The multidisciplinary nature of COVID-19 research
Ricardo Arencibia-Jorge* 1, Lourdes García-García2, Ernesto Galbán-Rodríguez3, Humberto Carrillo-Calvet4

ABSTRACT

Objective. We analyzed the scientific output after COVID-19 and contrasted it with studies published in the aftermath of seven epidemics/pandemics: Severe Acute Respiratory Syndrome (SARS), Influenza A virus H5N1 and Influenza A virus H1N1 human infections, Middle East Respiratory Syndrome (MERS), Ebola virus disease, Zika virus disease, and Dengue.

Design/Methodology/Approach. We examined bibliometric measures for COVID-19 and the rest of the studied epidemics/pandemics. Data were extracted from Web of Science, using its journal classification scheme as a proxy to quantify the multidisciplinary coverage of scientific output. We proposed a novel Thematic Dispersion Index (TDI) for the analysis of pandemic early stages.

Results/Discussion. The literature on the seven epidemics/pandemics before COVID-19 has shown explosive growth of the scientific production and continuous impact during the first three years following each emergence or re-emergence of the specific infectious disease. A subsequent decline was observed with the progressive control of each health emergency. We observed an unprecedented growth in COVID-19 scientific production. TDI measured for COVID-19 (29.4) in just six months, was higher than TDI of the rest (7.5 to 21) during the first three years after epidemic initiation.

Conclusions. COVID-19 literature showed the broadest subject coverage, which is clearly a consequence of its social,
Epidemics and pandemics have significant repercussions in human societies, spread beyond national borders, and affect many people in extensive areas (Porta Serra, 2014). Throughout history, humankind has suffered millions of deaths due to these diseases worldwide, undergoing vast economic and social losses (Fan, Jamison, and Summers, 2018; Huber, Finelli, and Stevens, 2018; Keogh-Brown and Smith, 2008; Kuhar and Fatović-Ferenčić, 2020). The early implementation of suppression or mitigation strategies has been crucial to face adverse effects (Kuhar and Fatović-Ferenčić, 2020; Pike et al., 2014). In such contexts, the integration of various sectors of society in decision-making processes has been paramount to balance health and economic priorities and minimize the impact in almost every aspect of social life.

COVID-19 pandemic is a disruptive experience for everyone, including scientists (Myers et al., 2020). The massive increase in scientific articles might have overcome researchers. Dozens of bibliometric studies have been published after the emergence and growing expansion of the new pandemic. Most of them include traditional bibliometric indicators, such as main contributing authors, journals, institutions, and countries (Chahrour et al., 2020; Darsono, Rohmana, and Busro, 2020; De Felice and Polimeni, 2020; Dehghanbanadaki et al., 2020; Kambhampati and Vaish, 2020). Thematic clusters have been identified through mapping techniques according to keywords co-occurrence, co-citation patterns, or international collaborations (El Mohadab, Bouikhalene, and Safi, 2020; Hamidah, Sriyono and Hudha, 2020; Herrera-Viedma et al., 2020). A few of them have compared bibliometric patterns of epidemic/pandemic diseases (Tao et al., 2020; Zhai et al., 2020; Zhang et al., 2020; Zhou and Chen, 2020). Descriptive statistics, co-occurrence networks analysis, and text mining have been among the explored methods to define the boundaries of COVID-19 research. However, previous reports have not been focused on research multidisciplinarity; and this can be solved using novel bibliometric indicators based on the thematic concentration of scientific output (Arencibia-Jorge, Vega-Almeida, and Carrillo-Calvet, 2020).

The current bibliometric study aims to analyze and compare scientific literature multidisciplinarity on COVID-19 and other 21st century epidemics/pandemics. We consider that we successfully overcame the limitations of previous reports.

METHODOLOGY

Data Recovery
Scientific output data of each epidemic/pandemic was retrieved from the Web of Science™
Search dates ranged as follows for the different diseases: 2003 to 2005 for Severe Acute Respiratory Syndrome (SARS); 2005 to 2007 for Influenza A virus H5N1 human infection; 2009 to 2011 for Influenza A virus H1N1 human infection; 2012 to 2014 for Middle East Respiratory Syndrome (MERS); 2014 to 2016 for Ebola and Zika virus diseases; 2015 to 2017 for Dengue; and January to June 2020 for COVID-19. A three-year period was observed for previous epidemics/pandemics, taking into account the duration of the stage of spread and subsequent control of each disease.

**Basic Indicators**

The following indicators were calculated for the set of articles of each epidemic/pandemic:

- **A**: articles published during the first three years (only the first six months were considered for COVID-19); **A Cit**: citing articles during the first three years after the outbreak; **Cit (mean)**: mean of citations received in one year by papers published during the same year (calculated for the first, second and third year after disease emergence); **WCs**: Web of Science subject categories; **Cit WCs**: Citing Web of Science subject categories.

**New indices proposed**

Following the Pareto principle, we used the WoS classification scheme to calculate the thematic concentration of articles generated by each disease and their citing articles (Arencibia-Jorge, Vega-Almeida, and Carrillo-Calvet, 2020):

- **Thematic concentration of scientific production (TCp)**: measure that expresses the core or minimum number of WCs that collect approximately 80% of articles.
- **Thematic concentration of citations (TCc)**: measure that expresses the core or minimum number of WCs that collect approximately 80% of citing articles.

Then, a novel Thematic Dispersion Index (TDI) was developed, with the aim to balance both WCs cores, using the following equation:

\[ \text{TDI} = \sqrt{\text{TCp} \cdot \text{TCc}} \]
TDI has the following property: \( TDI \geq 1 < 1x \), where: \( x \) equals the maximum number of WCs (currently 254). Values close to 1 (80\% of literature covered by only one WCs) express a high degree of disciplinary specialization or scientific production concentration, while higher values (80 \% of literature covered by more than one WCs) will increasingly determine multidisciplinary nature.

RESULTS

Despite the international scientific community reaction given the dramatic explosion of literature (more than 8,000 papers covered by WoS during the first six months, Figure 1a and 1c), the dynamic patterns of COVID-19 research could be qualitatively similar to those observed after other 21st century epidemics/pandemics.

We have confirmed that emerging infectious diseases with an epidemic/pandemic nature typically unleash an accelerated growth of scientific production during the three years after outbreak emergence (Figure 1b). This initial stage generates an avalanche of data and evidence derived from the damaging effects of the disease and the growth of incentives to face the problem. Subsequently, once the health emergency is under control or effective treatments have appeared, a gradual decrease in scientific production begins.

Likewise, the average annual citations received by articles published in those first three years
will increase, reaching its maximum peak generally the third year after outbreak emergence (Table 1). This phenomenon has occurred regularly during each of the seven epidemics/pandemics that preceded COVID-19. However, one remarkable aspect of the new pandemic is the multidisciplinary nature of research.

Five of the seven severe outbreaks before COVID-19 involved journal articles covering more than one hundred Web of Science subject categories (WCs). In the five cases, the thematic core of scientific production encompassed 10 to 25 WCs, mainly connected to biomedical domains (Table 1; Figure 2).

Therefore, research related to epidemic/pandemic diseases tends to go beyond disciplinary frameworks. We observed that after the eight epidemics/pandemics, the multidisciplinarity was related to the disease mortality, biomedical complexity, global dispersion, and economic impact. Hence, Influenza A-H5N1 and MERS, which affected fewer countries, had a lower multidisciplinary scope and concentrated the least number of articles published during the three years after outbreak emergence.

COVID-19, in a short time, has become the most extensive and deadliest pandemic of this century (more than 500,000 deaths during the first semester, and two million expected for the next year). COVID-19 research has increasingly involved different thematic domains (more than two hundred WCs) directly or indirectly related to the disease, prevention, diagnosis, treatment, and social response. The TDI reached by the current pandemic is the highest of all analyzed. It has nearly 30 WCs concentrating on the most massive volumes of articles and citations, reflecting the multidimensional impact that the disease has had in just six months.

<table>
<thead>
<tr>
<th>Epidemics/Pandemics (Start Year)</th>
<th>A (mean)</th>
<th>A Cit (mean)</th>
<th>Cit Year 1 (mean)</th>
<th>Cit Year 2 (mean)</th>
<th>Cit Year 3 (mean)</th>
<th>WCs</th>
<th>Cit WCs</th>
<th>TCp</th>
<th>TCc</th>
<th>TDI</th>
</tr>
</thead>
<tbody>
<tr>
<td>SARS (2003)</td>
<td>1871</td>
<td>3272</td>
<td>4.08</td>
<td>9.25</td>
<td>15.21</td>
<td>144</td>
<td>163</td>
<td>18</td>
<td>18</td>
<td>18.0</td>
</tr>
<tr>
<td>H5N1 (2005)</td>
<td>453</td>
<td>1858</td>
<td>3.41</td>
<td>9.01</td>
<td>14.00</td>
<td>71</td>
<td>139</td>
<td>9</td>
<td>11</td>
<td>9.9</td>
</tr>
<tr>
<td>H1N1 (2009)</td>
<td>3694</td>
<td>7628</td>
<td>2.69</td>
<td>6.27</td>
<td>10.9</td>
<td>134</td>
<td>173</td>
<td>11</td>
<td>13</td>
<td>12.0</td>
</tr>
<tr>
<td>MERS (2012)</td>
<td>712</td>
<td>1368</td>
<td>2.92</td>
<td>9.73</td>
<td>12.57</td>
<td>69</td>
<td>93</td>
<td>7</td>
<td>8</td>
<td>7.5</td>
</tr>
<tr>
<td>EBOLA (2014)</td>
<td>4344</td>
<td>6661</td>
<td>1.48</td>
<td>3.61</td>
<td>7.64</td>
<td>176</td>
<td>197</td>
<td>20</td>
<td>22</td>
<td>21.0</td>
</tr>
<tr>
<td>ZIKA (2014)</td>
<td>1934</td>
<td>1930</td>
<td>1.04</td>
<td>4.41</td>
<td>6.50</td>
<td>130</td>
<td>130</td>
<td>13</td>
<td>13</td>
<td>13.0</td>
</tr>
<tr>
<td>DENGUE (2015)*</td>
<td>7636</td>
<td>13748</td>
<td>0.69</td>
<td>4.62</td>
<td>8.07</td>
<td>169</td>
<td>205</td>
<td>15</td>
<td>22</td>
<td>18.2</td>
</tr>
<tr>
<td>COVID-19 (2020)</td>
<td>8608</td>
<td>8501</td>
<td>3.53</td>
<td>-</td>
<td>-</td>
<td>203</td>
<td>188</td>
<td>32</td>
<td>27</td>
<td>29.4</td>
</tr>
</tbody>
</table>

Table 1. Bibliometric measures of productivity, impact, and multidisciplinary scope after selected epidemics/pandemics (Source: WoS. Date of retrieval: June 30, 2020). (*Dengue has had pandemic characteristics during this century. More than 4,000 deaths were reported in 2015. The largest number of dengue cases ever reported globally was in 2019. All regions were affected).
Figure 2. Thematic concentration of articles and citations, and thematic dispersion index (TDI) of the eight analyzed pandemics during the first three years after disease emergence. Bubble’s size is proportionate to the scientific output. Blue bubbles: Severe Acute Respiratory Syndrome (SARS); Influenza A virus H5N1 and Influenza A virus H1N1 human infections, Middle East Respiratory Syndrome (MERS), Ebola virus disease, Zika virus disease, and Dengue. Red bubble: COVID-19.

DISCUSSION

Bibliometric multidisciplinary, interdisciplinary, or transdisciplinary analyses use two main approaches: A Bottom-up approach, based on clustering sets of articles according to a bibliographic criterion, either bibliographical coupling or co-citation networks; and a Top-to-bottom approach dependent on existing classifications schemes (Wagner et al., 2011). These strategies search for communication channels between authors, institutions, cited or citing authors, cited or citing journals, cited or citing documents, and mapping techniques to illustrate the diversity of research areas (Chen, 2017; Klavans and Boyack, 2011; Mochini et al., 2020). In this paper, we used the second approach to create a battery of indicators that expose the multidisciplinary nature of pandemics.

In just a few months, bibliometric indicators of COVID-19 have surpassed by far what has been observed in other epidemics/pandemics. There are five WoS subject categories (Medicine General & Internal, Infectious Diseases, Virology, Immunology, and Microbiology) shared by the thematic core of eight studied pandemics. COVID-19 research hotspots have
dealt with molecular virology, immunogenetics, epidemiology, medicine, imaging, pharmacology, environmental sciences, economics, anthropology, social sciences, philosophy, ethics, and other disciplines (El Mohadab, Bouikhalene, and Safi, 2020). Our bibliometric analysis demonstrated the growing thematic expansion with COVID-19.

Several reasons explain the multidisciplinary burst of COVID-19 research, as compared to other epidemics/pandemics. Its global distribution, incidence rates initially affecting high-income countries, high fatality rates in certain groups, wide-ranging clinical manifestations, lack of a vaccine or specific treatment during the first year, and psychological, social, political, and economic consequences, may altogether motivate the keen interest of a vast scientific community. Probably, the current technological resources plus extensive usage of social media and pre-print communications have also favored and accelerated communication between scientists from different disciplines and countries.

At the same time, COVID-19 is a respiratory infectious disease with numerous extrapulmonary manifestations that turn it into a multi-organ condition (Wang et al., 2020). Patients suffering from the most frequent chronic diseases have been significantly affected. Diabetes, hypertension, cardiovascular diseases, and cancer account for a highly vulnerable population with a worse prognosis contributing to death rates statistics. Hence, clinical consequences have been varied and severe, requiring the conjunction of different medical specialties. Multispecialty is a hallmark of COVID-19 research published in journals from various medical fields, seldom seen in previous epidemics/pandemics. The international race for the fast development and introduction of vaccines, its distribution and applying mechanisms, will probably increase the research on Immunology and clinical topics in 2021.

Additionally, this pandemic has strong links with the environment and OneHealth approaches. Leading approaches have included the origin of the virus, wildlife surveillance, risk reduction, and biosecurity. These topics have required the integration of researchers from Biological Sciences, such as Zoology, Genetics, Evolutionary Biology, Physiology, Biochemistry, and Cell Biology, in collaboration with ecologists, physicians, epidemiologists, and even in occupational health and safety specialists. Lockdown strategies have also required economists, sociologists, psychologists, psychiatrists, and researchers from Management, Commerce, Transport, and Tourism.

The techno-driven approach based on artificial intelligence and machine learning used to forecast outbreaks resurgence, SIR modeling, imaging processing, and diagnosis tools have also contributed. Internet of things (IoT) devices, big data, robotics, drone technologies, and multiple mobile applications have also been used to monitor mobility and track active cases and contacts.

Multidisciplinary research and development (R&D) approaches are crucial to effectively face a pandemic (Peters, Jandrić, and McLaren, 2020; Moradian et al., 2020). With the incorporation of specialists from various multinational teams, better conditions are created to develop interdisciplinary and transdisciplinary research on COVID-19 (Moradian et al., 2020). Thus, data analysts, computer specialists, robotics experts, engineers from various branches, sociologists, psychologists, philosophers, lawyers, information scientists, among others, have provided input to COVID-19 research.

The diversification of journal ecosystems and peer-review acceleration (Smart, 2020), together with the open exchange of experimental data following the Open Access and the more recent Open Science movements, could have also contributed to the explosion of COVID literature (Belli et al., 2020). The intensive and accelerated use of pre-print servers
has been another fingerprint of the new pandemic (Kozlakidis et al., 2020; Johanson et al., 2020). Overall, we testify an unprecedented phenomenon that just started, and we are unable to predict its extent and magnitude. Nevertheless, we consider that the productivity, impact, and multidisciplinary scope of COVID-19 literature has been extraordinary and will continue to grow for at least three more years.

LIMITATIONS

The main limitation of this study is its intrinsic bias. WoS-based results may differ from those obtained from other databases. However, previous studies have successfully used this data set (Zhang et al., 2020). To our knowledge, this is the first study that uses a journal classification scheme to analyze multidisciplinarity of COVID-19 from the bibliometric perspective. We consider that the WCs core offered an essential multidisciplinarity dimension as it covers the largest volume of articles and citations. The new metric we are proposing may be used to compare different levels of aggregation (institutions, individuals, countries), using complementary diversity measures to analyze the sets of publications (Moschini et al., 2020; Porter and Rafols, 2009). TDI would be relatively easy to implement in other internet-available databases with large journal classification schemes (e.g., Scopus or Dimensions).

CONCLUSIONS

In only six months, COVID-19 literature involved increasing thematic domains directly or indirectly related to the disease and its social, economic, and political consequences. The newly proposed indicator (TDI), based on a database source-aggregated classification system, allowed the study of multidisciplinarity, differentiating the thematic complexity of COVID-19 from the previous seven epidemics/pandemics.

Conflicts of interest

The authors declare that there is no conflict of interest.

Contribution statement

Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Supervision, Visualization, Writing – original draft, Writing – review & editing: Ricardo Arencibia Jorge.

Conceptualization, Formal Analysis, Investigation, Writing – original draft, Writing – review & editing: María de Lourdes García-García, Ernesto Galbán-Rodríguez.

Conceptualization, Methodology, Supervision, Validation, Project administration, Writing – review & editing: Humberto Carrillo-Calvet.

Statement of data consent

Data generated during the development of this study has deposited in Figshare and it can be accessed at https://doi.org/10.6084/m9.figshare.13440146

Funding

This research was supported by the project “Scientometrics, Complexity and Science of
Science”, at the Complexity Science Center of the National Autonomous University of Mexico (UNAM).

REFERENCES


Kuhar, M., & Fatović-Ferenčić, S. (2020). Victories and defeats: battles with pandemics caused by viruses during the last one hundred years. Liječnički vjesnik, 142(3-4), 107-113. https://doi.org/10.26800/LV-142-3-4-19


