Tourism Management Perspectives 31 (2019) 332–339

Contents lists available at ScienceDirect

Tourism Management Perspectives

journal homepage: www.elsevier.com/locate/tmp

Network text analysis of medical tourism in newspapers using text mining: The South Korea case☆

Sohyeon Kim[a], Won Seok Lee[b,*]

[a] Department of Leisure and tourism sciences, Kyonggi University, Room 403, Choong-Jung-Gwan, 24, Kyonggidae-ro 9-gil, Seodaemun-gu, Seoul, Republic of Korea
[b] Department of Tourism & Recreation, Kyonggi University, Room 403, Choong-Jung-Gwan, 24, Kyonggidae-ro 9-gil, Seodaemun-gu, Seoul, Republic of Korea

ARTICLE INFO

Keywords:
Text mining
Network text analysis
Medical tourism
Newspaper articles

ABSTRACT

The main purpose of this study was to investigate keywords related to medical tourism in daily and medical newspapers and to analyze networks between keywords. We collected newspaper articles by running a web crawler with Python. A total of 3802 articles from 4 medical newspapers and 1705 articles from 5 daily newspapers from 2009 to 2017 were reviewed. The clusters derived from daily newspapers focus mainly on medical tourists, medical care and service, regional vitalization, and field support. In contrast, the clusters derived from medical newspapers focus mainly on medical tourism marketing, introducing medical treatment overseas, the medical community's voice, and medical tourism vitalization with support from a government. We found that the content varied depending on the characteristics of the newspapers. This information is necessary to understand various perspectives on policymaking. The academic contributions and practical implications of the findings are discussed.

1. Introduction

Medical tourism is growing rapidly and has the potential to contribute both economically and to the tourism industry (Hudson & Li, 2012; Noree, Hanefeld, & Smith, 2016; Shetty, 2010). Medical tourism is expected to continue to grow more as healthcare costs increase in developed countries and as the privatization of healthcare spreads. Patients in developed countries are unable to afford expensive domestic medical care, and they often travel to developing countries to receive medical services (Connell, 2006; Fetscherin & Stephano, 2016). Insurance companies even encourage patients to embark on medical tours abroad to reduce costs (Álvarez, Chanda, & Smith, 2011; Connell, 2006; Fetscherin & Stephano, 2016). Numerous countries, including Korea, are making enormous efforts to promote medical tourism because the economic ripple effects of medical tourism in Korea are receiving increased attention (Kang, 2015). South Korea is a latecomer to medical tourism compared with other countries, but it is attempting to be fast follower. Government-led policy and the aggressive marketing of stakeholders (Korea Tourism Organization, 2016) support this goal. The Medical Service Act was revised in 2009 to promote medical tourism, and it is now possible to attract foreign patients to domestic medical institutions for commercial profit. Regulations on medical tourism have been eased, including the development of the hotel business 'Meditel' for medical tourists. A project was launched to establish a medical tourism cluster and to provide economic support to local governments from 2014 to 2017.

Considerable scholarly work has been conducted on how to promote medical tourism (Cho, 2006; Cho, 2015; Lee & Chung, 2007; Oh & Jeon, 2017), the quality of medical services and medical tourists' satisfaction (Lee & Lee, 2010; Yoon, Han, & Kim, 2009), and the attributes considered in the selection of medical tourism destinations (Kwak, 2011; Lee, Kim, & Lee, 2009). However, relatively few studies have been conducted on social perceptions of medical tourism, and little work has focused on both the medical community and public perspectives. Furthermore, there are a variety of stakeholders in medical tourism (Álvarez et al., 2011). This variety creates conflict in defining medical tourism and corresponding authorities or responsibilities in medical tourism-related policy among competent government departments, such as the Ministry of Health and Welfare and the Ministry of Culture, Sports and Tourism. The risk of conflicts of interest among ministries should be managed (House of Commons Public Administration and Constitutional Affairs Committee (PACAC), 2017). Therefore, it is necessary to study not only medical tourists but also various perceptions to establish a desirable direction for medical tourism policy.

In contrast to the use of traditional methods to identify perceptions and trends in society, big data analyses have become popular in recent

☆ This work was supported by Kyonggi University Research Grant 2017.
* Corresponding author.
E-mail address: lws798@kyonggi.ac.kr (W.S. Lee).

https://doi.org/10.1016/j.tmp.2019.05.010
Received 12 June 2018; Received in revised form 21 May 2019; Accepted 24 May 2019

years (LaValle, Lesser, Shockley, Hopkins, & Kruschwitz, 2011; Mendenhall, Brown, & Black, 2017; Walker, 2014). Text mining is drawing attention due to the increase in unstructured data such as blogs, tweets, and customer reviews (Chakraborty & Pagolu, 2014). Text mining is a technique for finding interesting information based on data mining, machine learning, natural language processing, information retrieval, and knowledge management (Feldman & Sanger, 2007). Because text mining can extract meaningful information from a large volume of text, the number of studies analyzing large quantities of text data crawled from newspapers or social networking sites (SNS), such as Facebook and Twitter, has increased in recent years. In particular, newspaper articles are published materials that reflect the author's preference for a specific object, which can help to identify social preferences (Choi, Lee, & Kwon, 2016). In other words, the opinions of various stakeholders on a particular subject can appear in newspaper articles. Moreover, articles offer deep insight because they are written even before events occur, while most SNS data are likely to be produced after events occur (Choi et al., 2016). In this study, network text analysis (NTA) was used to analyze newspaper articles. NTA is a variant of social network analysis, a method that identifies meaningful information by deriving the link structure between words from large volumes of text (Diesner & Carley, 2005). Text data were collected by crawling daily newspapers that include public perspectives and medical newspapers that include the perspectives of the medical community to identify various perspectives.

Medical tourists are a mixture of patients who travel solely to receive medical services due to serious problems and tourists who enjoy shopping, sightseeing, and cultural experiences but who also receive treatments such as cosmetic surgery (Korea Tourism Organization, 2016; Noree et al., 2016). Revealing various perspectives can help to identify which group is being targeted. By detailing the views of the public and medical community, the results of this study hold practical implications for aiding policy decisions on medical tourism.

2. Literature review

2.1. Medical tourism

Medical tourism is the outsourcing of medical care (Bies & Zacharia, 2007; Leahy, 2008) and a rising niche market for developing countries (Connell, 2006; Lautier, 2008). Formerly, the flow of medical tourism was from developing countries to developed countries (Lautier, 2008). However, in conjunction with the advancement of medical infrastructures in developing countries and the rising burden of receiving medical services for individuals in developed countries, the trend has reversed (Connell, 2006; Lautier, 2008; Turner, 2007). Due to high costs and long waiting times, patients in developed countries consider traveling abroad for medical treatment (Horowitz, 2007; Lautier, 2008). In addition, certain medical procedures, such as cosmetic surgery or dentistry, are not covered by insurance, making medical tourism a viable option for them (Bies & Zacharia, 2007; Connell, 2006; Garcia-Altes, 2005).

'Medical tourism' is easily confused with health tourism and medical travel (Connell, 2013). Sometimes, the phrase is used to refer to travel by medical staffing to other countries to offer medical treatment, but such travel is more akin to medical travel (Edelheit, 2008). A generally accepted definition for medical tourism is travel by patients overseas to receive medical care (Connell, 2006). This definition divides medical tourism into two main categories. One is defined as only travel from one country to another for medical treatment, with no mention of tourism activities (e.g., Horowitz, 2007; Lautier, 2008; Turner, 2007). Because not all patients enjoy tourism activities, medical experts do not perceive tourism to be a significant part of medical tourism (Jagyasi, 2008). However, other researchers have taken an interesting approach to the tourism sector in medical tourism (e.g., Connell, 2006; Connell, 2013). These authors argue that it is difficult not to address tourism in medical tourism because hotels, airlines, and food that medical tourists experience when they travel abroad are all part of tourism (Connell, 2006; Edelheit, 2008). The reason for the conflicting definitions of medical tourism is that there are a variety of stakeholders in the industry, who have different views on medical tourism and define it by considering only their own field (Álvarez et al., 2011). This difference could lead to confusion among the branches of the government supporting medical tourism. In Korea, the Ministry of Health and Welfare and the Ministry of Culture, Sports and Tourism are competent authorities of medical tourism. There is a conflict regarding authority and responsibility among the ministries; therefore, it has been noted that relevant policy has not been actively promoted (Kim, 2017a; Ko, 2012). For that reason, it is necessary to identify the areas in which each stakeholder is interested so that efficient resource distribution can be possible in policymaking.

The Internet is a source of information about public institutions and a medium that can influence stakeholders (Meijer, 2007). In particular, newspaper articles can be accessed from the Internet easily and reflect the author's preference on a certain subject (Choi et al., 2016). In other words, the opinions of various stakeholders on a particular subject can appear in newspaper articles and exert a secondary influence. Thus, to analyze the point of views on medical tourism, it is necessary to review content from newspaper articles. In addition, to explore the differences in content between stakeholders, newspapers should be classified into different types. News could not be written without considering the target audiences and appealing to them (Richardson, 2006). Accordingly, the discourse in articles could vary among different types of newspapers. As described above, the medical community tends not to consider the tourism sector; therefore, it is necessary to examine differences from the public views of medical tourism.

2.2. Text mining

Text mining (or text data mining) is the process by which new knowledge or patterns are found from a large quantity of text data by computer (Hearst, 1999; Hearst, 2003). Text mining is different from general data mining in that the former addresses unstructured data (such as email, newspaper articles, text documents) rather than structured data (Fan, Wallace, Rich, & Zhang, 2006). The text mining process is illustrated in Fig. 1 (Fan et al., 2006). When a set of documents is gathered, the required data are filtered out and preprocessed. In the preprocessing step, a string is split into several tokens, useless words are removed, and stemming can be performed. Then, text analysis, such as sentiment analysis, content analysis, keyword analysis, topic analysis or network analysis, can be used to discern interesting information from the data (Godnov & Redek, 2016).

Network analysis is one of the techniques used in text analysis. Network text analysis (NTA) was originally derived from social network analysis (SNA). SNA, starting from graph theory, is a method that focuses on structures and connections between nodes (Wasserman & Faust, 1994). The structure of networks consists of nodes and ties that connect nodes. Based on this concept, NTA constructs networks

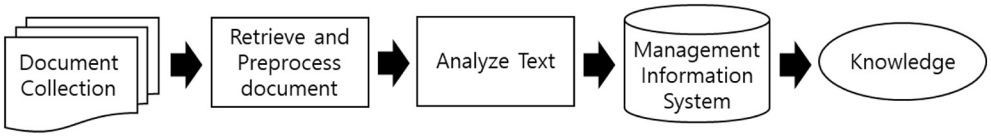

Fig. 1. Text mining process (Fan et al., 2006).

between words from textual data (Diesner & Carley, 2005). NTA is known as an effective method for analyzing a large number of texts and visualizing relationships, which can make it easy to obtain insight from the data (Diesner & Carley, 2005; Paranyushkin, 2010). In addition, NTA not only reveals the word that is most frequently mentioned but also how it is arranged (Paranyushkin, 2010). By analyzing how words are clustered in a network, NTA can discover that the same word may mean different things when grouped into different clusters (Park & Jung, 2013).

One of the indicators that can be found through the network structure derived from network analysis is centrality (Wasserman & Faust, 1994). Centrality refers to the influence of a word within a network structure. Degree centrality is a measure of the number of links nodes have in a network. If a node has connections with many other nodes, the node's degree centrality is high. Thus, a word with a high score plays an important role in a network. While degree centrality measures the number of nodes connected to a node, eigenvector centrality weights the degree centrality of connected nodes (Newman, 2008). A connection to a single node with high degree centrality has a higher eigenvector centrality than does a link to multiple nodes with low scores.

3. Methods

3.1. Research subjects and scope

The purpose of the study was to analyze newspaper articles on medical tourism by applying network text analysis to find connections between terms and explore perceptions of each stakeholder. The newspaper articles collected after searching for the phrase 'medical tourism' in the predefined newspaper websites were used as analysis sources. The time frame for the search was set from 1 January 2009 to 31 December 2017. The reason for choosing that time span is that the government started to actively support the medical tourism industry to attract foreign patients through the revision of the Medical Service Act in 2009 (Kim, Song, & Noh, 2012). The study was conducted in the following order: 1) collection of related articles and extraction of the body of articles, 2) preprocessing of the text, and 3) creation of a co-occurrence matrix between words and network analysis.

3.2. Data collection and analysis

To collect data, we employed Python, a computer programming language, and the 'BeautifulSoup' (Richardson, 2017) library. By running a web crawling code written in Python, we collected all articles that appeared after searching for 'medical tourism' on each newspaper's website.

We collected articles on medical tourism from *The Korean Doctors' Weekly*, *Dailymedi*, *Korea Medical News* and *Doctors News*. Other medical newspapers are published in Korea, but we chose these publications because they have been cited in medical tourism-related research and selected as preferred magazines for medical and welfare professionals (Chun & Lee, 2010; Kim, 2011; Lee, 2016; Lee & Park, 2016) and are preferred by health and welfare experts (Kham, 2010). A total of 4735 articles were collected, with 752 articles in *The Korean Doctors' Weekly*, 1463 articles in *Dailymedi*, 1611 articles in *Korea Medical News*, and 909 articles in *Doctors News*. We excluded articles that did not have the exact phrase 'medical tourism' in their content. As a result, a total of 86 articles were excluded: 33 articles in *The Korean Doctors' Weekly*, 11 articles in *Dailymedi*, 37 articles in *Korea Medical News*, and 5 articles in *Doctors News*. Research on newspaper articles has found that articles with fewer than 500 characters do not contain the necessary content and that their removal would not affect the study results (Ha & Lee, 2012; Kim, 2017b); thus, articles with fewer than 500 characters were excluded. Therefore, 589 articles in *The Korean Doctors' Weekly*, 1114 articles in *Dailymedi*, 1291 articles in *Korea Medical News*, and 808 articles in *Doctors News* were analyzed, for a total of 3802 articles.

With regard to daily newspapers, *Chosun Ilbo*, *Joongang Ilbo*, *Dong-A Ilbo*, *Hankyoreh*, and *Kyunghyang Shinmun* were selected based on the circulation of the newspapers in 2016 (Korea Audit Bureau of Certification, 2017). With 1618 articles in *Chosun Ilbo*, 1829 articles in *Joongang Ilbo*, 497 articles in *Dong-A Ilbo*, 152 articles in *Hankyoreh*, and 437 articles in *Kyunghyang Shinmun*, 4533 articles were initially collected. Articles were removed under the condition for removal indicated above. Therefore, a total of 1705 articles were analyzed: 394 articles in *Chosun Ilbo*, 421 articles in *Joongang Ilbo*, 441 articles in *Dong-A Ilbo*, 101 articles in *Hankyoreh*, and 348 articles in *Kyunghyang Shinmun*.

After excluding data, words are extracted through stop word removal, stemming, morphological analysis, and POS (part of speech) tagging using the R package 'KoNLP' (Korean Natural Language Processing) (Jeon & Kim, 2016), a Korean morphological analyzer. Because there is no Korean stop words list built in R, custom lists were used by referring to the English stop words list (e.g., over, we, because, and during). Numbers and punctuation were also removed. Using a morphological analyzer, sentences were divided into articles, and common nouns were extracted. As a result, 64,579 words and 52,100 words were extracted from medical newspaper and daily newspaper articles, respectively. Then, the Korean words were translated into English. During the translation process, the words did not appear to have been morphologically analyzed; however, all of the words underwent stemming and POS tagging in Korean. Furthermore, the word 'attract' means to attract tourists or customers in this paper. We created a term document matrix (TDM) using the R package 'tm' (Feinerer & Hornik, 2015). The top 100 keywords were extracted to create a co-occurrence matrix for network analysis. The co-occurrence matrix was created by multiplying the TDM with the TDM transposed matrix.

For network analysis, the relationships between keywords were quantified by analyzing the degree centrality and the eigenvector centrality using UCINET (Borgatti, Everett, & Freeman, 2002) and NetDraw (Borgatti, 2002). Next, a CONCOR (convergence of iterated correlations) analysis was performed using UCINET to derive clusters formed by similar words. CONCOR analysis is a common technique used in structural equivalence analysis, which analyzes the Pearson correlation of the co-occurrence matrix between keywords and then builds blocks of nodes to identify relationships among blocks (Kim & Kim, 2017).

4. Results

4.1. Eigenvector centrality of keywords

No significant differences in the top 30 keywords were observed based on degree centrality, eigenvector centrality or rank between daily and medical newspapers. Therefore, only eigenvector centrality is shown in Table 1 to compare keywords between daily and medical newspapers. Keywords that did not appear as the top 30 keywords in other newspapers are written in bold. Keywords with a gap of more than five ranks between newspapers are underlined.

Keywords such as 'Medical Tourism', 'Medical Care', 'Foreign Country', 'Hospital', 'Patient', 'Korea', 'China', 'Attract', 'Domestic', 'Scheme', 'Overseas', and 'Government' ranked in the top 30 among both daily and medical newspapers. Some keywords ranked highly in medical newspapers, but their ranks were relatively low in the daily newspapers and vice versa. 'Support' (0.121) was ranked 15th in medical newspapers, while it was ranked 29th (0.105) in daily newspapers. 'China' (0.139), 'Region' (0.124), 'Management' (0.123), and 'Service' (0.120) were ranked 10th, 15th, 16th, and 18th in daily newspapers, respectively, while they were ranked 29th, 23rd, 30th, and 28th in medical newspapers. In medical newspapers only, the keywords ranked in the top 30 were as follows: 'Medical Institution', 'Vitalization', 'Diagnosis', 'Overseas Patient', 'Progress', 'Target', 'Determined', 'Build', and 'Medical Service'. In daily newspapers only,

Table 1
Eigenvector centrality of top 30 keywords in newspapers.

Rank	Daily newspapers		Medical newspapers	
1	*Medical tourism*	0.273	*Medical tourism*	0.287
2	*Medical care*	0.205	*Patient*	0.230
3	*Foreign country*	0.196	*Attract*	0.217
4	*Hospital*	0.195	*Hospital*	0.205
5	*Attract*	0.191	*Medical care*	0.178
6	*Patient*	0.189	***Medical institution***	0.171
7	*Domestic*	0.178	*Foreign country*	0.170
8	*Korea*	0.152	*Domestic*	0.164
9	*Scheme*	0.15	***Vitalization***	0.146
10	*China*	0.139	*Scheme*	0.139
11	*Overseas*	0.139	*Overseas*	0.132
12	***Field***	0.135	*Government*	0.125
13	*Government*	0.134	*Schedule*	0.123
14	***Tourism***	0.132	***Diagnosis***	0.123
15	*Region*	0.124	*Support*	0.121
16	*Management*	0.123	***Overseas patient***	0.121
17	***Level***	0.121	*Korea*	0.120
18	*Service*	0.120	***Progress***	0.116
19	***Development***	0.116	***Target***	0.116
20	*Provide*	0.116	***Determined***	0.115
21	***Treatment***	0.115	*Push-ahead*	0.112
22	*Push-ahead*	0.114	*Related*	0.111
23	*Related*	0.113	*Region*	0.108
24	***Variety***	0.113	*Provide*	0.107
25	*Schedule*	0.112	*Need*	0.107
26	***Tourist***	0.109	***Build***	0.105
27	***Business***	0.107	***Medical service***	0.104
28	*Need*	0.107	*Service*	0.101
29	*Support*	0.105	*China*	0.101
30	***World***	0.103	*Management*	0.101

Keywords that appeared as the top 30 keywords only in daily newspapers (or medical newspapers) are written in bold. Keywords with a gap of more than five ranks between daily newspapers and medical newspapers are underlined.

the keywords ranked in the top 30 were as follows: 'Field', 'Tourism', 'Level', 'Development', 'Treatment', 'Variety', 'Tourist', 'Business', and 'World'.

4.2. CONCOR analysis results

CONCOR analysis was performed to visually derive connections and patterns within the network and to cluster keywords with similarities. As a result, eight clusters were grouped in daily newspapers and medical newspapers. Then, the keywords were regrouped into four groups to find the related cluster. Visualizations are shown in Figs. 2 and 3. Additional visualizations with the top 100 links in each cluster are shown in Figs. 4 and 5.

Keyword clusters in daily newspaper articles related to medical tourism were as follows, in order of ascending eigenvector score

- 1–1. Attracting medical tourists: Medical tourism, Foreign country, Attract, Target, Seoul, Medical tourist, Visit, Local, Russia, Public relations, Plastic surgery
- 1–2. Chinese Tourist Market: China, Tourism, Tourist, Market, The Person Interested, Chinese, Product, Hotel, Increase, Whole
- 2–1. Medical Care: Medical Care, Hospital, Patient, Overseas, Treatment, Diagnosis, Medical Institution, Doctor, Surgery, Specialty, Medical Service, Ministry of Health and Welfare, Medical Team
- 2–2. Service Level and Attention of Nation: Domestic, Korea, Government, Provide, Level, Service, Need, World, USA, Problem, Result, Start, Person, Nation, Explanation, Professor, Attention, Reason, Possibility
- 3–1. Regional Vitalization: Scheme, Region, Schedule, Push Ahead, Development, Business, Vitalization, Central, Create, Daegu, Selection, Build, Plan, Nationwide
- 3–2. Program and Facility Management: Management, Scale, Progress, Arrange, Determined, Program, Facility, Representative, Maximum, Participation, Prospect, Composition
- 4–1. Related Training and Growth: Field, Related, Variety, Competitiveness, Training, Growth, Global, Best, Role, Goal, International
- 4–2. Support Industry and Enterprise: Support, Industry, Enterprise, Advancement, Effect, Expansion, Establish, Investment, Promote, Content

Keyword clusters in medical newspapers articles related to medical tourism were as follows

- A-1. Attracting medical tourist and medical care: Medical tourism, Attract, Medical care, Foreign country, Overseas, Target, Provide, Build, Service, Variety, World, Medical tourist, Global, Medical technology, Excellent, Selection, Program, Continuous, Specialty, Tourism
- A-2. Marketing and PR: Scheme, Schedule, Progress, Region, China, The Person Interested, Take Place, Public Relations, Contract, Marketing, Activity, Busan, Composition, Cooperation, Representative
- B-1. Attention and Introduction to Treatment: Patient, Hospital, Diagnosis, Korea, Treatment, Attention, Director, USA, Professor, Start, Introduce, System
- B-2. Medical Travel by Medical Team: Visit, Russia, Local, Medical Team, Surgery, Director of Hospital
- C-1. Medical Institution Vitalization and Support by Ministry: Medical Institution, Vitalization, Support, Overseas Patient, Push Ahead, Related, Management, Arrange, Business, Expansion, Central, Development, Participation, Ministry of Health and Welfare, Plan, Reinforcement, Seoul, Foreign Patient, Establish, Industry
- C-2. Service Level and Role of Nation: Domestic, Determined, Medical Service, Level, Advancement, Competitiveness, Market, Prospect, Nation, Scale, Role, Main, Increase
- D-1. Problem Pointed Out by the Medical Community: Government, Need, Emphasis, Problem, Medical Community, Situation, Point Out
- D-2. Doctor's Effort to an Explanation: Doctor, Explanation, My Country, Effort, Chairperson, Result

5. Conclusions and implications

The purpose of the study was to understand the perspectives of the public and the medical community on medical tourism using text mining. Instead of using traditional methods such as surveys or interviews, medical and daily newspaper articles were collected by web crawling, and keywords related to medical tourism were extracted. Through network analysis and CONCOR analysis, the study extracted not only important keywords but also clusters of keywords.

More than half of the words overlapped in the top 100 keywords of the daily and medical newspapers. However, even if the same word is tied to another cluster, it can be used in a different sense (Park & Jung, 2013). 'Medical Tourism', 'Medical Care', 'Foreign Country', 'Hospital', 'Patient', 'Domestic', and 'Overseas' were ranked in the top 30 in both daily newspapers and medical newspapers. These words can be considered to be related to the definition of medical tourism.

In daily newspapers, 'China', 'Service', 'Region' and 'Management' were ranked more highly than in medical newspapers. The clusters derived from CONCOR analysis show similar results. First, clusters 1–1 and 1–2 are related to medical tourists. Clusters 1–1 indicate attraction of medical tourists. Cluster 1–2 indicate the Chinese tourist market. These results suggest that the targeted medical tourist market and its main interests, such as plastic surgery or hotels, are addressed in daily newspapers. Second, clusters 2–1 and 2–2 are related to medical care and service. Keywords concerning medical care, such as 'Treatment',

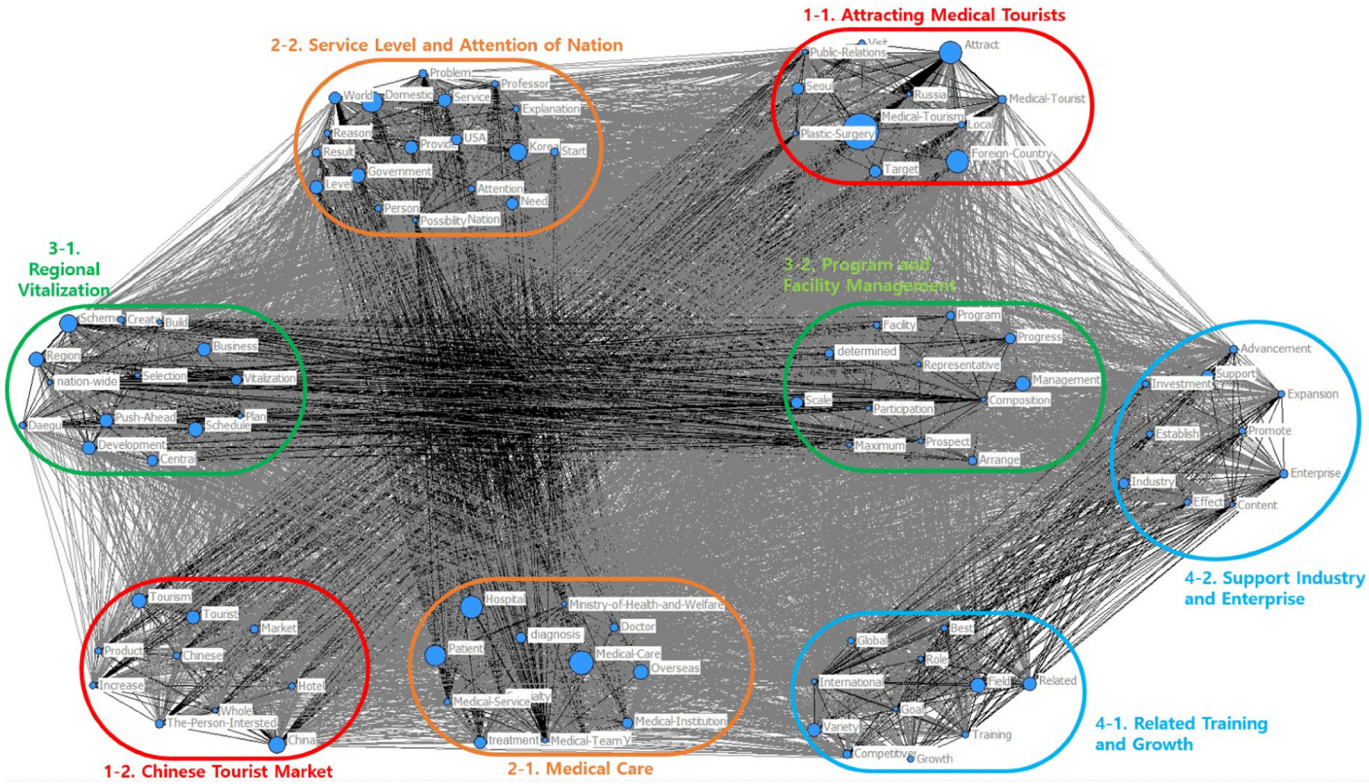

Fig. 2. Daily newspaper CONCOR analysis visualization.

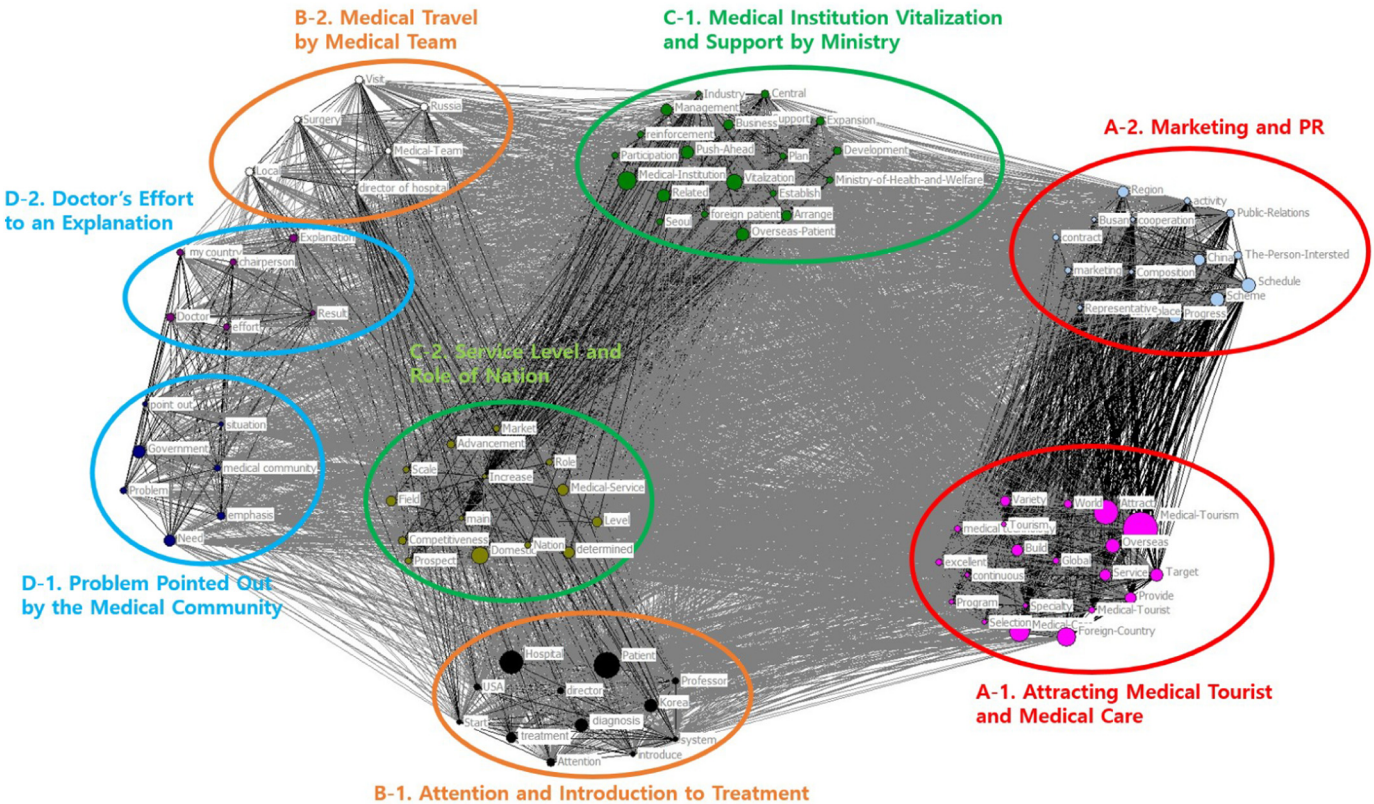

Fig. 3. Medical newspaper CONCOR analysis visualization.

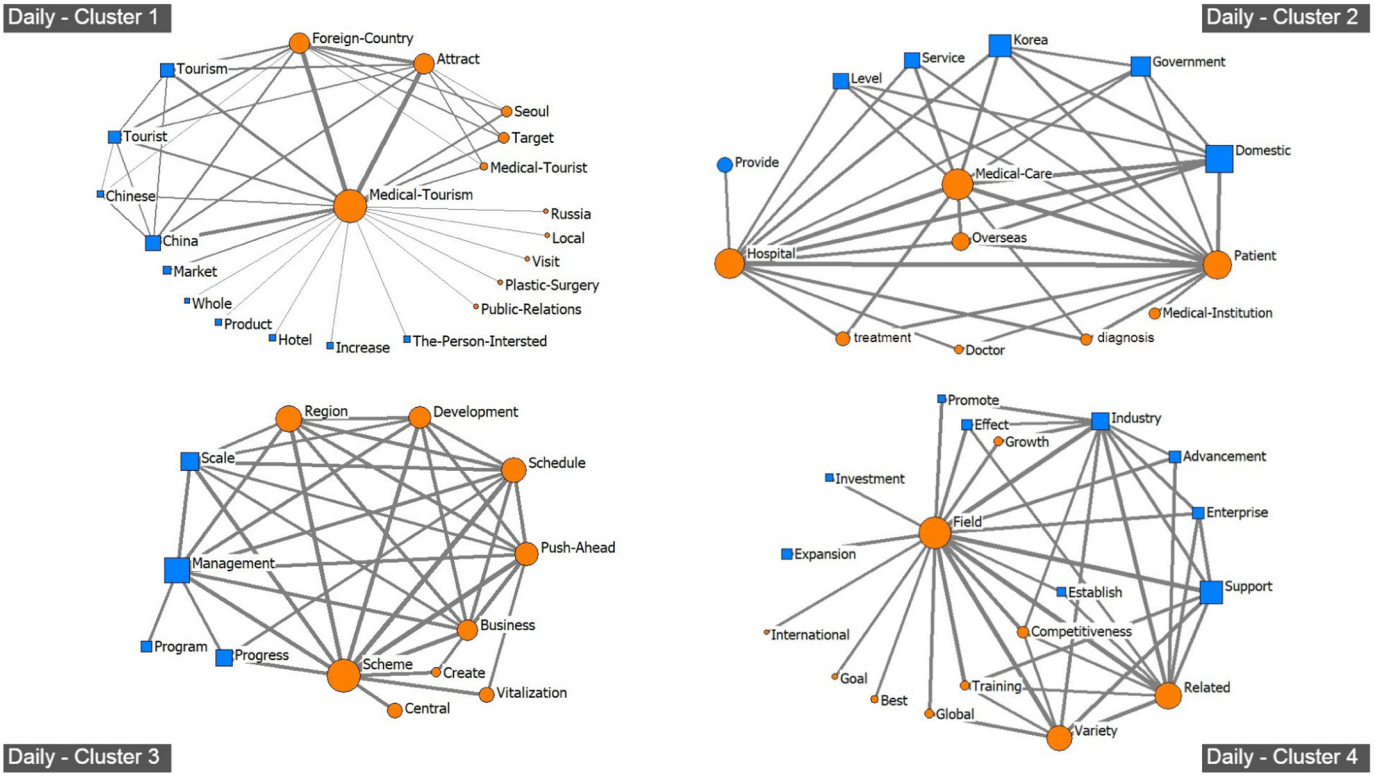

Fig. 4. Daily newspaper top 100 links visualization.

'Diagnosis' and 'Medical Service', appeared in cluster 2–1. Keywords related to service level and the attention of the nation appeared in cluster 2–2. Third, clusters 3–1 and 3–2 pertain to regional vitalization. Cluster 3–1 contains keywords such as 'Region'; 'Daegu', a city in Korea; 'Development'; and 'Vitalization'. Cluster 3–2 pertains to management programs and facilities. These clusters are grouped together; thus, programs and facility management can be interpreted as one way to activate the region. Finally, clusters 4–1 and 4–2 are relevant to field

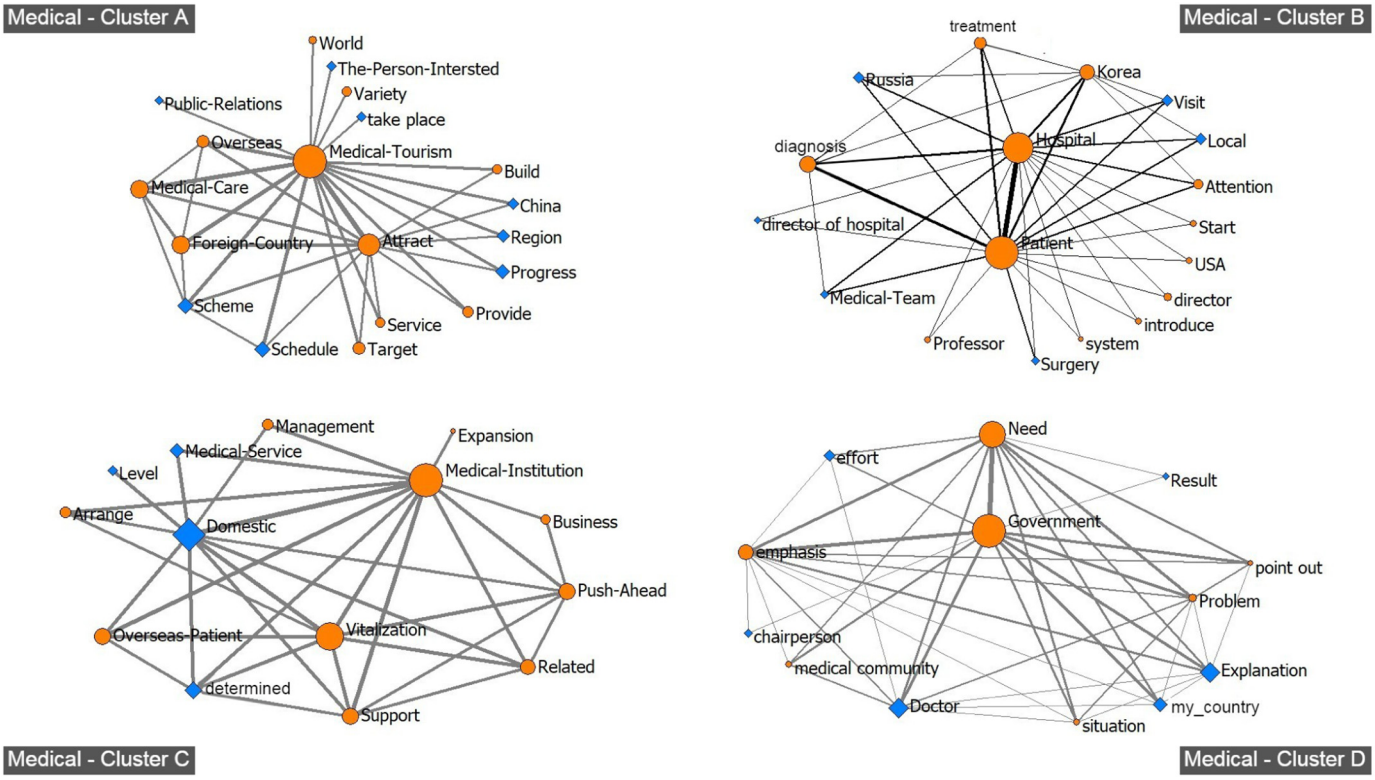

Fig. 5. Medical newspaper top 100 links visualization.

support. Keywords concerning field-related training and competitiveness appeared in clusters 4–1. Keywords concerning support of and investment in industry and enterprise appeared in cluster 4–2.

In medical newspapers, 'Support' was ranked more highly than in daily newspapers. Unlike in daily newspapers, in medical newspapers, 'support' more often refers to the subject of support than to those who receive support. First, clusters A-1 and A-2 derived from CONCOR analysis in medical newspapers are related to medical tourism marketing. Cluster A-1 indicates attraction of medical tourists and contains keywords regarding medical care. 'Region' and 'China' appeared in cluster A-2, but unlike in daily newspapers, in medical newspapers the words were grouped with the keywords 'Marketing' and 'Public Relations' keywords. Second, clusters B-1 and B-2 introduce medical treatment overseas. Cluster B-1 contains keywords that indicate attention and introduction of treatment into the USA. Cluster B-2 contains keywords that indicate medical travel by medical teams to Russia. Third, clusters C-1 and C-2 are associated with medical tourism vitalization with government support. Cluster C-1 indicates support and vitalization of medical institutions by the Ministry of Health and Welfare. Cluster C-2 indicates medical service level and the role of a nation. Finally, clusters D-1 and D-2 are related to the medical community's voice in medical tourism. Keywords related to problems indicated by the medical community are observed in cluster D-1, and keywords pertaining to doctors' effort to provide an explanation are found in cluster D-2. To determine the cause of the results, we reviewed the content of the collected articles and identified that doctors had an opportunity to explain medical tourism in person through a familiarization tour or briefing session.

Differences in clusters between daily and medical newspapers were observed. Unlike in medical newspapers, in daily newspapers, tourist-related words such as 'hotel' were also derived, which is consistent with previous studies indicating that the medical community rarely addresses the tourism sector (Jagyasi, 2008). Even the same word functions differently when grouped with other words. For example, 'Vitalization' and 'Development' were clustered with 'Region' in daily newspapers, while those words were clustered with 'Medical Institution' in medical newspapers. This result suggests that the purpose of medical tourism varies. The focus is on regional vitalization from public perspectives, but from the medical community's point of view, medical tourism focuses more on the vitalization of medical institutions. 'Support' was also clustered differently between the types of newspapers. In daily newspapers, 'Support' was grouped with 'Enterprise' and 'Investment', while it was clustered with 'Medical Institution', and 'Ministry of Health and Welfare' in medical newspapers. In other words, daily newspapers concentrated on private investment in and support for medical tourism, while medical newspapers discussed the support and the role of the nation and the Ministry. Therefore, the central goal and role of government must be clear when forming a consultative group with the medical community to promote medical tourism.

6. Limitations

Newspaper articles have the benefit of including diverse stakeholders' opinions about a particular subject and more information resources than the public (Choi et al., 2016), but they have limitations in that they do not present data reflecting the direct experience of medical tourists. Therefore, future studies should collect and analyze data that reflect the direct experiences of medical tourists from SNS, such as Twitter, blogs, and Facebook. Additionally, because not all newspaper articles on medical tourism were collected, there are limitations with respect to the generalization of the results of this study to perceptions of medical tourism.

Moreover, there are limitations due to natural language processing, which extracts keywords from unstructured newspaper articles. First, although we preprocessed data, there may be words that were not sufficiently processed due to the vastness of the data. The same limitation has affected other studies using text network analysis (Park, Ahn, & Park, 2018). Additionally, there may be limitations in interpreting the connections between words, as also described in other studies using text network analysis (Kim & Kim, 2017).

References

Álvarez, M. M., Chanda, R., & Smith, R. D. (2011). The potential for bi-lateral agreements in medical tourism: A qualitative study of stakeholder perspectives from the UK and India. *Glob. Health, 7*(1), 11. https://doi.org/10.1186/1744-8603-7-11.

Bies, W., & Zacharia, L. (2007). Medical tourism: Outsourcing surgery. *Math. Comput. Model. 46*(7), 1144–1159. https://doi.org/10.1016/j.mcm.2007.03.027.

Borgatti, S. P. (2002). *NetDraw software for network visualization.* Lexington, KY: Analytic Technologies.

Borgatti, S. P., Everett, M. G., & Freeman, L. C. (2002). *Ucinet for windows: Software for social network analysis.* Harvard, MA: Analytic Technologies.

Chakraborty, G., & Pagolu, M. K. (2014, March). Analysis of unstructured data: Applications of text analytics and sentiment mining. *SAS global forum* (pp. 1288–2014). .

Cho, K. H. (2006). An exploratory study on the activation plan of medical tourism in Korea. *Korean Journal of Hospitality & Tourism, 15*(4), 189–203.

Cho, S. S. (2015). A research on medical tourism activation plans in creative economy ages. *Journal of Tourism and Leisure Research, 27*(3), 253–272.

Choi, S., Lee, J., & Kwon, O. (2016). A morphological analysis method of predicting place-event performance by online news titles. *Journal of Society for e-Business Studies, 21*(1), 15–32. https://doi.org/10.7838/jsebs.2016.21.1.015.

Chun, H., & Lee, H. (2010). Market segmentation of overseas medical tourism and attributes of medical tourism destination. *Journal of Tourism and Leisure Research, 22*(6), 7–23.

Connell, J. (2006). Medical tourism: Sea, sun, sand and… surgery. *Tour. Manag. 27*(6), 1093–1100. https://doi.org/10.1016/j.tourman.2005.11.005.

Connell, J. (2013). Contemporary medical tourism: Conceptualisation, culture and commodification. *Tour. Manag. 34*, 1–13. https://doi.org/10.1016/j.tourman.2012.05.009.

Diesner, J., & Carley, K. M. (2005). Revealing social structure from texts: Meta-matrix text analysis as a novel method for network text analysis. In V. K. Narayanan, & D. J. Armstrong (Eds.). *Causal mapping for research in information technology* (pp. 81–108). IGI Global. https://doi.org/10.4018/978-1-59140-396-8.ch004.

Edelheit, J. S. (2008). Defining medical tourism or not. *Medical Tourism Magazine, 5*(1), 9–10.

Fan, W., Wallace, L., Rich, S., & Zhang, Z. (2006). Tapping the power of text mining. *Commun. ACM, 49*(9), 76–82. https://doi.org/10.1145/1151030.1151032.

Feinerer, I., & Hornik, K.. R package 'tm'. R foundation for statistical computing. (2015). URL http://tm.r-forge.r-project.org (last date accessed: 19 October 2018).

Feldman, R., & Sanger, J. (2007). *The text mining handbook: Advanced approaches in analyzing unstructured data.* Cambridge: Cambridge university press.

Fetscherin, M., & Stephano, R. M. (2016). The medical tourism index: Scale development and validation. *Tour. Manag. 52*, 539–556. https://doi.org/10.1016/j.tourman.2015.08.010.

Garcia-Altes, A. (2005). The development of health tourism services. *Ann. Tour. Res. 32*(1), 262–266. https://doi.org/10.1016/j.annals.2004.05.007.

Godnov, U., & Redek, T. (2016). Application of text mining in tourism: Case of Croatia. *Ann. Tour. Res. 58*, 162–166.

Ha, S. H., & Lee, M. K. (2012). A news frame analysis by the south Korean press on the livelihoods of a north Koreans. *Korean Journal of Communication and Information, 58*, 222–241.

Hearst, M. (2003). *What is text mining. SIMS.* UC Berkeley. URL https://www.jaist.ac.jp/~bao/MOT-Ishikawa/FurtherReadingNo1.pdf.

Hearst, M. A. (1999, June). Untangling text data mining. *Proceedings of the 37th annual meeting of the Association for Computational Linguistics on computational linguistics* (pp. 3–10). Association for Computational Linguistics. https://doi.org/10.3115/1034678.1034679.

Horowitz, M. D. (2007). Medical tourism-health care in the global economy. *Physician Executive, 33*(6), 24–30.

House of Commons Public Administration and Constitutional Affairs Committee (PACAC) (2017). Managing Ministers' and officials' conflict of interest: time for clearer values, principles and action. Thirteenth report of session 2016-17. URL https://publications.parliament.uk/pa/cm201617/cmselect/cmpubadm/252/252.pdf.

Hudson, S., & Li, X. (2012). Domestic medical tourism: A neglected dimension of medical tourism research. *Journal of Hospitality Marketing & Management, 21*(3), 227–246. https://doi.org/10.1080/19368623.2011.615018.

Jagyasi, P. (2008). Defining medical tourism: Another approach. *Medical Tourism Magazine, 6*, 9–11.

Jeon, H., & Kim, T. (2016). R package 'KoNLP'. R Foundation for Statistical Computing. URL https://cran.r-project.org/web/packages/KoNLP, Accessed date: 19 October 2018.

Kang, D. K. (2015). *The Economy Impact of Medical Tour by Foreigners. IOM migration research and training Centre.*

Kham, M. K. (2010). Most wanted magazines to read by field. *Sisa Journal.* (October 18) URL http://www.sisapress.com/journal/article/130408, Accessed date: 27 February 2018.

Kim, D., & Kim, I. (2017). An analysis of hotel selection attributes present in online reviews using text mining. *Journal of Tourism Sciences, 41*(9), 109–127. https://doi.org/10.17086/jts.2017.41.9.109.127.

Kim, H. M. (2011). Evaluating a brand image of Daegu city's medical tourism. *International Journal of Tourism Management and Sciences, 26*(1), 95–114.
Kim, J. E. (2017b). *Keyword and Topic analysis on the College and University Structural Reform Evaluation Using Big Data.* Seoul, South Korea: Unpublished Doctoral Dissertation. Seoul National University.
Kim, K. H., Song, H. J., & Noh, G. Y. (2012). Comparing the media frames of medical tourism. *Kookmin Social Science Revies, 24*(2), 1–28.
Kim, Y. K.. 일본은 대사관까지 의료관광 홍보하는데... *Joongangilbo (2017)*. (February 02) URL https://news.joins.com/article/21199847 ((last date accessed: 19 October 2018)).
Ko, T. G. (2012). *Medical tourism management. Seoul:* 무역경영사. https://doi.org/10.978.89468/05187 ((last date accessed: 19 October 2018)).
Korea Audit Bureau of Certification (2017). *A paid circulation of the 2017 daily newspaper.* (Seoul, South Korea).
Korea Tourism Organization (2016). *Korea Medical Tourism Marketing 2016.* (Wonju, South Korea).
Kwak, D. Y. (2011). A study on product choice attributes of medical tourists. *Korea Journal of Tourism and Hospitality Research, 25*(4), 85–101.
Lautier, M. (2008). Export of health services from developing countries: The case of Tunisia. *Social Science & Medicine, 67*(1), 101–110. https://doi.org/10.1016/j.socscimed.2008.01.057.
LaValle, S., Lesser, E., Shockley, R., Hopkins, M. S., & Kruschwitz, N. (2011). Big data, analytics and the path from insights to value. *MIT Sloan Manag. Rev. 52*(2), 21.
Leahy, A. L. (2008). Medical tourism: The impact of travel to foreign countries for healthcare. *Surgeon, 6*(5), 260–261. https://doi.org/10.1016/s1479-666x(08)80047-9.
Lee, E. M., Kim, W. I., & Lee, G. H. (2009). Medical tourism selection attributes and development strategy for Daegu. *Korean Tournal of Tourism Research, 24*(2), 109–125.
Lee, H. G., & Lee, J. C. (2010). The effect of quality of medical service on medical customer satisfaction and sustainable medical tourism. *Korean Journal Tourism Research, 25*(3), 279–296.
Lee, H. J. (2016). Study on the national qualification examination for international medical tour coordinator which students majoring in healthcare recognize. *International Journal of Tourism Management and Sciences, 31*(2), 45–60.
Lee, M. B., & Park, S. H. (2016). The effect of engagement in Korean dramas and cosmetic surgery programs on the attitudes of Chinese people regarding cosmetic surgery and their behavior intentions. *International Journal of Tourism Management and Sciences, 31*(3), 115–133.
Lee, W. K., & Chung, B. W. (2007). A case study for the activation of medical tourism. *Journal of Medical Tourism, 21*(2), 389–406.
Meijer, A. J. (2007). Publishing public performance results on the internet: Do stakeholders use the internet to hold Dutch public service organizations to account? *Gov. Inf. Q. 24*(1), 165–185. https://doi.org/10.1016/j.giq.2006.01.014.
Mendenhall, R., Brown, N. M., & Black, M. L. (2017). The potential of big data in rescuing and recovering black women's contributions to the Du bois-Atlanta school and to American sociology. *Ethn. Racial Stud. 40*(8), 1231–1233. https://doi.org/10.1080/01419870.2017.1285421.
Newman, M. E. J. (2008). Mathematics of networks. In S. N. Durlauf, & L. E. Blume (Eds.). *The new palgrave dictionary of economics* (pp. 465–470). (2nd ed.). London: Palgrave Macmillan.
Noree, T., Hanefeld, J., & Smith, R. (2016). Medical tourism in Thailand: A cross-sectional study. *Bull. World Health Organ. 94*(1), 30–36. https://doi.org/10.2471/blt.14.152165.
Oh, S. S., & Jeon, M. S. (2017). A study on the revitalization of Korea medical tourism through Hallyu culture. *Journal of Convergence Tourism Contents, 3*(2), 41–56.
Paranyushkin, D. (2010). Text network analysis. http://issuu.com/deemeetree/docs/text-network-analysis.
Park, C., & Jung, C. (2013). Text network analysis: Detecting shared meaning through socio-cognitive networks of policy stakeholders. *Journal of Governmental Studies, 19*(2), 73–108.
Park, E. J., Ahn, D. W., & Park, C. S. (2018). Text network analysis of newspaper articles on life-sustaining treatments. *J. Korean Acad. Community Health Nurs. 29*(2), 244–256. https://doi.org/10.12799/jkachn.2018.29.2.244.
Richardson, J. (2006). *Analysing newspapers: An approach from critical discourse analysis.* Basingstoke: Palgrave.
Richardson, L. (2017). Python library 'Beautiful Soup'. URL https://www.crummy.com/software/BeautifulSoup/.
Shetty, P. (2010). Medical tourism booms in India, but at what cost? *Lancet, 376*(9742), 671–672. https://doi.org/10.1016/s0140-6736(10)61320-7.
Turner, L. (2007). 'First world health care at third world prices': Globalization, bioethics and medical tourism. *BioSocieties, 2*(3), 303–325. https://doi.org/10.1017/s1745855207005765.
Walker, S. J. (2014). Big data: A revolution that will transform how we live, work, and think. *Int. J. Advert. 33*(1), 181–183. https://doi.org/10.2501/ija-33-1-181-183.
Wasserman, S., & Faust, K. (1994). *Social network analysis: Methods and applications. Vol. 8.* Cambridge university press.
Yoon, S. M., Han, J. S., & Kim, H. R. (2009). A study on the effects relation of medical service quality, service value and clients' satisfaction for the medical tourism. *Journal of the Korea Service Management Society, 10*(1), 137–157.

Sohyeon Kim BS, Sohyeon Kim is currently pursuing her Master's degree in leisure and tourism sciences at Kyonggi University. Her research interests include text mining and issues of discrimination against tourists.

Won Seok Lee Ph.D, Won Seok Lee is an assistant professor in the department of tourism and recreation at Kyonggi University in Korea. He earned his Ph.D. in recreation, park, and tourism management from the Pennsylvania State University. His research interests are tourism-economics.