



Analysis of Social Needs for Doctors and Medicine through a Keyword Analysis of Newspaper Articles (2016–2020)

Hanna Jung¹, Jea Woog Lee², Geon Ho Lee³

¹Department of Medical Education, Yonsei University College of Medicine, Seoul, Korea

²Chung-Ang University, Institute of Sports Convergence, Seoul, Korea

³Department of Family Medicine, Daegu Catholic University School of Medicine, Daegu, Korea

The purpose of this study was to explore, using topic modeling, the social value of doctors and medicine demanded by society as reflected in published newspaper articles in Korea. Ultimately, this study aimed to reflect social needs in the process of developing the Patient-Centered Doctor's Competency Framework in Korea. For this purpose, a total of 2,068 newspaper articles published from 2016 to 2020 were analyzed. Through topic modeling of these newspaper articles over the past 5 years, 18 topics were derived and divided into four categories. Focusing on the derived topics and keywords, the topics derived in specific years and the proportion of topics by year were analyzed. The results of this study make it possible to grasp the needs of society projected through the press for doctors and medicine. Due to the nature of the press, topics that frequently appeared in newspaper articles were mainly social phenomena related to requirements for doctors, particularly dealing with economic and legal aspects. In particular, it was confirmed that doctors are now required to have a wider range of competencies that go beyond their required medical knowledge and clinical skills. This study helped to establish doctors' competencies by analyzing social needs for doctors through the latest research methods, and the findings could help to establish and improve doctors' competencies through ongoing research in the future.

Keywords: Doctor's competency; Healthcare industry; Health policy; Patient-centered care; Topic modeling

Introduction

In the early 20th century, the Flexner Report catalyzed reforms in medical education that emphasized basic science knowledge and prioritized the training of medical education and clinical skills. These reforms laid the groundwork for the modern medical education system. By the 1990s, the concept of physician competencies had become integral to medical education, leading to the definition and incorporation of specific competencies for physicians in many countries [1]. In the United States, the Accreditation Council for Graduate Medical Educa-

tion (ACGME) oversees residency programs and introduced six core competencies in 1999 through the ACGME Outcome Project. These competencies have undergone revisions up until 2019 [2]. Canada undertook the CanMEDS 2000 Project in 1996, a multidisciplinary effort that defined seven key competencies. These have been updated periodically, with the latest iteration being the CanMEDS 2015 Physician Competency Framework [3]. The United Kingdom has maintained four competencies since the General Medical Council established Good Medical Practice in 1995, with updates continuing through to 2019 [4]. Similarly, Korea has defined and developed physician competencies through various studies, including "RESPECT 100, Development of a Generic Curriculum for Graduate Medical Education in Korea (Korean Institute of Medical Education and Evaluation, 2008)," "Survey on General and Specialty Competencies Education and Training in Specialty Boards (Korean Institute of Medical Education and Evaluation, 2012)," "Competency of Residents as Learners and Teachers (Korean Institute of Medical Education and Evaluation, 2013)," "Reorganizing Training Courses by Specialty for the Efficient Training of Specialty Physicians (Korean Academy of Medical Sciences, 2013)," and "Korean doctor's role (Ministry of Health and Welfare, 2014)" [5].

Received: December 26, 2023 **Accepted:** December 27, 2023

Corresponding author: Geon Ho Lee

Department of Family Medicine, Daegu Catholic University School of Medicine, 33 Duryugongwon-ro 17-gil, Nam-gu, Daegu 42472, Korea
Tel: +82-53-650-4696 Fax: +82-53-650-4122 E-mail: totoslee@cu.ac.kr

This is a secondary publication of an article published in the Korean Medical Education Review 2022; 24(2): 103-112, <https://doi.org/10.17496/kmer.2022.24.2.103>, with the approval of the editor and the author. The articles are identical. Either citation can be used when citing this article.

As such, physicians' competencies have become a critical issue both in Korea and internationally, with a focus on the perspectives of experts and stakeholders in its identification and development. However, the concept of physicians' competencies is influenced by the political, social, and economic context. Therefore, to establish a relevant definition of physicians' competencies, it is essential to monitor healthcare-related phenomena in contemporary society and to understand current social issues [1]. This means that physician competence must be viewed through a multifaceted lens that includes political, social, and economic considerations, in addition to professional perspectives. To achieve this, an analysis of the social value of doctors and healthcare as portrayed in the media, such as newspapers and news reports, is necessary. In response to this need, this study aimed to analyze major Korean newspaper articles to assess social demands for physicians' competencies in the process of developing a framework for "patient-centered physician competencies in Korea," the first subtask of the "A Study on Improving the System of Graduate Medical Education to Enhance Patient-Centered Performance," led by Jeon et al. [6] with support from the National Evidence-based Healthcare Collaborating Agency.

Newspapers are the oldest form of mass media [7,8] and have the advantage of reflecting social values and zeitgeist, including the interests and needs of the public [8]. They play a pivotal role in guiding social discourse by highlighting the prevailing interests and issues of the time. In essence, media texts, including newspapers, serve as instruments for the social construction of frames and discourses on specific topics, thereby reflecting the opinions, beliefs, and values of the broader society at that moment [9]. Analyzing newspaper articles through this lens can provide insights into the public's interest in and perception of the social value of doctors and healthcare. This, in turn, can inform challenges and directions for future research and policy development. Despite this potential, there is a notable absence of studies that have examined the social value of doctors and medical care through media channels such as newspaper articles, underscoring the importance and necessity of such research.

Newspapers represent a vast repository of textual data that record daily events and societal issues over time. Analyzing newspaper articles can reveal trends through time-series data, tracing developments from the past to the present [10,11]. Topic modeling is a useful approach for examining newspaper articles, given their nature. It is a big data analysis technique employed to identify recurring themes within large volumes of unstructured text, such as that found in newspapers [12]. The use of topic modeling to analyze unstructured data is significant because it can uncover new insights or value not obtainable through traditional methods of analyzing structured data [13]. Due to these advantages, topic

modeling has been adopted as an analytical tool across various research domains [14].

Therefore, this study explored the social value of doctors and healthcare as perceived and demanded by society by using topic modeling to analyze Korean newspaper articles. Ultimately, the goal of this study is to incorporate the social needs in the process of developing a framework for a patient-centered doctor's competency framework in Korea.

Methods

1. Data collection

In this study, we collected newspaper articles spanning a recent 5-year period, from January 1, 2016, to December 31, 2020, to conduct topic modeling on physician competencies. We utilized the news trend analysis website Big Kinds (www.bigkinds.or.kr) to search for and retrieve the articles. Big Kinds is a news big data analysis system provided by the Korea Press Foundation. It currently archives and offers access to news from 54 media companies in the form of big data, which facilitates the analysis of a vast quantity of news. Furthermore, Big Kinds allows researchers to tailor their news extraction by setting specific parameters such as the time period, media company, search area, and search method, as well as by inputting relevant search terms. Leveraging these features, we were able to selectively extract the news articles pertinent to our study. The detailed criteria and methodology for the collection are as follows. The study analyzed news outlets including Kyunghyang Shinmun, Kookmin Ilbo, Tomorrow's Newspaper, Donga Ilbo, Culture Daily, Seoul Shinmun, World Daily, Chosun Ilbo, JoongAng Ilbo, Hankyoreh, and Hankook Ilbo. These are classified by Big Kinds as 11 major Korean general dailies. We focused on news content provided by Big Kinds, specifically targeting media companies that publish newspaper articles. The search parameters were set to include the terms "doctor AND (competency OR role) AND (medical)" with the search term processing utilizing "morphological analysis." The search yielded a total of 9,781 newspaper articles, with an annual breakdown as follows: 1,412 articles in 2016, 1,447 in 2017, 1,766 in 2018, 1,724 in 2019, and 3,432 in 2020.

When conducting a morphological analysis search, all articles sharing the same morphology as the entered search term are retrieved, regardless of their subject matter. Consequently, it became necessary to re-categorize the newspaper articles by their titles or topics. In this study, we removed duplicates and irrelevant articles from the initial collection obtained through Big Kinds, using Microsoft Excel (Microsoft Corp., Redmond, WA, USA). A total of 6,589 newspaper articles were discarded. Additionally, we eliminated 1,124 articles that were adver-

tisements or promotional in nature. Ultimately, we analyzed a total of 2,068 newspaper articles, which included 378 from 2016, 294 from 2017, 371 from 2018, 432 from 2019, and 593 from 2020.

2. Data processing

Since the collected newspaper articles consist of text sentences and paragraphs, keyword analysis necessitates a data preprocessing step to break down the data into individual words, rendering them suitable for analysis. To this end, the current study undertook data preprocessing to extract keywords that not only encapsulate the societal demand for doctors and medical care as depicted in the newspaper articles but are also compatible with the input requirements of the analysis program. We utilized the IMC Textom, a web-based big data analysis program developed by The IMC Inc. in Daegu, Korea, for this data preprocessing task. Textom was selected for its efficiency in word extraction, particularly its feature that automatically transforms verb forms into their corresponding noun forms. In our study, data preprocessing involved conducting detailed morphological analysis and text mining using the functionalities provided by Textom.

In some instances, the words extracted by Textstorm may accurately reflect the content of the newspaper article, yet their grammatical form is that of a verb rather than a noun, or they fail to meaningfully describe the article's content. In our study, we utilized Textom's automation feature to convert words from their verb form to nouns. Additionally, we eliminated stopwords from the extracted words. These stopwords include conjunctions, prepositions, adverbs, and articles that lack specific meaning, as well as noun-type words that do not convey meaningful content. Examples of such stopwords are "however," "like," "one after another," "on the contrary," "this," and "those." Lastly, we excluded words that appeared fewer than 5 times from our study, as their infrequent occurrence may not provide a clear indication of overarching trends.

Words with low term frequency-inverse document frequency (TF-IDF) values were removed. TF-IDF is not simply a measure of word frequency across all newspaper articles; it gauges how often a word appears in a specific article relative to a corpus. This metric reflects the relevance of a word to a single document or topic. A low TF-IDF score indicates that a word is common across all documents in the corpus, making it less useful for describing a specific topic due to its ubiquity. Conversely, a high TF-IDF score signifies that a word is frequent in a particular document but rare in others. In essence, the more common a word is across all documents, the higher its TF-IDF score will be, and the closer the score is to zero, the more common the word is. For this study, we set the threshold for TF-IDF at 0.5 or lower. Additionally, we standardized terms with the same meaning but different forms among

the derived words.

3. Topic modeling

To extract specific topics, we conducted topic modeling on the final selection of keywords following data preprocessing. Topic modeling is an analytical technique that uncovers the relationships between words within large-scale documents and derives structured topics. This technique encompasses methods such as latent semantic analysis, probabilistic latent semantic analysis, and latent Dirichlet allocation (LDA), with LDA being the most widely recognized topic model [15]. Recently, there has been a surge in research papers employing LDA techniques within the realm of medical education [16-18]. LDA topic modeling identifies the extent to which words in one newspaper article co-occur in others, as well as the intermediary relationships between these words. It clusters words that meet the researcher-defined criteria and designates them as a topic group if they are deemed representative of a specific topic. In our study, we applied topic modeling to newspaper articles from 2016 to 2020, categorizing topics related to societal needs for physicians and medical services based on the words identified post-data preprocessing. The topic modeling was executed using the Net Miner ver. 4.2 software (Cyram Inc., Seongnam, Korea).

4. Topic categorization

The topics identified by year through topic modeling were categorized in two steps, all of which involved discussion by the researcher. In Step 1, categorization occurred on an annual basis. For each year, the researcher assigned names to each topic through discussion, and topics that were the same or similar were grouped together as a single topic. Step 2 involved categorization based on the outcomes of Step 1, grouping together topics with similar themes across different years. The research process is illustrated in Figure 1.

Results

1. 2016–2020 topic modeling and topic categorization results

From January 1, 2016, to December 31, 2020, a total of 2,068 newspaper articles related to the competence and role of physicians were published: 378 in 2016, 294 in 2017, 371 in 2018, 432 in 2019, and 593 in 2020 (Table 1). Topic modeling performed on these articles yielded 10 topics in 2016, nine topics in both 2017 and 2018, 11 topics in 2019, and 13 topics in 2020. When the topics identified through topic modeling shared similar keywords or themes, they were combined and categorized for each respective year. Consequently, eight top-

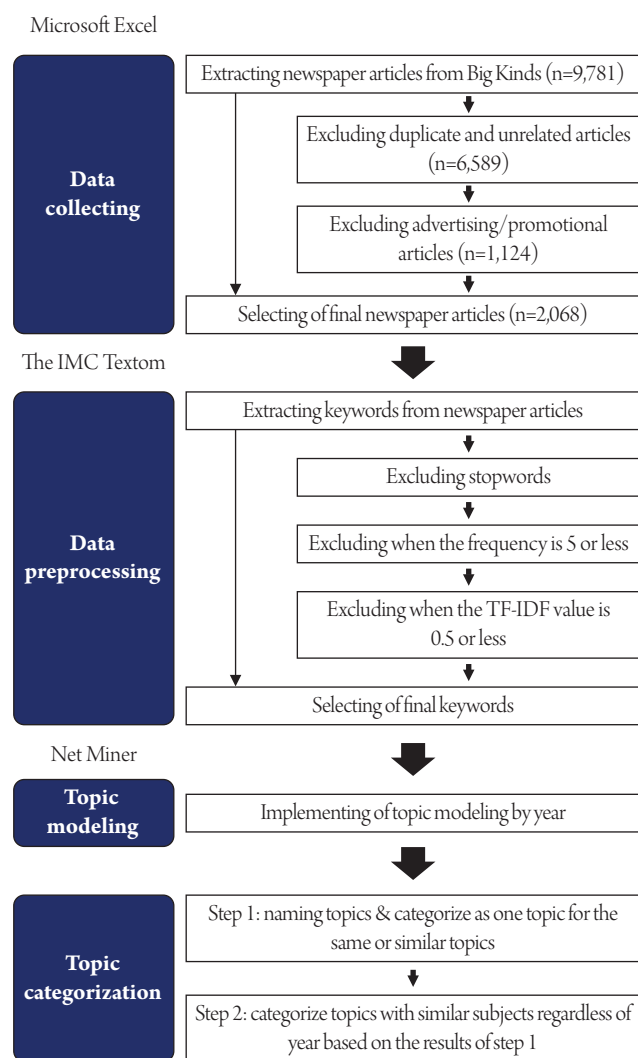


Figure 1. The procedure of the study. TF-IDF, term frequency–inverse document frequency.

ics were selected for 2016, seven for 2017, six for 2018, eight for 2019, and nine for 2020. The finalized topics were then assigned appropriate names.

The eight topics identified in 2016 were “digital healthcare,” “healthcare system,” “communicable disease control,” “medical technology,” “global healthcare,” “publicness in healthcare,” “medical volunteering,” and “patient safety.” The seven topics identified in 2017 were “patient safety,” “medical volunteering,” “publicness in healthcare,” “palliative care,” “convergence,” “healthcare system,” and “digital healthcare.” In 2018, the six topics were “digital healthcare,” “healthcare system,” “patient-centeredness and good doctors,” “publicness in healthcare,” “medical science research,” and “global healthcare.” In 2019, the eight topics

were “digital healthcare,” “medical technology,” “publicness in healthcare,” “patient safety,” “global healthcare,” “healthcare policy,” “medical volunteering,” and “palliative care.” Finally, the nine topics that emerged in 2020 were “digital healthcare,” “publicness in healthcare,” “conflict between medical professionals and the government,” “COVID-19: telemedicine,” “telemedicine,” “medical volunteering,” “healthcare policy,” “COVID-19: lack of medical resources,” and “COVID-19: dedication of medical staff.”

As shown above, a total of 18 topics have been identified over the past 5 years. These topics are “digital healthcare,” “communicable disease control,” “medical technology,” “global healthcare,” “publicness in healthcare,” “medical volunteering,” “patient safety,” “palliative care,” “convergence,” “healthcare system,” “patient-centeredness and good doctors,” “medical science research,” “healthcare policy,” “conflict between medical professionals and the government,” “telemedicine,” “COVID-19: lack of medical resources,” “COVID-19: dedication of medical staff,” and “COVID-19: telemedicine.” Among these, “digital healthcare” and “publicness in healthcare” have been the most consistently discussed topics over the 5-year period. In contrast, certain topics emerged in specific years: “communicable disease control” appeared in 2016, “convergence” in 2017, “patient-centeredness and good doctors” and “medical science research” in 2018, and “conflict between medical professionals and the government,” “telemedicine,” “COVID-19: lack of medical resources,” “COVID-19: dedication of medical staff,” and “COVID-19: telemedicine” in 2020.

The 18 topics were grouped into four categories based on their similarities: healthcare industry, healthcare policy, patient-centered medicine, and infectious diseases. Specifically, the topics “digital healthcare,” “global healthcare,” “medical technology,” “convergence,” “medical science research,” and “telemedicine” were categorized under “healthcare industry.” The topics “publicness in healthcare,” “healthcare system,” “healthcare policy,” and “conflict between medical professionals and the government” were grouped under “healthcare policy.” The topics “medical volunteering,” “patient safety,” “palliative medicine,” “patient-centeredness,” and “good doctors” were classified as “patient-centered medicine.” Lastly, “communicable disease control,” “COVID-19: lack of medical resources,” “COVID-19: dedication of medical staff,” and “COVID-19: telemedicine” were categorized as “infectious diseases.”

2. Main keywords by topic from 2016 to 2020

The main keywords for each topic are listed below (Appendix 1). The topics identified for the 5-year period from 2016 to 2020 are “digital healthcare” and “publicness in healthcare.” The main keywords for “digital healthcare” are summarized as follows: “Watson,” “diagnostics,”

Table 1. The proportion of topics by year

Topic	Year				
	2016	2017	2018	2019	2020
Healthcare industry					
Digital healthcare	57 (15.1)	63 (21.4)	85 (22.9)	89 (20.6)	85 (14.3)
Global healthcare	43 (11.4)	-	78 (21.0)	58 (13.4)	-
Healthcare technology	77 (20.4)	-	-	92 (21.3)	-
Convergence	-	59 (20.1)	-	-	-
Medical science research	-	-	36 (9.7)	-	-
Telemedicine	-	-	-	-	51 (8.6)
Healthcare policy					
Publicness in healthcare	38 (10.1)	44 (15.0)	67 (18.1)	42 (9.7)	39 (6.6)
Healthcare system	25 (6.6)	34 (11.6)	57 (15.4)	-	-
Healthcare policy	-	-	-	28 (6.5)	42 (7.1)
Conflict between medical professionals and the government	-	-	-	-	71 (12.0)
Patient-centered care					
Medical volunteering	49 (13.0)	34 (11.6)	-	38 (8.8)	38 (6.4)
Patient safety	31 (8.2)	32 (10.9)	-	51 (11.8)	-
Palliative care	-	28 (9.5)	-	34 (7.9)	-
Patient-centered and good doctors	-	-	48 (12.9)	-	-
Communicable diseases					
Communicable disease control	58 (15.3)	-	-	-	-
COVID-19: lack of medical resources	-	-	-	-	97 (16.4)
COVID-19: dedication of medical staff	-	-	-	-	84 (14.2)
COVID-19: telemedicine	-	-	-	-	86 (14.5)
Total	378 (100.0)	294 (100.0)	371 (100.0)	432 (100.0)	593 (100.0)

Values are presented as number (%).

COVID-19, coronavirus disease 2019.

“platform,” “artificial intelligence (AI),” “data,” “precision,” “advanced,” “convergence,” “internet of things (IoT),” “hub,” “robotics,” “healthcare,” “personalized,” “smart,” “video,” “augmented reality,” “reading,” “telemedicine,” “big data,” “prediction,” “digital healthcare,” “innovation,” “digital,” “remote work,” “deep learning,” “Industry 4.0.” The keywords for “publicness in healthcare” include “public medical school,” “establishment,” “public,” “universal,” “workforce,” “public healthcare,” “elderly,” “long-term care,” “dementia,” “welfare,” “dementia care,” “for-profit,” “medical welfare,” “healthcare disparities,” “public health scholarships,” “vulnerable,” “strike,” “resident,” “medical association,” “local government,” and “private.”

“Medical volunteering” was identified in 4 years, with the exception of 2018, and the main keywords were “mission,” “service,” “doctors without borders,” “Africa,” “religion,” “medical service,” “relief,” “free,” “overseas,” “free hospital,” “welfare center,” “Bangladesh,” “refugee,” “poor,” “donation,” “Gambia,” “Vietnam,” “rural,” “Ethiopia,” “Morocco,” “Tae-Seok Lee,” “Afghanistan,” “heart center,” “poverty,” and “Cambodia.”

The topics that emerged in 3 years were “healthcare system,” “patient

safety,” and “global healthcare.” The main keywords for “healthcare system” from 2016 to 2018 were “rural,” “health center,” “medicine,” “facility,” “network,” “service,” “community care,” “primary care,” “health,” “healthcare,” “national primary care,” “prevention,” “dementia,” “long-term care,” “elderly,” “primary care system,” “home visits,” “nursing home,” “public guardianship,” “national accountability system,” and “neighborhood doctor.” “Patient safety” is a topic that emerged in 2016, 2017, and 2019, with the following keywords: “mediation,” “dispute,” “arbitration,” “accident,” “lawsuit,” “proceeding,” “discipline,” “arbitrator,” “negligence,” “civil,” “judgment,” “suspension,” “disposition,” “show doctor,” “Namki Baek,” “illegal,” “death,” “medicine,” “violation,” “medical law,” “ethics,” “medical dispute,” “unlicensed,” “court,” “illegal procedure,” “bioethics,” “manipulation,” “professional ethics,” “surrogate surgery,” “abortion,” and “license.” “Global healthcare” was derived in 2016, 2018, and 2019, and the keywords are: “advance,” “world,” “leading,” “market,” “attract,” “global,” “overseas,” “medical tourism,” “China,” “network,” “Mongolia,” “agreement,” “exchange,” “Southeast Asia,” “underdeveloped,” “spread,” “Russia,” “medical export,” “training,” “Korean Wave,” “visit,”

“medical market,” “international patients,” “United States,” “Myanmar,” “Bahrain,” “Middle East,” “tourism,” “Uzbekistan,” “Zambia,” and “Bangladesh.”

The topics identified in 2 years were “medical technology,” “palliative care,” and “healthcare policy.” “Medical technology” was identified in 2016 and 2019, with keywords including: “endoscopy,” “wearable,” “medical device,” “precision,” “lesion location,” “minimally invasive,” “laparoscopic,” “regenerative,” “procedure,” “transplant,” “biomarker,” “cochlear implant,” “face lift surgery,” “bionic,” “trauma center,” “gastrectomy,” and “tumor.” “Palliative care” was derived in 2017 and 2019, with the following keywords: “hospice,” “death preparation,” “death,” “terminal cancer,” “palliative care,” “cessation,” “Life-sustaining treatment,” “euthanasia,” “dignity,” “end of life,” “palliative,” and “terminal.” “Convergence” is a topic that only emerged in 2017, with keywords including “MERS,” “infectious disease,” “response,” “measures,” “countermeasures,” “inadequacies,” and “public health system.” In 2019 and 2020, the main keywords for “healthcare policy” were “public,” “health insurance,” “medical fee,” “public health,” “institution,” “critical care center,” “health insurance premium,” “medical helicopter,” “cooperative,” “hospitalist,” “emergency room,” “vulnerable,” and “Medicare.”

The topics identified in a single year were “communicable disease control,” “patient centeredness and good doctors,” “medical science research,” “conflict between medical professionals and the government,” “telemedicine,” “COVID-19: lack of medical resources,” “COVID-19: dedication of medical staff,” and “COVID-19: telemedicine.” “Communicable disease control” is a topic that was only identified in 2016, and it included keywords related to the Middle East respiratory syndrome (MERS) outbreak in 2015, with the following main keywords: “MERS,” “infectious disease,” “response,” “measures,” “countermeasures,” “inadequacies,” and “public health system.” “Patient-centered and good doctors” and “medical research” were derived only in 2018. The main keywords of “patient-centered and good doctors” were “total care,” “patient-centered,” “safety,” “communication,” “good doctors,” “medical quality,” “personalized care,” and the main keywords of “medical science research” were “scientist,” “training,” “research and development,” “research doctor,” “research-oriented hospital,” “technology development,” and “multidisciplinary.”

The topics that emerged in 2020 alone were: “Conflict between medical professionals and the government,” “telemedicine,” “COVID-19: lack of medical resources,” “COVID-19: dedication of medical staff,” and “COVID-19: telemedicine.” The main keywords of “conflict between medical professionals and the government” were “collective leave,” “strike,” “collective action,” “residents,” “medical association,” “gap,” “agitation,” “examination,” and “refusal.” The main keywords of the “telemedi-

cine” were “remote,” “industry,” “digital,” “online,” and “sensors.” In 2020, three topics related to COVID-19 emerged. The main keywords for “COVID-19: lack of medical resources” were “medical staff,” “lack,” “gap,” “urgent,” “beds,” “workforce,” “limits,” “financial difficulties,” and “burnout”; those for “COVID-19: dedication of medical staff” were “sacrifice,” “dedication,” “service,” “medical staff,” “medical volunteering,” “voluntary,” “workforce,” “lack,” “urgent,” “collaboration,” “cooperation,” “dispatch,” “support”; and those for “COVID-19: telemedicine” were “remote,” “telemedicine,” “contactless,” “diagnosis,” “regulation,” “permission,” “information and communication technology (ICT),” “response,” “infection,” “safety.”

3. Distribution of topics by year

The distribution of topics by year from 2016 to 2020 is presented in Table 1. In 2016, the category with the highest number of newspaper articles was “medical technology,” accounting for 77 articles, or 20.4% of the total. The topic “communicable disease control” was unique to 2016, comprising 58 articles (15.3%) that were associated with the 2015 MERS outbreak. The concept of “publicness in healthcare” was consistently explored from 2016 to 2020, with key issues encompassing public healthcare, the creation of public medical schools, and the public health scholarship program. In terms of public healthcare, the government unveiled the Public Healthcare Implementation Plan and Public Healthcare Plan in 2016, followed by the Comprehensive Plan for the Development of Public Healthcare in 2018. These initiatives aimed to bolster national accountability for public healthcare and to mitigate regional disparities in access to essential medical services. Subsequently, media coverage highlighted these regional disparities, focusing on essential healthcare, public healthcare, and medical services. Notably, in 2018, articles pertaining to “publicness in healthcare” accounted for 18.1% of the total, a higher proportion than in other years. This surge can be attributed to the heightened coverage of conflicts between healthcare professionals and the government during that time. The government’s proposal to establish a national public medical college, intended to alleviate the workforce shortage in essential healthcare, was met with resistance from medical professionals, resulting in noteworthy discord.

Among the seven topics identified in 2017, “digital healthcare” (63, 21.4%) and “convergence” (59, 20.1%) had the highest proportions. Notably, the share of “digital healthcare” topics saw a significant increase of 6.3% points, rising from 15.1% in 2016 to 21.4% in 2017. This surge can be attributed to the deployment of the artificial intelligence system “Watson for Oncology” (Watson) by Gachon University Gil Hospital in December 2016, which spurred a substantial number of related arti-

cles the following year. Furthermore, the category of “convergence” emerged exclusively in 2017. This emergence is likely linked to the integration of Watson into digital healthcare, suggesting the potential for blending multidisciplinary care and collaboration, as well as merging medicine and information technology.

The topics “patient-centeredness and good doctors” (48 articles, 12.9%) and “medical science research” (36 articles, 9.7%) were only identified in 2018 and are associated with initiatives launched by the government that year. Specifically, the government announced the Patient-centered Medical Technology Optimization Research Project and the Strategy for Fostering Research Doctors and Hospital Innovation. It appears that articles related to these projects were prominently featured in 2018. Furthermore, there was a notable increase in articles pertaining to “global healthcare,” rising from 11.4% in 2016 to 21.0% in 2018. This surge can be seen as a response to a press release from the Ministry of Health and Welfare, which reported an uptick in the number of foreign patients visiting Korea in 2018.

“Palliative care” appeared as a topic in 2017 (28 cases, 9.5%) and 2019 (34 cases, 7.9%). These years correspond to significant legislative milestones in the realm of hospice and palliative care, as well as decisions regarding life-sustaining treatment for end-of-life patients. Specifically, a law was enacted on August 4, 2017, and subsequently underwent a partial amendment on March 28, 2019. This legislative activity likely accounts for the increased attention to palliative care in scholarly articles during 2017 and 2019.

In 2020, the landscape of healthcare and the challenges faced by doctors were unlike any other year. After the first case of COVID-19 was confirmed in South Korea, a continuous stream of issues related to the pandemic emerged. Additionally, there were significant disruptions such as collective leave taken by medical professionals due to disputes with the government, resident physician strikes, medical students jointly suspending their studies, and boycotts of the medical licensing examination. Consequently, the most prominent topics related to COVID-19 were “COVID-19: lack of medical resources” (97 mentions, 16.4%), “COVID-19: dedication of medical staff” (84 mentions, 14.2%), and “COVID-19: telemedicine” (86 mentions, 14.5%). These topics stood out in frequency when compared to other issues.

Discussion

To cultivate the doctors needed for society and to train them properly, it is necessary to establish appropriate competencies for physicians. The shift toward competency-based, performance-oriented training systems for residents necessitates the precise definition of these compe-

tencies. With this objective in mind, the patient-centered physician competencies in Korea framework was developed in 2021, supported by the Korea Healthcare Research Institute [6]. As part of the development of this framework, this study employed topic modeling to analyze major Korean newspaper articles. This analysis aimed to capture the expectations and requirements of Korean society regarding doctors and medical care, and the findings were incorporated into the competency framework for doctors.

Based on an analysis of major Korean newspaper articles from the past 5 years, spanning 2016 to 2020, we have identified societal expectations of doctors and healthcare as reflected in these publications. Given the media’s focus, many of the topics discussed in the articles centered on social phenomena, including the expectations placed on doctors, and also covered a range of economic and legal issues.

Topic modeling yielded 18 distinct topics over a 5-year period. We examined the primary keywords associated with each topic and tracked the prevalence of these topics annually. Our analysis of the main keywords revealed an increasing expectation for doctors to possess a broader range of competencies beyond the essential medical knowledge and clinical skills traditionally required. The necessity for doctors to acquire a diverse set of competencies, in addition to foundational skills, has been previously addressed in the literature, and various countries have proposed frameworks for what constitutes basic competencies for physicians [2-4,19,20].

The core competencies of patient-centered physicians in Korea relate to the physician’s role as an expert, communicator, collaborator, healthcare leader, professional, and scholar. Based on these core competencies, 15 of the 18 identified topics can be included in the competencies of healthcare leaders, professionals, and communicators. Eleven of the topics (“digital healthcare,” “global healthcare,” “telemedicine,” “publicness in healthcare,” “healthcare system,” “healthcare policy,” “conflict between medical professionals and the government,” “communicable disease control,” “COVID-19: lack of medical resources,” “COVID-19: dedication of medical staff,” “COVID-19: telemedicine”) could be included in the competencies of healthcare leaders, four of the topics (“healthcare technology,” “patient safety,” “palliative care,” “patient centeredness and good doctors”) could be included in the competencies of professionals, and two of the topics (“palliative care,” “patient centeredness and good doctors”) could be included in the competencies of communicators. These findings are similar to previous studies that have shown that the public’s expectations and satisfaction with doctors’ social competencies include poor communication skills with patients and caregivers and lower-than-normal ratings of leadership skills and responsibility as a member of society [21].

In addition, three of the 18 topics—“convergence,” “medical science research,” and “medical volunteering”—were found to be socially relevant issues, which can be judged as reflecting social demands for doctors and medical care. In addition, these topics are included in doctors’ competencies in terms of content, but they are not included in the core competencies and detailed competencies of the Korean patient-centered physician competencies. The reason for this exclusion is that the defined competencies of doctors are restricted to the core competencies that all physicians should possess and to areas that are amenable to education and evaluation. It is anticipated that the inclusion of these topics in the competency framework could be reconsidered following further research.

Over the 5-year period from 2016 to 2020, “digital healthcare” and “publicness in healthcare” were identified as topics each year. Digital healthcare, being closely linked to industrial potential as well as the health and lives of people, has received significant government attention. In response, the government has developed an industrial policy and implemented various strategies, including the 2017 Healthcare Development Strategy, which is aligned with the Fourth Industrial Revolution. Consequently, digital healthcare has been frequently featured in newspaper reports. Publicness in healthcare has gained increased media attention, particularly during the COVID-19 pandemic. This concept focuses on reducing health and healthcare utilization inequalities by addressing gaps in healthcare for medically vulnerable populations and regions. It also aims to establish a safety net capable of effectively responding to national disasters, catastrophes, and emergencies. In the patient-centered doctor’s competency framework in Korea, digital healthcare is incorporated as a sub-competency under the core competencies of healthcare leaders, preparing them for future changes. Similarly, publicness in healthcare is addressed within the sub-competencies of social activities for promoting health and enhancing equity in healthcare.

The 18 topics were analyzed based on the proportion of topics that emerged each year, along with keywords associated with policies or projects implemented in that year, or issues that arose within the medical field. Given the nature of newspapers to reflect contemporary social issues and political challenges, it is not feasible to precisely determine the competencies society expects from doctors through keyword analysis of newspaper articles [22]. However, it is important to pay attention to the topics identified in newspaper articles, particularly those that have persisted over an extended period, those that have faded without resolution, or those that are newly emerging. This analysis can provide insight into the issues that doctors should be aware of within society and the social phenomena currently affecting the Korean medical community. It can also stimulate and refine discourse on doctors’ potential roles in society.

The importance of this discussion is heightened in the information

age, where the general public frequently acquires medical knowledge through newspaper articles, news broadcasts, and social networking services (SNS). Consequently, there is an increasing need for physicians to engage with a variety of societal issues and phenomena [23-26]. With the rise in smartphone usage and the proliferation of social media, it has become crucial for doctors to communicate not only with individual patients but also with the wider public and media outlets [27,28]. Effective communication with society at large is now recognized as a key competency for physicians. Specifically, doctors should serve as informed experts who can evaluate various media content using their professional knowledge, helping to prevent the spread of indiscriminate or biased medical information. This necessitates discussions and education regarding the roles doctors can play within society.

Limitations of this study include the fact that we analyzed newspaper articles over a 5-year period from 2016 to 2020, that we were unable to analyze all newspapers in Korea, and that the categorization of topics was based on the subjective views of the researchers.

This study was conducted to examine the feasibility of the competency-oriented, outcome-based patient-centered doctor’s competency framework in Korea, which was developed to improve the resident training system. This study used cutting-edge research methods to explore the societal demands for physicians using a variety of media to help define the competencies of physicians. Continued research on this topic can help define and improve the competencies of physicians needed in the future. Furthermore, the physician competencies explored in this study can be used as evidence for the development of medical education curricula.

ORCID

Hanna Jung	https://orcid.org/0000-0001-5051-3953
Jea Woog Lee	https://orcid.org/0000-0001-8112-407X
Geon Ho Lee	https://orcid.org/0000-0003-0696-3804

Conflict of interest

Hanna Jung is an Editorial Board member of KMER, but was not involved in the peer reviewer selection, evaluation, or decision process of this article. Except for that, no other potential conflict of interest relevant to this article was reported.

Authors’ contribution

Hanna Jung: collected data, wrote the manuscript; Jea Woog Lee:

data analysis; Geon Ho Lee: critical revision of the article; and final approval of the version to be published.

References

- Albanese MA, Mejicano G, Mullan P, Kokotailo P, Gruppen L. Defining characteristics of educational competencies. *Med Educ.* 2008;42(3):248-55. <https://doi.org/10.1111/j.1365-2923.2007.02996.x>
- Edgar L, McLean S, Hogan SO, Hamstra S, Holmboe ES. The Milestones guidebook. Chicago (IL): Accreditation Council for Graduate Medical Education; 2020.
- Frank JR, Snell L, Sherbino J. CanMEDS 2015 Physician Competency Framework. Ottawa (ON): Royal College of Physicians and Surgeons of Canada; 2015.
- General Medical Council. Good Medical Practice. London: General Medical Council; 2020.
- Lee SW. Concept and development of resident training program for general competencies. *Korean Med Educ Rev.* 2017;19(2):63-9. <https://doi.org/10.17496/kmer.2017.19.2.63>
- Jeon WT, Jung H, Kim YJ, Kim C, Yune S, Lee GH, et al. Patient-centered doctor's competency framework in Korea. *Korean Med Educ Rev.* 2022;24(2):79-92. <https://doi.org/10.17496/kmer.2022.24.2.79>
- Jung YJ, Kim JH, Yoo JW. Research on the process to a popularization of dancesports through the mass media. *Korea J Phys Educ.* 2003;42(3):355-64.
- Kim HK. A study on the social recognition of library [dissertation]. Seoul: Chung-Ang University; 2011.
- Clarke J. Portrayal of childhood cancer in English language magazines in North America: 1970-2001. *J Health Commun.* 2005;10(7):593-607. <https://doi.org/10.1080/10810730500267605>
- Na K, Lee J. Trends of South Korea's informatization and libraries' role based on newspaper big data. *J Korea Contents Assoc.* 2018;18(9):14-33. <https://doi.org/10.5392/JKCA.2018.18.09.014>
- Han S. An analysis of news trends for libraries in Korea: based on 1990-2018 news big data. *J Korean Soc Inf Manag.* 2019;36(3):7-36. <https://doi.org/10.3743/KOSIM.2019.36.3.007>
- Cho KW, Bae SK, Woo YW. Analysis on topic trends and topic modeling of KSHSM Journal papers using text mining. *Korean J Health Serv Manag.* 2017;11(4):213-24. <https://doi.org/10.12811/kshsm.2017.11.4.213>
- Yang YH. Analysis on types and trends of public conflicts using topic modeling. *Korea Local Adm Rev.* 2021;35(2):159-88. <https://doi.org/10.20484/klog.23.3.18>
- Kang B, Song M, Jho W. A study on opinion mining of newspaper texts based on topic modeling. *J Korean Soc Libr Inf Sci.* 2013;47(4):315-34. <https://doi.org/10.4275/KSLIS.2013.47.4.315>
- Blei DM, Ng AY, Jordan MI. Latent Dirichlet allocation. *J Mach Learn Res.* 2003;3:993-1022.
- Zhang R, Pakhomov S, Gladding S, Aylward M, Borman-Shoap E, Melton GB. Automated assessment of medical training evaluation text. *AMIA Annu Symp Proc.* 2012;2012:1459-68.
- Borakati A. Evaluation of an international medical E-learning course with natural language processing and machine learning. *BMC Med Educ.* 2021;21(1):181. <https://doi.org/10.1186/s12909-021-02609-8>
- Moon H, Kim S, Lee J, Cheon J. Examination of change in perception toward virtual medical education after COVID-19 pandemic in the U.S. using Twitter data. *Proceeding of the Annual Convention of the Association for Educational Communications and Technology*; 2021 Nov 2-6; Chicago, USA. Bloomington (IN): Association for Educational Communications and Technology; 2021. p. 172-80.
- Carraccio C, Wolfsthal SD, Englander R, Ferentz K, Martin C. Shifting paradigms: from Flexner to competencies. *Acad Med.* 2002;77(5):361-7. <https://doi.org/10.1097/00001888-200205000-00003>
- Shah N, Desai C, Jorwekar G, Badyal D, Singh T. Competency-based medical education: an overview and application in pharmacology. *Indian J Pharmacol.* 2016;48(Suppl 1):S5-9. <https://doi.org/10.4103/0253-7613.193312>
- Heo YJ. Doctors' competency and empowerment measures desired by the state and society. *J Korean Med Assoc.* 2014;57(2):121-7. <https://doi.org/10.5124/jkma.2014.57.2.121>
- Collins PA, Abelson J, Pyman H, Lavis JN. Are we expecting too much from print media?: an analysis of newspaper coverage of the 2002 Canadian healthcare reform debate. *Soc Sci Med.* 2006;63(1):89-102. <https://doi.org/10.1016/j.socscimed.2005.12.012>
- Dutta-Bergman MJ. Primary sources of health information: comparisons in the domain of health attitudes, health cognitions, and health behaviors. *Health Commun.* 2004;16(3):273-88. https://doi.org/10.1207/S15327027HC1603_1
- Dutta-Bergman MJ. Developing a profile of consumer intention to seek out additional information beyond a doctor: the role of communicative and motivation variables. *Health Commun.* 2005;17(1):1-16. https://doi.org/10.1207/s15327027hc1701_1
- Prosser H. Marvelous medicines and dangerous drugs: the representation of prescription medicine in the UK newsprint media. *Public Underst Sci.* 2010;19(1):52-69. <https://doi.org/10.1177/0963662508094100>

26. Furstrand D, Pihl A, Orbe EB, Kingod N, Sondergaard J. "Ask a doctor about coronavirus": how physicians on social media can provide valid health information during a pandemic. *J Med Internet Res*. 2021;23(4):e24586. <https://doi.org/10.2196/24586>
27. Kim CJ, Kwon I, Han HJ, Heo YJ, Ahn D. Korean doctors' perception on doctor's social competency: based on a survey on doctors. *J Korean Med Assoc*. 2014;57(2):128-36. <https://doi.org/10.5124/jkma.2014.57.2.128>
28. Smailhodzic E, Hooijsma W, Boonstra A, Langley DJ. Social media use in healthcare: a systematic review of effects on patients and on their relationship with healthcare professionals. *BMC Health Serv Res*. 2016;16(1):442. <https://doi.org/10.1186/s12913-016-1691-0>

APPENDICES

Appendix 1. Main keywords by topic from 2016 to 2020

토픽	주요 키워드				
	2016	2017	2018	2019	2020
의료산업					
디지털헬스케어	왓슨, 정밀, 인공지능, 헬스케어, 진단, 원격 의료, 영상, 원격, 플랫폼	진단, 첨단, IoT, 맞춤형, 왓슨, AI 빅데이터, 로봇	플랫폼, 융합, 허브, AI, 증강현실, 로봇, IoT, 데이터, 왓슨, 셀스케어, 플랫폼, 정밀, 채택, 스마트	AI, 인공지능, 로봇, 스마트, 왓슨, 영상, 판독, 예측, 진단, 혁신, 헬스케어, 디지털, 딥러닝, 4차산업	데이터, 진단, AI, 영상, 판독, 원격진료, 인공지능, 디지털헬스케어, 스마트, 플랫폼, 판독
글로벌헬스케어	진출, 세계, 선도, 시장, 유치, 글로벌, 해외, 의료관광, 중국, 세계, 네트워크, 몽골, 협약, 교류, 동남아	-	후진국, 전파, 러시아, 의료수출, 진출, 연수, 해외, 한류, 유치, 방문, 의료시장, 해외환자, 미국	해외, 진출, 교류, 미얀마, 바레인, 러시아, 교류, 중동, 관광, 우즈베크, 잠비아, 중동, 방글라데시	-
의료기술	내시경, 웨어러블, 의료기기, 정밀, 병변위치, 최소침습, 복강경, 재생, 기술	-	-	이식, 내시경, 바이오마커, 인공와우, 안면거상술, 생체이식, 외상센터, 위절제, 종양	-
융합	-	협진, 의공학, 발전, 다학제, 융합, 첨단, 공학	-	-	-
의과학연구	-	-	과학자, 양성, 연구개발, 연구의사, 연구중심병원, 기술개발, 다학제	-	-
원격의료	-	-	-	-	원격, 산업, 디지털, 온라인, 센서
보건의료정책					
의료의 공공성	공공의대, 설립, 공공, 보편, 인력, 공공의료	공공, 공공의료, 노인, 장기요양, 치매, 복지, 치매케어	영리화, 공공의료, 의료복지, 공공의대, 의교격차	공공, 복지, 공중보건, 장학제도, 취약	공공, 파업, 전공의, 의협, 지자체, 민간
의료체계	농촌, 보건소, 의약품, 시설, 네트워크, 서비스, 지역의료	지역의료, 1차의료, 보건, 건강관리, 국민주치의, 예방, 치매, 장기요양, 노인	주치의제도, 왕진, 치매, 용양원, 공공후견, 국가책임제, 노인, 왕진, 1차의료, 동네의사	-	-
의료정책	-	-	-	공공, 건강보험, 진료비, 보건, 기관	중증센터, 건보료, 닥터헬기, 협동조합, 입원전담, 응급실, 취약계층, 메디케어, 공공
의정사태	-	-	-	-	집단휴진, 파업, 집단행동, 전공의, 의협, 공백, 선동, 시험, 거부

(Continued on next page)

Appendix 1. Continued

토픽	주요 키워드				
	2016	2017	2018	2019	2020
환자중심의료					
의료봉사	선교, 봉사, 국경없는 의사회, 아프리카, 종교	의료봉사, 구호, 무료, 선교, 해외, 무료병원, 봉사, 복지관	-	봉사, 방글라데시, 난민, 빈민, 기부, 감비아, 아프간, 베트남, 농촌	봉사, 에티오피아, 아프리카, 모로코, 이태석, 아프가니스탄, 심장센터, 가난, 캄보디아
환자안전	백남기, 불법, 사망, 의약품, 위반, 의료법, 윤리, 의료분쟁, 무허가, 법원	불법시술, 의료분쟁, 생명윤리, 조작, 직업윤리, 소송, 대리수술, 낙태, 면허	-	조정, 분쟁, 중재, 사고, 소송, 절차, 징계, 중재원, 과실, 민사, 판결, 정지, 처분, 쇼닥터	-
완화의료	-	호스피스, 죽음준비, 죽음, 말기암, 완화의료, 중단	-	죽음, 호스피스, 연명의료, 안락사, 존엄, 임종, 완화, 말기	-
환자중심 및 좋은의사	-	-	토털케어, 환자중심, 안전, 의사소통, 좋은 의사, 의료질, 맞춤진료	-	-
감염병					
감염병방역	메르스, 감염병, 대응, 방안, 대책, 부실, 방역체계	-	-	-	-
코로나19: 의료자원 부족	-	-	-	-	의료진, 부족, 공백, 긴급, 병상, 인력, 한계, 경영난, 번아웃
코로나19: 의료진 헌신	-	-	-	-	희생, 헌신, 봉사, 의료진, 의료봉사, 자발, 인력, 부족, 긴급, 협업, 협조, 파견, 지원
코로나19: 원격의료	-	-	-	-	원격, 원격진료, 비대면, 진단, 규제, 허용, ICT, 대응, 감염, 안전