used the term SNA and identified semantic relationships between keywords to identify research trends. SNA measures density and centrality, which represent the degree of connectivity between nodes. The higher the number of links to a node that represents a keyword, the higher the result. Centrality is the most commonly sought-after attribute in network analysis, as it expresses the degree to which a connection is central to the overall network. This allows for an understanding of where each node is located and what role it plays in the overall structure. Therefore, words with high degree centrality are the most frequent words and play leading roles in building logic about a topic (Choi et al., 2021; Hwang & Kim, 2019; Hwang & Shim, 2020).

Research trend analysis identifies trends in research and thereby provides useful information to guide future research. Recently, a growing number of studies have utilized SNA to identify research trends or trends in social phenomena via language analysis (Kang & Namkung, 2021; KCI, 2022; Kim et al., 2021; Lee et al., 2017).

Therefore, this study applied SNA to the main keywords to examine the connectivity and centrality between keywords and to identify the research trends of academic articles in the field of veterinary science published in J Vet Sci via the extracted results.

2. Methods

2.1 Analysis target

This study analyzed scientific articles published in J Vet Sci from 2020 to 2022 to identify research trends in veterinary science. By confirming the research trends related to this from 2020, the year after the first outbreak of African swine fever in Korea, we selected this period as it could serve as an indicator for the analysis of veterinary research trends using semantic network analysis. A total of 98 articles from 2020, 84 articles from 2021, and 101 articles from 2022 were collected from the Research Information Sharing Service (RISS), and the English keywords presented in the articles were analyzed.

2.2 Analysis methods

After the keywords were collected, they were refined using the Notepad++ program to unify case as well as singular and plural forms such that they would be recognized as one keyword. Then, the main words were refined and extracted again using the Textom program (Textom 6.0, TheIMC, Korea) and analyzed in the following manner. First, in a keyword frequency analysis, term frequency (TF) indicates how often a particular word appears within a document, and term frequency-inverse document frequency (TF-IDF) quantifies the importance of a word within a document, with a higher value indicating that the word is considered more meaningful and important to the document. Second, N-gram analysis is a statistical language analysis model that extracts N consecutive elements from a string and counts the words that occur simultaneously in a sentence and their frequencies. Third, centrality analysis quantifies the relative importance of a particular word in the network, which can be used to identify more influential words in terms of their structural positions. Degree centrality is the simplest and most basic indicator in the centrality index, which evaluates the index by the number of all edges directly connected to a node. In this analysis, the results were normalized to a range of values from 0 to 1 to enable comparison with other networks. Fourth, co-occurrence keyword analysis identifies two or more words that occur together within a certain range or distance, which can help in understanding the context in which they are used together.

3. Results

3.1 Keyword frequency analysis

Table 1 shows the 19 words with TF $\geq$ 10 in the keywords of the target articles and their respective TF-IDF values. The TF of the livestock species was the highest for “canine (65)”, and this was followed by “swine (43)”, “feline (25)”, “avian (22)”, “bovine (20)”, and “equine (15)”. The TF of the other items was the highest for “virus (51)”, and this was followed by “cell (23)”, “disease (21)”, “fever (19)”, and “analysis (17)”.

Table 1. Term frequency and term frequency-inverse document frequency of keywords

Rank	Keyword	TF	TF-IDF
1	canine	65	97.65
2	virus	51	96.10
3	swine	43	86.34
4	feline	25	62.12
5	cell	23	61.87
6	avian	22	61.68
7	disease	21	55.64
8	bovine	20	54.02
9	fever	19	53.43
10	analysis	17	50.14
11	equine	15	47.81
12	protein	13	40.05
13	gene	12	38.97
13	molecular	12	38.97
13	PCR	12	38.97
13	African	12	37.93
17	resistance	11	36.77
17	stem	11	36.77
19	phylogenetic	10	34.68

3.2 N-gram analysis

As a way to extract N consecutive elements from a string, this study analyzed bigrams (two consecutive words) and trigrams (three consecutive words), as shown in Table 2. For the bigrams, “swine” occurred 16 times with “-fever,” 12 times with “-African,” and five times with “-circovirus,” whereas “stem” occurred 11 times with “-cell” and five times with “-mesenchymal.” For the trigrams, “African-swine-fever” occurred 12 times, “swine-fever-virus” eight times, “magnetic-resonance-imaging” five times, and “mesenchymal-stem-cell” five times, consecutively.

Table 2-1. Bigrams of keywords

Rank	Keyword 1	Keyword 2	N
1	swine	fever	16
2	African	swine	12
3	stem	cell	11
4	fever	virus	9
5	phylogenetic	analysis	6
5	staphylococcus	aureus	6
7	immune	response	5
7	feline	calicivirus	5
7	magnetic	resonance	5
7	resonance	imaging	5
7	swine	circovirus	5
7	case	report	5
7	real	time	5
7	mesenchymal	stem	5
7	disease	virus	5

Table 2-2. Tigrams of keywords

Rank	Keyword 1	Keyword 2	Keyword 3	N
1	African	swine	fever	12
2	swine	fever	virus	8
3	magnetic	resonance	imaging	5
3	mesenchymal	stem	cell	5
5	swine	reproductive	respiratory	4
5	reproductive	respiratory	syndrome	4
5	classical	swine	fever	4
5	real	time	PCR	4
9	respiratory	syndrome	virus	3
9	enzyme	immunosorbent	assay	3
9	staphylococcus	aureus	nanogel	3
9	case	report	feline	3
9	ct	scan	canine	3

3.3 Centrality analysis results

The centrality analysis analyzed the centrality of connections, which is an index of the degree to which a node is connected to other nodes, as shown in Table 3. The centrality was the highest for “disease (0.79),” and this was followed by “analysis (0.74)” and “virus (0.74).” The centrality of the livestock species terms was the highest for “swine (0.53),” and this was followed by “canine (0.47),” “bovine (0.47),” “feline (0.47),” and “avian (0.32).”

Table 3. Centrality index of keywords

Rank	Keyword	Degree centrality	Rank	Keyword	Degree centrality
1	disease	0.79	5	phylogenetic	0.47
2	analysis	0.74	11	cell	0.42
3	virus	0.74	11	molecular	0.42
4	swine	0.53	11	protein	0.42
5	canine	0.47	14	African	0.32
5	bovine	0.47	14	avian	0.32
5	feline	0.47	14	gene	0.32
5	fever	0.47	17	resistance	0.21
5	PCR	0.47	17	stem	0.21

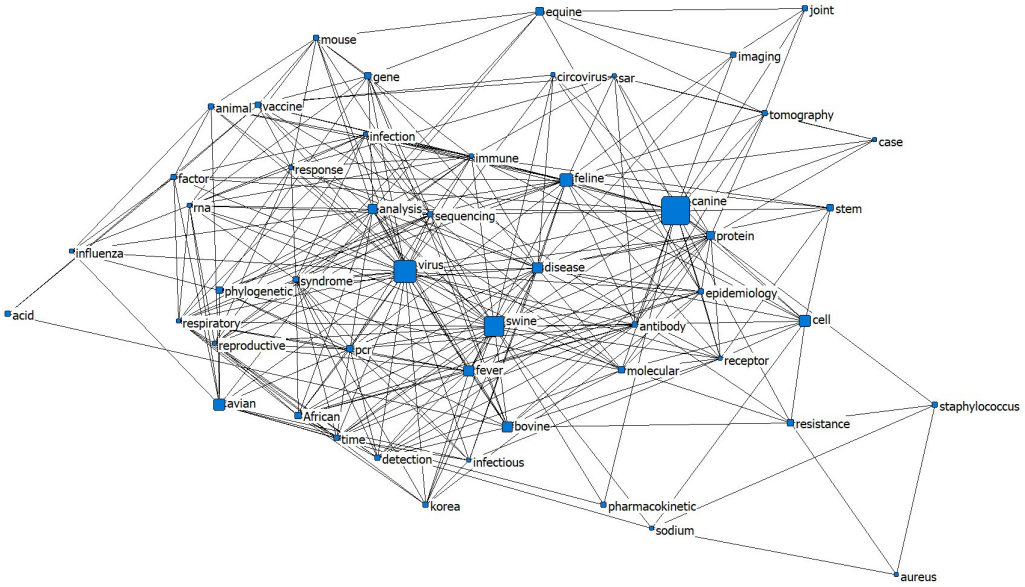

Fig. 1. Visualization results of the centrality reflecting TF

3.4 Co-occurrence keyword analysis

Table 4. demonstrates the results of arranging keywords with a frequency of co-occurrence exceeding 7 times. The highest frequency of co-occurrence was ‘swine’, which co-occurred 23 times with ‘-fever’, 21 times with ‘-virus’, and 17 times with ‘-African’. The keywords with the frequency of co-occurrence at 15 times were “cell-stem,” “fever-African,” “virus-avian,” and “virus-fever.”

Table 4. Co-occurrence of keywords

Keyword 1	Keyword 2	Weight (TF)
swine	fever	23
virus	swine	21
swine	African	17
cell	stem	15
fever	African	15
virus	avian	15
virus	fever	15
virus	PCR	11
swine	PCR	10
canine	cell	9
equine	joint	8
avian	influenza	7
virus	African	7
virus	syndrome	7
virus	infectious	7

4. Discussion

A keyword is a word or phrase that summarizes the essential content within a body or title that describes its content. Therefore, these keywords serve as keys with which to easily find the desired information when searching for related studies (Lee et al., 2017). Therefore, this study identified research trends in articles published in J Vet Sci from 2020 to 2022 based on SNA and targeted the keywords presented therein.

The analysis of TF and TF-IDF values allowed for the extraction of keywords by analyzing the frequencies of occurrence and weights of keywords as well as the selection of important words. After checking for words with TF $\geq$ 10, the TF-IDF was found to be proportional to the TF. Moreover, “canine” had the highest TF among the livestock species terms, and this was followed by “swine.” Among the other terms, “virus” had the highest TF, and this was followed by “cell” and “disease”, indicating that research related to these topics is being conducted.

The N-gram analysis allows for the extraction of themes by checking and connecting the relationships of words that are related to each other through N consecutive elements. The bigram and trigram analyses revealed that “African swine fever” and “stem cells” were the most popular research topics. Other key research topics included “phylogenetic analysis, Staphylococcus aureus, immune response, feline calicivirus, magnetic resonance imaging, swine circovirus, case report, mesenchymal stem cell, real time polymerase chain reaction (PCR), and disease virus”.

The centrality analysis shows the influences between keywords that play important roles in the network. The centrality analysis considers the degree, betweenness, closeness, and Eigenvector centralities (Hwang & Kim, 2019). This study analyzed the centralities of connections for major keywords and found that “disease,”, “analysis,”, and “virus” had the highest centrality values, whereas “swine” had the highest centrality value among the livestock species terms. The degree centrality index presented in this study evaluates the index as the number of all edges that are directly connected to a node. Furthermore, each relationship between nodes is analyzed via classification into in-degree and out-degree, according to the direction of the relationship. In-degree is a measure of the popularity of a node, and out-degree is a measure of the impact of a node. Thus, while “canine” and “virus” had higher frequencies of occurrence, the centrality analysis revealed that research on the topics of “swine” and “disease” had a high level of interest and impact.

The results of the co-occurrence keyword analysis showed in which context two or more words were used together. The research topics of “African-swine-fever” and “stem-cells” showed a high frequency of co-occurrence. There were 15 cases of co-occurrences of “avian” and “virus”, further confirming that virus research is a major topic in the field of avian research.

This study utilized big data analysis methods, such as TF, TF-IDF, N-gram, centrality analysis, and co-occurrence keyword analysis, which are rarely used in veterinary research, to identify the research trends of articles and suggest key research topics. Starting with this study, in future research, a variety of domestic and foreign articles should be selected for analysis, and the time period should be expanded to increase the value of the analytical results and provide results that can be used as the main statistical data for veterinary research.

5. Conclusion

This study used SNA to analyze 283 academic articles published in the English journal *(J Vet Sci*) of KSVS from 2020 to 2022. The conclusions of the study follow. When TF-IDF and TF were combined, the top 10 research keywords were “canine”, “swine”, “feline”, “avian”, and “bovine” for livestock species and “virus”, “cell”, “disease”, “fever”, and “analysis” for non-livestock species. The results of an N-gram analysis identified “African swine fever” and “stem cells” as the top research topics. Other key research topics included “phylogenetic analysis, *Staphylococcus aureus*, immune response, feline calicivirus, magnetic resonance imaging, swine circovirus, case report, mesenchymal stem cell, real time PCR, and disease virus” In addition, the centrality analysis showed that “disease”, “analysis”, and “virus” ranked highest in terms of research interest and impact, whereas “swine” ranked highest in terms of livestock species.

In conclusion, the results of this study are expected to provide evidence for identifying research trends in veterinary science and provide useful information for future research.

Notes

ORCID

Jae-Hee Roh : https://orcid.org/0000-0003-4223-1096

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

Choi, J. H., Park, J. K., & Kim, M. Y. (2021). Analysis of research trends related to diagnosis of ASD through keyword network analysis: Focusing on domestic academic journals published from 2011-2020. *Journal of Behavior Analysis and Support, 8* (1), 115-135.

Hwang, S., & Hwang, D. R. (2018). A study on the research trends in arts management in Korea using topic modeling and semantic network analysis. *Journal of Arts Management and Policy, 47*, 5-29.

Hwang, S., & Kim, M. K. (2019). An analysis of artificial intelligence (AI)-related studies’ trends in Korea focused on topic modeling and semantic network analysis. *Journal of Digital Contents Society, 20* (9), 1847-1855.

Hwang, S., & Shim, J. W. (2020). Semantic network analysis of “smart city” in newspaper articles-from 2016 to 2019. *Journal of Digital Contents Society, 21* (5), 941-950.

Hwang, S. I., & Park, Y. W. (2019). An analysis of arts management-related studies’ trend in Korea using topic modeling and semantic network analysis. *Journal of Applied Mathematics*

and Physics, 50, 5-31.

Jang, J. H., Won, B. Y., Jang, G. W., Kim, S. A., Oh, S. H., Kim, Y. J., Ha, M. O., Lee, Y. E., Lee, G. R., Park, S. S., & Eom, S. (2017). Trend analysis of research in the journal of Korean society dental hygiene from 2001 to 2015. *Journal of Korean Society of Dental Hygiene, 17* (4), 693-704.

Jung, H., & Lee, B. G. (2020). Research trends in text mining: Semantic network and main path analysis of selected journals. *Expert Systems with Applications, 162*, 113851.

Kang, J. W., & Namkung, Y. (2021). Understanding consumers' perceptions of the fresh-food delivery platform service based on big data: Using text mining and semantic network analysis. *Korean Journal of Hospitality and Tourism, 30* (2), 37-52.

Kang, S. J., Jeong, H. Y., & Lee, Y. S. (2018). A semantic network analysis on parents' perception of children's play space: Focusing on playground and kids café. *Journal of Early Childhood Education Research, 38* (2), 281-304.

KCI. (2022, November 6). *KCI portal: Article search*. Retrieved from http://www.kci.go.kr/kci-portal/po/search/poArtiSear.kci

KCI. (2024, June 25). *KCI portal*. Retrieved from http://www.kci.go.kr/kciportal/po/search/poTotalSearList.kci.

Kim, K. S., Jang, S. G., & Lee, K. S. (2021). A network analysis of research topics and trends in end-of-life care and nursing. *International Journal of Environmental Research and Public Health, 18* (1), 313.

Kim, Y. J. (2018). Comparison of author keywords and medical subject heading terms in the Journal of Korean Society of Dental Hygiene from 2001 to 2015. *Journal of Korean Society of Dental Hygiene, 18* (6), 1047-1055.

KSVS. (2024, June 25). *Journal of KSVS*. Retrieved from http://www.ksvs.or.kr/journal/journal.do.

Lee, S., Kim, S. S., & Chae, D. H. (2017). Analysis of qualitative research on science education trend in Korea using semantic network analysis. *Journal of Korean Society of Earth Science Education, 10* (3), 290-307.

Lee, S. S. (2014). A content analysis of journal articles using the language network analysis methods. *Journal of the Korean Society for Information Management, 31* (4), 49-68.

Park, H. W., & Leydesdorff, L. (2004). Understanding the KrKwic: A computer program for the analysis of Korean text. *Journal of the Korean Data Analysis Society, 6* (5), 1377-1387.

Reardon, S. (2014). Text-mining offers clues to success. *Nature, 509* (7501), 410.

Scopus Preview. (2024, November 21). *Scopus portal*. Retrieved from https://www.scopus.com/sourceid/18327.