

Comparative Analysis of News Big Data related to SARS-CoV, MERS-CoV, and SARS-CoV-2 (COVID-19)

Jae-Hyun Woo*

*Senior Researcher, Dept. of Medical Devices Industry, Dongguk University, Seoul, Korea

[Abstract]

This paper intends to draw implications for preparing for Post-Corona in the health field and policy fields as the global pandemic is experienced due to COVID-19. The purpose of this study is to analyze the news and trends of media companies through temporal analysis of the three infectious diseases, SARS-CoV, MERS-CoV, and SARS-CoV-2 (COVID-19), in which the domestic infectious disease preventive system was active throughout the first year of the outbreak. To this end, by using the news analysis program of the Korea Press Foundation 'Big Kinds', the number of news articles per year was digitized based on the period when each infectious disease had an impact on Korea, and major trends were implemented and analyzed in a word cloud. As a result of the analysis, the number of articles related to infectious diseases peaked when the World Health Organization (WHO) declared a warning and (suspicious) confirmed cases occurred. According to keyword and word cloud analysis, 'infectious disease outbreak and major epidemic areas', 'prevention authorities', and 'disease information and confirmed patient information' were found to be the main common features, and differences were derived from the three infectious diseases. In addition, the current status of the infodemic was identified by performing word cloud analysis on information in uncertainty. The results of this study are significant in that they were able to derive the roles of the health authorities and the media that should be preceded in the event of a new disease epidemic through previously experienced infectious diseases, and areas to be rearranged.

▶ **Key words:** SARS, MERS, COVID-19, NEWS, BIGKINDS

-
- First Author: Jae-Hyun Woo, Corresponding Author: Jae-Hyun Woo
 - *Jae-Hyun Woo (highambition@naver.com), Dept. of Medical Devices Industry, Dongguk University
 - Received: 2021. 06. 03, Revised: 2021. 08. 24, Accepted: 2021. 08. 24.

[요 약]

본 논문은 COVID-19로 인해 세계적인 팬데믹(Pandemic)을 경험하게 되면서 보건 분야, 정책 분야 등에 있어 포스트 코로나(Post-Corona)를 준비하기 위한 시사점을 도출하고자 한다. 국내 감염병 방역체계가 가동되었던 SARS-CoV, MERS-CoV, SARS-CoV-2(COVID-19)의 3개 감염병에 대해 발병 1년간의 시기적인 분석을 통해 언론사 뉴스 및 트렌드를 분석해보자는 것이다. 이를 위해 한국언론진흥재단 ‘빅카인즈’ 뉴스 분석 프로그램을 활용하여 각 감염병이 국내에 영향이 미치던 시기를 기준으로 1년간의 뉴스 기사 건수를 수치화하고 주요 트렌드를 워드클라우드로 구현하여 분석하였다. 분석 결과, 감염병과 관련한 기사 건수는 세계보건기구(WHO)의 경고 선언 및 (의심)확진자 발생 시점에 정점을 기록하였다. 키워드와 워드클라우드 분석에 따르면 감염병에 대한 ‘발병지 및 주요 유행지역’, ‘방역당국’, ‘질병정보 및 확진자 정보’ 등이 주요한 공통점으로 나타났다. 3개 감염병에서 차이점을 도출하였다. 아울러, 불확실 정보에 대하여 워드클라우드 분석을 수행함으로써 인포데믹 현황을 파악하였다. 본 연구결과는 앞서 경험하고 있는 감염병을 통해서 새로운 질병이 대유행할 시 선행되어야 하는 보건당국, 언론의 역할 및 재정비되어야 할 영역을 도출할 수 있었다는 점에서 의의를 갖는다.

▶ **주제어:** SARS, MERS, COVID-19, 뉴스, 빅카인즈

I. Introduction

2020년 세계보건기구(World Health Organization)는 2019년 처음 존재가 확인된 이후 현재 세계를 강타 중인 COVID-19의 전 세계적 대유행(pandemic)을 선언하였다 [1]. 이는 앞서 2020년 1월 30일에 국제보건 규칙에 근거하여 선포한 ‘국제 공중보건위기상황’과는 별도의 높은 수준의 경보이다[2].

현재까지 진행 중인 COVID-19(SARS-CoV-2, 2019)는 인간에게 중증의 급성호흡기 감염을 유발하는 중증 급성 호흡기 증후군 코로나 바이러스로서, 이전에 국내에도 유행했던 중증 급성 호흡기 증후군 코로나바이러스(severe acute respiratory syndrome coronavirus, SARS-CoV, 2003), 중등 호흡기 증후군 코로나바이러스(Middle East respiratory syndrome coronavirus, MERS-CoV, 2012) 등과 동일한 베타코로나 바이러스에 속한다[3-8].

이러한 질병들은 전염성이 매우 높아 사람들의 일상생활 및 경제활동에 상당한 영향을 미치고 있으며, 특히 COVID-19는 강한 전염성으로 인해 전 세계적으로 보건의료체계 붕괴와 경제성장을 둔화 시켜 우리가 많은 상황이다. 또한 COVID-19의 확산은 금융, 관광, 물류·유통, 교통, 문화, 요식업 등과 같이 제공 주체의 변환 프로세스에서 고객과의 대면 과정을 통해 프로세스 산출물이 전달되는 서비스 산업에 막대한 피해를 주고 있다. 이와 같은 상황을 종합해보면, 서비스업이나 제조업을 막론하고 산출물

을 생산하는 과정에서 사람이 관여하는 정도가 높은 프로세스의 산업일수록 COVID-19로 인한 피해는 상대적으로 높다고 볼 수 있다[9-10].

이처럼, 감염병의 대 확산은 방역시스템을 포함하여 각 산업계를 위한 정부의 역할 등이 높게 연계될 수밖에 없다. 이에 따라, 정부는 국내 발병 초기부터 K-방역 체계를 가동하여 감염자를 검사로 빠르게 찾아내고, 확진자의 동선을 ICT 기술을 동원하여 추적하는 방식으로 연쇄 감염의 고리를 차단하며, 환자를 제때에 격리하고 치료하는 방역 기법인 이른바 3T(Test-Trace-Treat)를 실행하고 있다[11]. 아울러, 국내에서는 2020년 3월 23일 사회적 거리두기를 시작으로 2020년 11월 7일 사회적 거리 두기 개편을 통해 코로나19 공존 시대에 지속 가능한 거리 두기 체계를 마련하여 운영하고 있으며[12-13], 백신·치료제 개발 기업에 대한 정부 R&D 지원과 중장기적 관점에서의 보건 의료체계 및 관련 법적·제도를 운용 및 확대해 나아가고 있다[14-16]. FDA는 South Korea's Response to COVID-19" 보고서를 발간하면서 한국의 국가 대응 전략 등을 분석하여 제공하였으며, 이러한 감염병 대응 체계들은 사스(2003), 메르스(2015)를 겪으면서 얻은 교훈을 통해 기초의 틀을 다져놓았기에 가능하다고 분석하였다[17].

그러나 COVID-19는 기존 2개 감염병과는 달리, 무증상 확진 및 장기화되는 특징점 등에 따라 현행 제도의 공백 영

역, 불확실 정보에 대한 이슈[1, 18]가 상당수 급증한 양상을 보이고 있으며, COVID-19를 경험하면서 실행된 새로운 정책 및 제도의 보완 요구 또는 정보의 검증 연구의 필요성이 간간히 연구되고 있다[2, 15, 19-21]. 그러나 선행 연구들에서도 개별 감염병에 대한 이슈 및 시사점을 조망하는 연구가 주를 이룬다는 한계점이 있다. 이에 따라, 우리가 앞서 경험한 3개 감염병을 통하여 과거와 현재를 종단할 수 있는 시계열분석을 수행함으로써 위드코로나(With-Corona) 및 포스트코로나(Post-Corona)에 반영하여 진보적인 방향으로 나아갈 수 있는 시사점 도출이 필요한 시점이다.

한편, 이러한 주요 과정과 결과들을 포함한 정보 및 정책 등은 핵심적인 시각의 틀을 제공할 수 있다는 점에서 일반 대중에게 언론 기사를 통해 제공된다[22-23]. 언론사 뉴스 기사는 특정 분야 및 정책 등을 둘러싼 쟁점을 분석할 때 대표적으로 활용되는 원자료이며, 관련 주요 이슈를 독자들에게 전달함으로써 해당 분야의 정책 및 행동론에 대한 관심 및 이해 수준에 큰 영향을 준다[22]. 이에 따라, 21세기 오늘날은 다양한 분야에서 빅데이터 분석이 시도되면서 언론사 뉴스 또한 주요 분석 대상으로 활용되고 있으며, 앞서 제시한 문제를 해결하기 위해 이를 활용하여 분석을 수행하고자 한다. 빅카인즈(BIGKinds)를 이용한 뉴스 추출 방식은 전통적인 프레임 분석 내 주관성 개입과 많은 작업 시간의 단점들을 보완하는 방식으로써 최근 다양한 분야에서 활용하고 있으며, 특히 방대한 뉴스의 양을 통해 연구의 질적 수준을 높이고, 자연어 처리 방식을 활용함에 따라 코더 간 신뢰도를 측정할 필요가 없다는 강점을 가지고 있다[24-26]. 이를 통해, 본 연구에서는 국내 감염병 방역체계가 가동되었던 SARS-CoV, MERS-CoV, SARS-CoV-2(COVID-19)의 3개 감염병에 대해 발병 1년간의 시기적인 분석을 통해 언론사 뉴스 및 주요 트렌드를 분석하여 주요 공통사항 및 특징점을 도출하고자 한다. 아울러, 뉴스 데이터 분석을 통하여 감염병에 대한 불확실한 정보의 공유 현황을 고찰함으로써 정책에 환원할 수 있는 종합적인 시사점을 도출하는 것을 목적으로 하고자 한다.

본 연구의 목적 달성을 위해 설정된 구체적인 연구 문제들은 다음과 같다.

첫째, 3개 감염병 관련 보도량의 차이가 있는가?

둘째, 각 언론사의 3개 감염병 관련 시기별 뉴스 트렌드는 무엇인가?

셋째, 각 언론사 간의 3개 감염병 관련 시기별 주요 키워드 및 워드 클라우드에 차이가 나타나는가?

넷째, 감염병 관련 주요 불확실한 정보는 무엇이 있었는가? 이다.

II. Preliminaries

1. Disease information

국내 유행했던 중증 호흡기 감염 질환은 코로나바이러스의 감염과 관련된 임상 증후군으로 상부 호흡기 질환, 폐렴, 급성 호흡곤란 등 중증도가 가변적인 호흡기 증후군이 공통적인 특징이다[6].

SARS는 2002년 11월 중국 광둥성에서 전에 알려지지 않았던 새로운 코로나바이러스(novel coronavirus)가 등장하여 최초로 발병하였다. 국내에서는 2003년 4월 10일에 첫 사례 발생 후 114일간 3명의 추정 확진자를 기록하며 종식되었다. MERS는 2012년 4월 사우디아라비아에서 처음 보고되었으며 국내에서는 2015년 5월 20일에 첫 확진자가 발생하였다. 218일간 국내에 영향을 미치며 186명의 누적 확진자를 기록한 후 종식되었다. COVID-19는 2019년 12월 중국 후베이성 우한시에서 최초 보고되었으며, 국내에서는 2020년 1월 20일에 첫 확진자가 발생하였다. 해당 감염병은 2020년 12월까지 누적 확진자 62,593명을 기록하였으며, 현재 시점인 2021년 7월 31일에도 지속하여 유행 중이다[15, 27].

세계보건기구(WHO)는 전염병 경보 체계를 6단계로 구분하고 있는데, 높은 수준 단계로 사람 간 빠르게 전염병이 확산되는 국지적 상태를 에피데믹(4단계), 세계적 대유행을 팬데믹(5~6단계)로 정의하고 있다[28]. 우리가 경험한 사스와 메르스는 에피데믹, COVID-19는 팬데믹 경보가 발령된 바 있다. 국내에서는 위기 경보로 발령되고 있으며, 「재난 및 안전관리 기본법」에 의거하여 관심, 주의, 경계, 심각 단계로 구분하고 있다. 상기 법이 제정된 이후의 감염병인 MERS는 주의단계, COVID-19는 심각 단계까지 발령되었다[29-31].

Table 1. Comparison table of disease information

국내 발생현황	SARS	MERS	COVID-19
최초 확진자 발생 일자	2003.04.10.	2015.05.20.	2020.01.20.
최상위 경보단계 일자 (추정)	2003.03.16.(경보)	2015.05.20(주의)	2020.02.23(심각)
확진자 수	3명	186명	62,593명 (20.12.기준)
유행기간	114일	218일	지속유행 (21.07.기준)

Source: Korea Disease Control and Prevention Agency

2. News Big Data Analytics Using BIGKINDS

빅카인즈 서비스는 종합일간지, 경제지, 지역일간지, 방송사 등을 포함한 국내 최대의 기사 DB(총 54개 언론사)에 빅데이터 기술을 접목한 뉴스 분석 시스템이다. 빅카인즈 서비스 시스템은 기본적으로 비정형 텍스트로 구성된 뉴스 정보를 정형화된 데이터로 전환하여 관련 사회현상을 다양하게 분석할 수 있는 기초 자료로 활용하는 데 도움을 준다 [32]. 빅카인즈 서비스의 개념도는 텍스트 기사, 사진, PDF 등의 뉴스 수집 시스템과 형태소, 개체명, 네트워크를 활용한 뉴스 분석 시스템, 뉴스 저장 시스템으로 구성되어 있다. 빅카인즈에서는 텍스트 마이닝의 분석 결과 및 시각화의 결과물을 ‘키워드 트렌드’, ‘연관어 분석’ 등으로 제공해 준다. ‘키워드 트렌드’는 검색한 키워드가 포함된 뉴스 건수를 연간, 월간, 일간 등 기간별로 확인하는 방법이며, ‘연관어 분석’은 검색 결과 중 분석 뉴스와 연관성(가중치, 키워드 빈도수)이 높은 키워드를 최대 1,000건 내에서 시각화하여 보여 준다[26]. 가중치 정보는 사용자가 입력한 질의어에 기반하여 검색한 결과에 대해 공식적 분석과 워드 클러스터링을 하여 동적으로 상호 작용하는 시맨틱 네트워크를 생성하는 토픽랭크 알고리즘에 기반해 출력된 결과이며, 키워드 빈도수는 가중치 순으로 선정된 연관어를 해당 연관어가 등장한 횟수 기준으로 재정렬한 결과이다[21].

III. The Research Methodology

1. Data collection target

본 연구는 세계적 호흡기 질환 대유행과 관련된 언론사들의 뉴스 기사들을 수집하기 위하여, 한국언론진흥재단 (Korea Press Foundation) 빅카인즈(Bigkinds) 뉴스 분석 서비스[26]를 활용하였다. 본 연구의 데이터 수집 기간은 SARS-CoV에 대해서 2003년 1월에서 12월까지, MERS-CoV는 2015년 1월에서 12월까지, COVID-19는 2020년 1월에서 12월까지로 설정하였다. 이 같은 데이터 수집 기간 설정은 국내에 대유행한 시기와 관련된 언론사 뉴스 보도의 양상을 비교하기 위해서이다.

2. Major research methods

본 연구에서 주요하게 활용한 구체적인 연구 방법은 빅카인즈의 뉴스 검색과 키워드 트렌드 분석, 워드클라우드 생성 등이었다. 첫째, 뉴스 검색에서는 분석 대상 기간 (2003년, 2015년, 2020년 분리)을 설정한 후, "SARS", "SARS-CoV", "사스"와 "MERS", "MERS-CoV", "메르스"

그리고 세계보건기구(WHO)에서 공식 명칭을 지정하기 이전에 통용되던 "우한 폐렴"을 포함하여, "SARS-CoV-2", "COVID-19", "코로나"로 검색을 진행하였다. 둘째, 키워드 트렌드는 검색한 키워드와 관련된 뉴스 건수를 기간별로 확인하는 방법이다. 본 연구에서는 월간(Monthly) 단위로 관련 뉴스 건수를 분석하였다. 셋째, 검색 용어와 관련된 연관어 분석을 진행하였다. 본 연구에서는 상기 키워드들을 기준으로 언론사들의 뉴스 기사 내용에 포함된 단어들의 빈도와 가중치를 분석하였다. 가중치는 분석 뉴스 내에서 토픽 랭크 알고리즘을 기준으로 유사도가 높은 키워드 순으로 표출된다. 넷째, 관련 연관어 분석 결과는 가장 시각적이며 직관적으로 가독성이 높은 워드클라우드 형식으로 비교 제시하였다[32-34]. 마지막으로 추출한 뉴스 데이터 내에서 불확실한 정보를 분석하였다. 단, 불확실한 정보라는 개념은 가짜뉴스, 허위사실, 검증되지 않은 정보 등이 혼용되어 사용되고 있으며, 명확한 정의가 정립되지 않은 단계임에 따라[19, 35-36], 법률용어인 "허위사실", "방송통신위원회의 권고 용어인 "정정보도", 언론에서 활용되고 있는 "팩트체크"로 국한하여 검색함으로써 해당 감염병에 대해 직관적으로 비교할 수 있는 워드클라우드 분석을 수행하였다.

Table 2. Search terms and set period

구분	검색어	기간
SARS	"SARS", "SARS-CoV", "사스"	2003-01-01 ~ 2003-12-31
MERS	"MERS", "MERS-CoV", "메르스"	2015-01-01 ~ 2015-12-31
COVID-19	"우한 폐렴", "SARS-CoV-2", "COVID-19", "코로나"	2020-01-01 ~ 2020-12-31
불확실한 정보	"허위사실", "정정보도", "팩트체크"	감염병 기간별

IV. Research Results

1. News number comparison

: SARS vs MERS vs COVID-19

본 연구는 언론사 뉴스 기사 빅데이터 시스템인 빅카인즈 서비스를 이용하여, "SARS", "SARS-CoV", "사스"와 "MERS", "MERS-CoV", "메르스" 그리고 "우한 폐렴", "SARS-CoV-2", "COVID-19", "코로나" 의 키워드를 설정한 기간으로 검색하여 비교하였다.

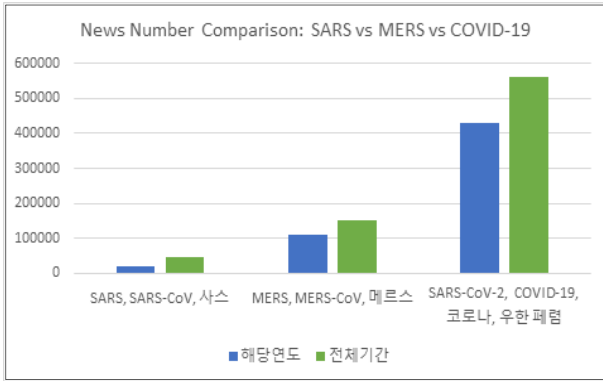


Fig. 1. News Number Comparison : SARS vs MERS vs COVID-19

SARS는 2003년 18,430건의 뉴스가 검색되었으며, 비교적 국내에 영향이 높았던 MERS 및 COVID-19는 각각 110,396건과 417,242건을 기록하였다.

2. News trend(monthly) comparison : SARS vs MERS vs COVID-19

뉴스 트렌드는 관련 검색어를 포함한 뉴스 건수의 변동을 말한다. 본 연구에서는 해당 감염병에 대하여, 설정한 기간 동안의 주요 키워드를 포함한 언론사의 뉴스 건수를 월 간 단위 변동 상황을 비교하여 분석하였다.

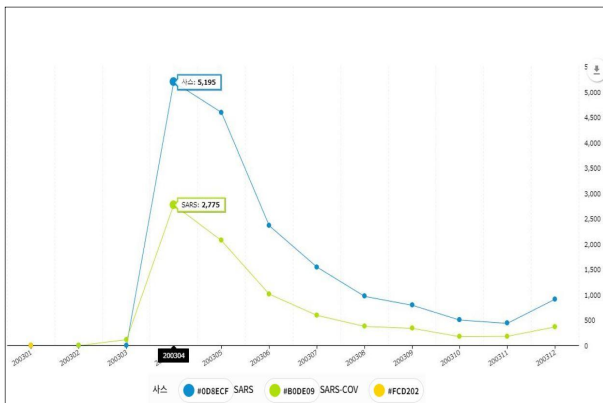


Fig. 2. News Trend on SARS

SARS-CoV는 국내에 3월에서 4월까지 뉴스 건수가 폭발적으로 증가한 월간이었다. 해당 기간에는 세계보건기구(WHO)의 경고 메시지, 국내 사스 감시체계 작동, 국내 첫 의심 환자 발생 등이 뉴스 트렌드의 주요한 원인이었다. SARS-CoV의 키워드는 전체 기간 중 0건을 기록하였다.

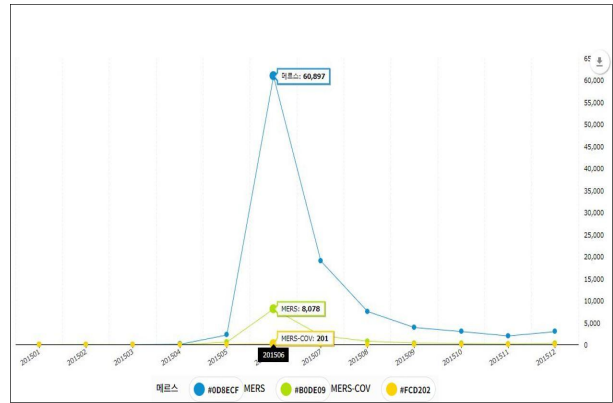


Fig. 3. News Trend on MERS

MERS-CoV는 국내에 5월에서 7월까지 뉴스 건수가 폭발적으로 증가한 월간이었다. 해당 기간에는 국내 첫 확진자 발생, 방역체계 가동 등이 뉴스 트렌드의 주요한 원인이었다.

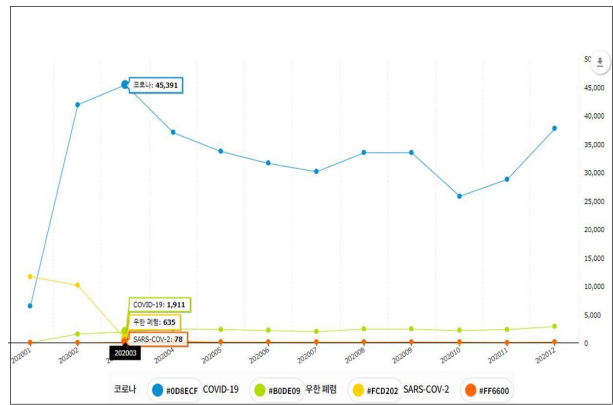


Fig. 4. News Trend on COVID-19

COVID-19는 국내에 1월에서 2월까지 뉴스 건수가 폭발적으로 증가한 월간이었으며, 3월에 정점을 기록한 후 꾸준한 뉴스 건수를 나타내었다. 1월과 2월 사이에는 국내 첫 확진 환자가 발생하였으며, 3월에는 세계보건기구(WHO)의 팬데믹(Pandemic) 선언과 국내 감염병 확산 추세 등이 뉴스 트렌드의 주요한 원인이었다.

3. Keywords comparison : SARS vs MERS vs COVID-19

이 부분에서는 2번 및 3번 연구 문제에 대한 답을 제시한다. 2003년, 2015년, 2020년으로 설정한 기간 내 언론사 뉴스 기사 제목 및 내용에서 검색 키워드를 모두 합산한 주요 핵심어(키워드)를 추출하여 제시한다. 3개 감염병 질환의 상위 출현 빈도 키워드를 비교함으로써 종단적인 분석 작업이 진행되었다.

Table 3. Keyword frequency and weight(SARS)

Rangking	Keyword	Frequency	Weight
1	중국	1544	223.82
2	홍콩	935	154.16
3	급성호흡기증후군	814	58.36
4	대만	607	97.06
5	베이징	563	96.01
6	SARS·중증	477	82.99
7	감염자	428	87.06
8	SARS·중증급성 호흡기증후군	401	558.19
9	의심환자	375	110.67
10	싱가포르	347	52.3
11	세계보건기구	300	132.64
12	국립보건원	241	122.68
13	추정환자	235	76.56
14	캐나다	224	33.15
15	보건당국	167	36.44
16	환자들	122	24.12
17	상하이	118	24.08
18	광등성	104	21.18
19	감염지역	97	21.47
20	전문가들	85	19.33

SARS(2003)에서 추출된 키워드(상위 20위까지)는 주로 주요 대유행 지역명(중국, 홍콩, 대만 등)과 질환명(급성호흡기증후군, 중증급성호흡기증후군 등) 그리고 질병의 대응조직과 환자 정보 등이 주요한 키워드를 나타내었다. "중국"이 가장 높은 키워드 빈도수를 보였으며, "SARS·중증급성호흡기증후군"이 가장 높은 가중치를 나타내었다.

Table 4. Keyword frequency and weight(MERS)

Rangking	Keyword	Frequency	Weight
1	증후군	1189	17.97
2	중등호흡기증후군	1076	572.99
3	확진자	427	71.76
4	의료기관	315	18.7
5	충북	302	29.35
6	사망자	300	72.64
7	보건복지부	294	41.68
8	감염자	263	53.1
9	MERS-CoV	232	119.59
10	국민안심병원	213	27.06
11	보건당국	196	44.83
12	질병관리본부	167	24
13	확진 판정	155	21.47
14	중동지역	149	23.19
15	사스	135	14.4
16	치사율	133	22.06
17	잡복기	126	22.79
18	격리대상자	118	31.9
19	사우디아라비아	114	24.79
20	확산 방지	113	33.52

MERS(2015)에서 추출된 키워드(상위 20위까지)는 주로 질병에 대한 정보(증후군, 중등호흡기증후군 등)와 주요

발병지 및 국내 확산지역(사우디아라비아, 충북), 그리고 질병의 대응조직과 감염자에 대한 정보 및 질병정보 등이 주요한 키워드를 나타내었다. "증후군"이 가장 높은 키워드 빈도수를 보였으며, "중등호흡기증후군"이 가장 높은 가중치를 나타내었다.

Table 5. Keyword frequency and weight(COVID-19)

Rangking	Keyword	Frequency	Weight
1	중국	3996	264.4
2	확진자	1235	213.58
3	사망자	854	134.43
4	바이러스 감염증	575	118.81
5	사스	502	47.29
6	후베이성	439	37.11
7	대구	416	31.12
8	감염자	361	37.24
9	발병	298	31.98
10	확진 판정	295	41.05
11	질병관리본부	278	37
12	보건당국	233	32.05
13	메르스	228	17.29
14	의심 환자	203	34.24
15	태국	197	21.3
16	세계보건기구	188	51.94
17	武漢	164	97.26
18	중국 정부	145	28.16
19	중국 당국	120	24.64
20	청와대	110	13.49

COVID-19(2020)에서 추출된 키워드(상위 20위까지)는 주로 발병 추정지역 정보(중국, 후베이성, 武漢 등)와 감염자 정보(확진자, 사망자, 의심 환자 등), 그리고 국내외 감염병 대응조직과 해외 확산 첫 지역 및 국내에서 확산이 급증한 지역명 등이 주요한 키워드를 나타내었다. "중국"이 가장 높은 키워드 및 빈도수를 나타내었다.

4. Wordcloud comparison of related words
: SARS vs MERS vs COVID-19

워드클라우드 분석 대상의 키워드, 개념 등을 가장 직관적으로 확인시켜주는 기법이다. 워드클라우드는 분석 대상 텍스트 및 언어 정보에서 많이 나타날수록 해당 단어를 크고 잘 보이게 표현하는 방법이다. 3개 감염병에 대한 전체 뉴스 기사의 워드클라우드 결과는 [그림 5~7]과 같으며, SARS에서는 "중국", MERS에서는 "증후군", COVID-19에서는 "중국"이 가장 시각적으로 부각되었다.



Fig. 5. Wordcloud of related words on SARS



Fig. 8. Wordcloud of uncertain information on SARS



Fig. 6. Wordcloud of related words on MERS

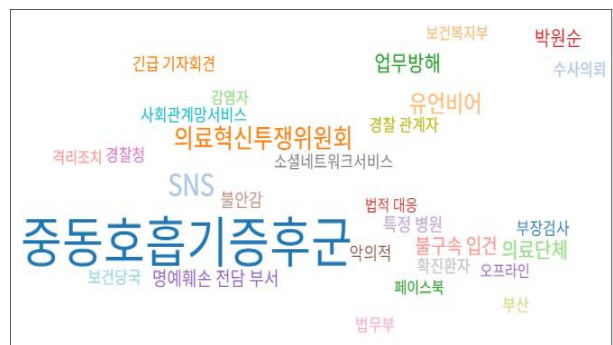


Fig. 9. Wordcloud of uncertain information on MERS



Fig. 7. Wordcloud of related words on COVID-19



Fig. 10. Wordcloud of uncertain information on COVID-19

5. Wordcloud comparison of uncertain information: SARS vs MERS vs COVID-19

불확실한 정보에 대한 뉴스 건수는 SARS(7건), MERS(555건), COVID-19(1,436건)으로 나타났다.

분석 뉴스 내에서 토픽 랭크 알고리즘을 기준으로 의미적 유사도가 높은 키워드 순으로 표출한 워드클라우드 결과는 [그림 8~10]과 같다.

Table 6. Summary of Key Results

구분	SARS	MERS	COVID-19
뉴스 건수	18,430건	110,396건	417,242건
뉴스 트렌드 정점시기	2003.03~2003.04	2015.05~2015.06	2020.02~2020.03
주요 공통 키워드 내용	발병지, 유행지역, 방역당국, 질병정보, 확진자정보, 보건당국		
주요 특징점 키워드 내용	홍콩, 대만, 베이징 등 국외지역 감염병 정보 다수	중동 호흡기 증후군, 중동지역, 치사율	사스-메르스 연계, 의심환자, 청와대, 대구(집단 감염)
주요 불확실한 정보 키워드	정정요구, 악의적 비판기사, 의료단체, 경찰청, 법무부, 가짜뉴스, 유튜브, 구속영장, 방송통신위원회		

V. Conclusions

본 연구는 점차 세계적 감염병의 발병 및 확산에 자유로울 수 없으며 한국이 경험한 주요 호흡기질환에 대한 시사점을 도출하기 위해 수행되었다. 이를 위해 3개 감염병에 대해서 국내 확산에 따른 팬데믹 기간을 1년 단위로 설정하고 각 질병에 대한 비공식 명칭, 원인체 명칭을 주요 키워드로 도출하여 주요 언론사 뉴스 기사들은 어떤 차이점을 가지고 있는지를 뉴스 빅데이터 분석 서비스인 빅카인즈를 활용하여 객관적으로 분석하였다. 연구에 따른 상세 결과는 다음과 같다.

첫 번째, 빅카인즈 서비스를 이용하여 첫 번째와 두 번째 문제를 분석한 결과 3개 감염병의 기사 건수는 18,430(SARS), 110,396(MERS), 417,242(COVID-19)로 큰 차이를 보였다. 이는 국내 감염병의 종식 선언 일자 및 확진자의 유무 및 수치에 따라 정보 확산 목적이 주요 원인으로 사료된다. 또한 감염병별 뉴스 건수가 정점을 기록하는 해당 월은 세계보건기구(WHO)의 팬데믹 등의 경고 선언과 (의심)확진자가 발생한 것으로 나타났다.

두 번째, 3번 문제에 대한 키워드와 워드클라우드 분석에 따르면 감염병에 대한 발병지 및 주요 유행지역, 방역당국, 질병정보 및 확진자 정보 등이 주요한 공통점으로 나타났으며, 3개 감염병에서 상위 빈도수 및 가중치가 높은 것으로 도출된 키워드는 '감염자', '보건당국'이었다.

한편, 국내 확진자가 발생했던 'MERS'와 'COVID-19'에서는 증상 및 예방과 관련한 키워드와 '사망자'가 더욱 빈출되었으며, 'SARS', 'MERS'의 질병정보와 연계하여 분석하는 뉴스 등이 포함되는 차이점을 나타내었다. 특히, COVID-19에서는 세계적 대유행에 따라, 정부의 정책과 방역체계 등과 밀접한 연관성을 보임으로써 '청와대', '중국 정부' 등이 등장하며 지난 2개 감염병과는 차이점을 보였다.

세 번째, 선행연구들을 통하여 설정한 검색어 기반으로 불확실 정보를 분석한 결과, 각 감염병에 대한 전체 뉴스 건수와 비례하는 결과 비중을 보였다. 공통적인 내용으로는 확진자 정보, 질병 대응조직에 대한 내용이 주를 이루며 검증되지 않은 정보에 대한 정정 요구, 악의적 비판 기사에 대한 대응이 있었다. 한편 MERS에서는 의료 단체, 경찰청, 법무부 등의 정부 관련 조직 키워드가 등장함으로써 불확실한 정보에 법적 대응하는 뉴스 및 SNS에 등장하는 허위사실들에 대한 이슈가 새로이 등장하였다.

현재까지 이어지고 있는 COVID-19에서는 MERS 때와는 달리 2.5배 이상의 뉴스 건수를 기록하였으며 가짜뉴

스, 유튜브, 구속영장, 방송통신위원회 등이 새로이 등장하였고, 감염병 관련 가짜뉴스에 대한 정의 및 사회관계망서비스의 키워드 비중이 높은 결과를 보였다.

본 연구결과가 시사하는 바는 다음과 같다.

첫째, 감염병이 확산되는 상황 속에서 대중에게 주요 정보를 제공하는 언론사 기사에 대한 연구는 부족한 실정으로서, 본 연구는 SARS, MERS, COVID-19와 관련 키워드를 통해 비교 연구를 수행함으로써 공통점 및 차이점을 알게 되었다는 점에서 그 가치가 있다.

둘째, 국내에 감염병이 유행하는 초기 시점에서는 다채널 확립화된 정보가 공유되는 시스템이 구축되어야 한다. 연구를 통해 분석한 결과 3개 감염병에 대해서 질병 정보, 확진자 정보, 보건당국의 브리핑이 주요 공통점으로 나타났다. 이에 따라, 상기 정보들에 대해서는 신뢰성 있는 질병정보, 투명한 확진자 정보, 명확한 보건당국의 브리핑 내용들을 우선순위 목표로 설정하여 정보 수신자의 편향과 왜곡이 없도록 신속하게 공유되는 시스템이 마련되어야 할 것이다.

셋째, 질병의 확진자 발생률 증가와 유행 기간이 중장기화되는 경우, 정부 역할론의 중요성을 인식하여 체계적인 통합 컨트롤 타워 기능이 구현되어야 하며 예방·진단·치료 영역에 대한 명확한 정보의 생산이 이루어져야 한다. 본 연구결과를 통해서 청와대, 질병관리본부 등 정부의 정책 및 지침안 등과 질병 간의 차이, 예방·진단·치료 등의 질병에 대한 대응 정보 관련 키워드가 증가한 양상으로써 주요 차이점이자 특징적인 내용을 보여주었다. 또한 선행 연구들을 통해 초기 감염병에 대한 인포데믹이 포함된 예방 정보의 범람[19-20,35-36], '거리 두기 단계별 기준 및 방역조치' 및 '환자 정보 공개 범위', '백신 접종 체계' 등을 포함한 정책적 이슈들에서 법적·제도적 한계들[37-38]이 지속해서 제기되고 있다. 이는 앞서 경험한 사스와 메르스에서 감염병의 방역에 집중하였던 정부와 보건 의료 기관의 역할과는 달리하는 제도적 보완 영역으로 도출할 수 있으며, 충분한 사회적 합의의 필요성이 증대되었다. 특히 메르스를 통해 얻은 '환자 정보의 투명한 공개'는 COVID-19초기에 구체적인 정보를 제공하면서 정상적으로 작동하였으나 심각한 개인 정보의 위협과 사회적 불안감이 제기되면서 제도적 재정비가 필요한 영역이 되었다 [39-40]. 이에 따라 사회적인 충분한 공감대 형성과 예방·진단·치료 영역을 포함한 제도적 보완을 통하여 감염병 대유행 시 정부와 국민이 동행하는 사회적 대응 체계 협력 거버넌스를 마련할 필요가 있다.

넷째, 감염병과 관련하여 불확실 정보에 대한 정확하고 엄중한 대응이 중요하다.

COVID-19의 진원지, 감염매개체, 활용 가능한 백신·치료제 등의 불명확함과 시기적으로 장기화되는 특징점에 따라, 불확실한 정보가 범람하고 있는 실정이다. 세계보건기구(WHO)는 가짜뉴스의 범람을 경고하며, 인포메이션(information)과 에피데믹(epidemic)을 합친 인포데믹(infodemic)이라는 용어를 통해 공중보건을 위협하는 정보 전염병에 대해 경고하였으며[28], 본 연구결과를 통해 확인한 결과에서도 불확실한 정보에 대한 정정 요구, 법적 대응, 허위사실에 대한 팩트체크 등이 주요하게 확인되었다. 특히 유튜브, SNS 등 사회관계망 서비스가 활발한 오늘날 시점에서는 검증되지 않은 불확실한 정보들이 급속도로 확산되는 경향을 나타내었다. 같은 맥락으로 미국에서는 페이스북, 구글, 마이크로소프트(MS), 트위터 등이 공동성명을 내고 의료당국과 공조하여 COVID-19에 대한 허위정보를 색출함으로써 공인된 정보의 콘텐츠를 확대하는 대응 시스템을 마련한 바 있다[41]. 이처럼 국내에서도, 의료계·정부·미디어 매체가 불확실한 정보에 대한 신속한 검증 및 대응이 이루어질 수 있는 가치 인포데믹방역시스템(prevention system of epidemics)이 마련될 필요가 있으며, 정보 취약계층을 비롯하여 대중이 쉽게 접근하고 자체 검증이 가능한 경로에 대한 충분한 홍보 및 교육이 이루어져야 할 것이다.

본 연구의 제한점으로는 COVID-19는 2021년을 비롯하여 현재 시점에서도 진행되고 있는 질병인 점과 전문학술 정보, 보건당국의 정책 등 전문가 시각이 아닌, 일반 뉴스 기사 분석으로 다소 피상적인 분석이 진행된 점을 들 수 있으며, 불확실한 정보에 대한 정의가 정립되지 않은 상태임에 따라, 제한적인 분석이 수행되었다. 따라서, 후속 연구로는 해당 3개 감염병에 대해 국내 대유행 전주기 동안의 대응 체계 및 의료물자 공급에 대한 체계적 문헌 검토 수행과 이슈 되었던 불확실 정보에 대한 팩트체크 연구를 통해서 미흡했었던 법적·제도적 부분에 대한 발본색원 연구가 수행된다면, 차후 유사한 감염병 대유행 시 국민-언론 미디어-정부-의료계의 위기 대응 역할론 정립 및 커뮤니케이션 고도화 방안이 마련될 것으로 기대한다.

REFERENCES

- [1] Dong Kyun Park, "The Problems and Lessons of Emergency Management in Korea: With the Case of Covid-19," Korean Association for Public Security Administration, Vol. 17, No. 3, pp. 127-150, Aug. 2020. DOI: 10.25023/kapsa.17.3.2020.8.0127
- [2] Ji Na Park, "Covid-19 and the International Health Regulations : Focusing on the Limitations and Problems," Seoul International Law Journal, Vol. 27, No. 1, pp. 1-28, Jun. 2020. DOI: 10.18703/silj.2020.06.27.1.1
- [3] Drosten C, Günther S, Preiser W, van der Werf S, Brodt HR, Becker S, et al. "Identification of a novel coronavirus in patients with severe acute respiratory syndrome," N Engl J Med, Vol. 348, No. 20, pp. 1967-1976, Apr. 2003. DOI: 10.1056/NEJMoa030747
- [4] Chan JF, Lau SK, To KK, Cheng VC, Woo PC and Yuen KY. "Middle East respiratory syndrome coronavirus: another zoonotic betacoronavirus causing SARS-like disease," Clin Microbiol Rev, Vol. 28, No. 2, pp. 465-522, Mar. 2015. DOI: 10.1128/CMR.00102-14
- [5] Zhou P, Yang XL, Wang XG, Hu B, Zhang L and Zhang W, "A pneumonia outbreak associated with a new coronavirus of probable bat origin," Nature, Vol. 579, No. 7798, pp. 270-273, Feb. 2020. DOI: 10.1038/s41586-020-2012-7
- [6] Eun Joong Kim and Dong sup Lee, "Coronaviruses: Sars, Mers and Covid-19," The Korean Journal of Clinical Laboratory Science, Vol. 52, No. 4, pp. 297-309, Dec. 2020. DOI: 10.15324/kjcls.2020.52.4.297
- [7] Mi Na Lee and Ju hyun Hong, "Semantic Network Analysis of Government's Crisis Communication Messages During the Mers Outbreak," The Journal of the Korea Contents Association, Vol. 16, No. 5, pp. 124-136, May 2016. DOI: 10.5392/jkca.2016.16.05.124
- [8] Jin Won Jeong and Jun Hui U, "Understanding Severe Acute Respiratory Syndrome (Sars)," Association of Internal Medicine, Vol. 65, No. 2, pp. 154-159, Aug. 2003. UCI: G704-000582.2003.65.2.026
- [9] Je Bum Pyun, and Seung Beom Kim, "Assessment of Covid-19 Response of the Medical Institutions Based on Iso Public Service Quality Management Framework," Journal of the Korea Industrial Information Systems Research, Vol. 25, No. 6, pp. 69-84, Dec. 2020. DOI: 10.9723/jksiiis.2020.25.6.069
- [10] Jin Sung Rha, "Analysis on Issues Related to Supply Chain Management in the Era of Covid19 Using Network Text Analysis," Journal of the Korea Industrial Information Systems Research, Vol. 25, No. 6, pp. 109-123, Dec. 2020. DOI: 10.9723/jksiiis.2020.25.6.109
- [11] Sun Ju Ahn, Hae Pum Park, Seung Youg Song, Ji Young Ryu, and Su Hwa Kim, "Non-Pharmaceutical Standard Models for Managing Pandemic," Society for Standards Certification and Safety, Vol. 11, No. 1, pp. 45-67, Mar. 2021. DOI: 10.34139/jscs.2021.11.1.45
- [12] Yeon Jeong Heo, So Hee Nam, Jae Sim Jeong, and Yeon Hee Kim, "A Comparison of the Perception of and Adherence to the

- Covid-19 Social Distancing Behavior Guidelines among Health Care Workers, Patients, and General Public," *Journal of Korean Biological Nursing Science*, Vol. 23, No. 1, pp. 55-63, Feb. 2021. DOI: 10.7586/jkbns.2021.23.1.55
- [13] Eun Jung Kim, Hye Min Sim, Jae Woong Won, and Bum Joon Kang, "Mapping the Covid-19 Issues from an Urban Perspective in South Korea - Text Mining Analysis Focused on Newspaper Articles," *Journal of the Urban Design Institute of Korea Urban Design*, Vol. 21, No. 6, pp. 163-179, Dec. 2020. DOI: 10.38195/judik.2020.12.21.6.163
- [14] KISTEP, Direction of Government R&D Support in Response to Infectious Diseases Such as Covid-19, https://www.kistep.re.kr/board.es?mid=a10306010000&bid=0031&list_no=42038&act=view.
- [15] Suk Man Kim, Sang Yong Park, Min Woo Lee, and Chul Woong Kang, "Comparison of Domestic and International Government Policies in Pandemic Circumstances and Crises: Based on Covid-19, Sars, Mers," *Journal of The Korean Society of Physical Medicine*, Vol. 16, No. 1, pp. 123-141, Feb. 2021. DOI: 10.13066/kspm.2021.16.1.123
- [16] Sung Man Hong, "A Study on the Conflict Amplification in Covid-19 Infectious Disease Disaster," *The Journal of Korean Policy Studies*, Vol. 20, No. 4, pp. 1-32, Dec. 2020. DOI: 10.46330/jkps.2020.12.20.4.1
- [17] U.S. Food and Drug Administration, South Korea's Response to COVID-19, <https://www.fda.gov/media/149334/download>.
- [18] Deuk-Sang Ma, "COVID-19, infodemic, and evidence-based dentistry," *Journal of Korean Academy of Oral Health*, Vol. 45, No. 2, pp. 49-50, Jun. 2021. DOI: 10.11149/jkaoh.2021.45.2.49
- [19] Seung Hee Ryu, and Chung Joo Chung, "Fact-Checking Status and Characteristics: Focusing on Covid-19 in Korea," *Discourse and Policy in Social Science*, Vol. 13, No. 2, pp. 271-298, Oct. 2020. DOI: 10.22417/dpss.2020.10.13.2.271
- [20] Joo-Hyun Park, "A Comparative Study on the 'Corona19' News Frame Based on Ideological Orientation of Media," *Korean Journal of Journalism and Communication Studies*, Vol. 64, No. 4, pp. 40-58, Aug. 2020. DOI: 10.20879/kjcs.2020.64.4.002
- [21] Na-Hyeon Kim and Hayoung Oh, "Analysis of COVID-19 Pandemic based on Massive Big Data Analysis," Vol. 25, No. 4, pp. 495-500, Apr. 2021. DOI: 10.6109/jkiice.2021.25.4.495
- [22] Todd, G. and Tuchman, G., "Making news: A study in the construction of reality," Vol. 9, No. 1, pp. 99, Jan. 1980. DOI: 10.2307/2065604
- [23] Pan, Z., and Kosicki, G. M., "Framing analysis: An approach to news discourse," *Political communication*, Vol. 10, No. 1, pp. 55-75, May 1993. DOI: 10.1080/10584609.1993.9962963
- [24] Babbie, R., "The practice of social research(11th ed.)," Cengage Learning, 2007.
- [25] Eun Byul Lee, Jin O John and Ji Sun Park, "A Study of Multicultural Space in Seoul : Analysing the Coverage of Foreign Communities with News Big Data Analytics 'BigKinds' for 27 Years," *Journal of Media Economics & Culture*, Vol. 15, No. 2, pp. 7-43, May 2017. DOI: 10.21328/JMEC.2017.5.15.2.7
- [26] Seong Wook Kwon, "Keyword Analysis of Covid-19 in News Big Data :Focused on 4 Major Daily Newspapers," *Journal of The Korea Society of Computer an Information*, Vol. 25, No. 12, pp. 101-107, Dec. 2020. DOI: 10.9708/jksci.2020.25.12.101
- [27] Korea Disease Control and Prevention Agency, Infectious disease outbreak information, <https://www.kdca.go.kr/>.
- [28] World Health Organization, Newsroom, <https://www.who.int/emergencies/diseases/novel-coronavirus-2019>.
- [29] Introduction of Korean Law Information Center, Framework Act on the Management of Disasters and Safety, <https://www.law.go.kr>.
- [30] Ministry of Food and Drug Safety, Infectious Disease Disaster' Crisis Response Practical Manual, <https://www.mfds.go.kr/docviewer/skin/doc.html?fn=20210701095846967.pdf&rs=/docviewer/result/data0010/14855/1/202108>.
- [31] Korea Transport Institute, Korea Report of COVID-19 Responses, https://english.koti.re.kr/user/bbs/BD_selectBbs.do?q_bbsCode=1054&q_bbscttSn=20200706184425103.
- [32] Korea Press Foundation, Manuel for BIGKinds's users(version 2.9), <https://www.bigkinds.or.kr/>.
- [33] Ysjang, and Hgkang, "The first step in big data analytics: coding with R," Life and Power Press. 2018.
- [34] T. Gottron, "Document word clouds: Visualizing web documents as tag cloud to aid users in relevance decisions," *Research and Advanced Technology for Digital Libraries-Lecture Notes in Computer Science*, 5714, 94-105, 2009.
- [35] Moon Han Lee, "Criminal Punishment for Expressing False Information that Harms the 'Public Interest' such as Fake News Related to COVID-19, and Freedom of Press and Publication under the Constitution," *SNU Law Review*, Vol. 11, pp. 201-251, Apr. 2021.
- [36] Hyoung-Jee Kim, EunRyung Chong, Eun-mee Kim, Soeun Yang, Jae Woo Lee and Minji Kang, "Fake News and Fact Check News Differences : Focusing on News Usage, Perception, and Literacy in Multi-Media Environments," *Korean Journal of Communication & Information*, Vol. 101, pp. 231-267, Jun. 2020. DOI: 10.46407/kjci.2020.06.101.231
- [37] Lee, J., "A Legal study for Reasonable response to Coronavirus(COVID-19) Pandemic : Suggestions on the Scope of Information Use by Infectious Disease Patients," *Kangwon Law Review*, Vol. 61, No. 2, pp. 439-477, Oct. 2020. DOI: 10.18215/kwlr.2020.61.439
- [38] Sun-Jae Hwang, Jung-ah Gil and Seulki Choi, "Vaccination Acceptance for COVID-19: Implications of Trust in Government," *Korea Journal of Population Studies*, Vol. 44, No. 2, pp. 95-120, Jun. 2021. DOI: 10.31693/KJPS.2021.06.44.2.95

- [39] Kyung-lyul Lee, "Dilemma on the protection and urgent use personal information for public safety," *Korean Journal of Comparative Criminal Law*, Vol. 23, No. 1, pp. 137-165, Apr. 2021. DOI: 10.23894/kjcccl.2021.23.1.005
- [40] Geun Oak Lee, "The Legal Harmony between Personal Information Protection and Public Interest under the Official Information Disclosure Act : Focusing on Privacy Infringement under the COVID-19 Pandemic," *Korean Journal of Communication & Information*, Vol. 103, pp. 145-176, Oct. 2020. DOI: 10.46407/kjci.2020.10.103.145
- [41] Pérez-Escoda, A., Jiménez-Narros, C., Perlado-Lamo-de-Espinosa, M. and Pedrero-Esteban, L. M., "Social networks' engagement during the COVID-19 pandemic in Spain: health media vs. healthcare professionals," *International journal of environmental research and public health* Vol. 17, No. 14, pp. 5261, Jul. 2020. DOI: 10.3390/ijerph17145261

Authors



Jae-Hyun Woo received a B.S. degree in Biomedical Engineering from Konkuk University, Korea, in 2015. He is currently attending an Integrated Master's and Doctorate Course in the Department of Medical Device

Industry at Dongguk University. Jae-Hyun Woo is currently a Senior researcher in the Department of Research Institute for Commercialization Biomedical Convergence Technology at Dongguk University. He completed training as a trainer at Singapore-Stanford BIODESIGN in 2016. He is interested in regulatory science and technical policy studies of medical devices.