

# A bibliometric review of COVID-19 from december 2019 to july 2020

LUIS EDUARDO PINO-VILLARREAL, IVÁN CAMILO TRIANA-AVELLANEDA, MATEO BARROS-BARRAZA, LUCÍA CAROLINA VIOLA-MUÑOZ • BOGOTÁ, D.C. (COLOMBIA)

DOI: <https://doi.org/10.36104/amc.2020.1879>

## Abstract

**Objective:** to perform a bibliometric analysis of COVID-19 from December 2019 to June 30, 2020.

**Methods:** a bibliometric review of all the literature and COVID-19 related material in the PubMed, Scopus and Lilacs databases from December 2019 to June 30, 2020. Articles were classified by categories (objective, country, specialty) and compared to the previous year's literature.

**Results:** a total of 27,373, 16,944 and 1,083 publications on COVID-19 were found in PubMed, Scopus and Lilacs, respectively. The main medical specialty by search was pulmonology. The most frequently found objective was treatment (50.0%). The country with most publications was the United States (28.9%) on PubMed and Scopus. During the COVID-19 span, publications on the disease represented 0.06 of the total.

**Conclusions:** there was a significant volume of COVID-19 related publications during the study period, equivalent to 6% of the total publications, which is significant for a single disease. (*Acta Med Colomb* 2020; 45. DOI: <https://doi.org/10.36104/amc.2020.1879>).

**Key Words:** *COVID-19, bibliometric review, descriptive statistics.*

Dr. Luis Eduardo Pino-Villarreal: Hematólogo y Oncólogo Clínico. Magíster en Oncología Molecular y en Administración de la Salud MBA. Departamento de Oncología Hospital Universitario Fundación Santa Fe de Bogotá, Facultad de Medicina, Universidad de los Andes; Dr. Iván Camilo Triana-Avellaneda: Asistente de Investigación en Oncología. Departamento de Oncología, Hospital Universitario Fundación Santa Fe de Bogotá; Mateo Barros-Barraza: Estudiante de Medicina, Facultad de Medicina Universidad de los Andes; Dra. Lucía Carolina Viola-Muñoz: Especialista en Medicina Interna y Neumología, Departamento de Oncología, Fundación Neumológica Colombiana. Bogotá, D.C. (Colombia).

Correspondencia: Dr. Iván Camilo Triana-Avellaneda. Bogotá, D.C. (Colombia).

E-mail: [camilotrianaa@gmail.com](mailto:camilotrianaa@gmail.com)

Received: 15/V/2020 Accepted: 17/VII/2020

## Introduction

In December 2019, the first known case of a novel coronavirus infection occurred, originating in the city of Wuhan, China (1). On December 31 of the same year, the regional office of the World Health Organization (WHO), from its office in Beijing, reported a cluster of similar cases of respiratory infection, specifically in the same city of Wuhan. All the recorded cases had a common history of contact with the city's seafood market.

On January 8, 2020, the Centers for Disease Prevention and Control (CDC), in its first bulletin, described the investigation of acute respiratory infection cases in China. At the time of that report, cases had not been identified outside of the Asian country.

In the WHO report on January 20, 2020, four cases were reported outside of China in Thailand, Japan and Korea, all of which were exported from Wuhan. Out of the 278 Chinese reports, six deaths from the disease were notified (2).

A novel coronavirus was identified in bronchoalveolar lavage (BAL) samples of patients with the acute respiratory infection; using real-time polymerase chain reaction (PCR), viral nucleic acid was identified. Likewise, genetic sequencing of the ribonucleic acid (RNA) extracted from the BAL was performed (2). The novel coronavirus, the

causative germ of the infection identified in three patients from Wuhan with a compatible clinical picture, corresponds to a betacoronavirus, subgenus sarbecovirus of the *coronaviridae* family. Using similar techniques, a short time later these findings were confirmed with nine more cases, which, in turn, allowed the development of a PCR test to identify the infection (3).

Three months after the first descriptions, the disease is now a pandemic which has spread throughout all continents except Antarctica.

In Colombia, to date, there are more than 150,000 known cases and almost 6,000 deaths (as of July 14, 2020). Little is known regarding the clinical behavior outside of the cases initially reported in China; the spread of the virus in other countries has not had the same characteristics, especially in Italy (4), and there is still no specific treatment for the infection today.

Healthcare personnel caring for infected patients have a high risk of contracting the disease, similar to the situation in previous epidemics of similar coronaviruses such as the severe acute respiratory syndrome (SARS) in 2002-2003 in Hong Kong, Toronto, Singapore and Guangzhou (5).

There are no clear pharmacological strategies for preventing the infection. Non-pharmacological interventions

which have shown a reduction in infection transmission have been hand washing and hygiene, social isolation, mandatory quarantine for travelers to high risk zones or following unprotected contact with infected individuals, and closure of sites where large groups of people may gather (schools, theaters, sporting events) (6).

Having been declared a pandemic by WHO in March 2020, COVID-19 has led to focusing the scientific abilities of the different countries almost exclusively on the study of this disease. Likewise, social distancing measures designed to mitigate the impact of the pandemic have been implemented worldwide. Despite only having been on the scientific scene for four months, the SARS-CoV-2 virus is producing data at an enormous rate with a variety of publications of all types including rapid-progression studies in search of effective therapeutic interventions and vaccines. It is important to understand the bibliometric impact of the pandemic to be aware of the volume of information published, the topics covered, the countries which are producing this literature and the potential impact generated by the diversion of priorities towards this disease, as an opportunity cost for research in other medical topics.

## Methods

A bibliometric review was performed from December 1, 2019 to June 20, 2020 (article cutoff date). To accomplish this, the following databases were consulted: PubMed, Scopus and Lilacs (Spanish acronym for “Latin American and Caribbean Literature in Health Sciences”), to search for articles according to the specified search criteria. All published articles with the key words “COVID-19” and/or “SARS-CoV-2” were filtered; the search term was “COVID-19 or 2019-novc”. The articles found were classified by specialty, by the focal objective of the article, by source country of the publication and by journal of publication. Concurrently, a review was performed of the various medical articles published in the same databases from December 2018 to June 30, 2019, to have an initial comparison of the universe of publications outside of the COVID-19 scenario. The MAIA data platform (in-house platform) and the VOSviewer program were used for data analysis.

## Results

For the period from December 1, 2019 to June 30, 2020, a total of 45,400 articles on COVID-19 and/or SARS-CoV-2 were found in the three databases. Table 1 shows the distribution of these articles by database as follows: PubMed 27,373 articles, Scopus 16,944 and Lilacs 1,083 articles.

For the period between December 1, 2018 and June 30, 2019, 743,175 articles were found in all medical literature for two databases (PubMed and Lilacs), since a filter by specific months was unable to be applied on Scopus. Comparing the pre-COVID-19 period (2018-2019) with the COVID-19 period (2019-2020), a difference of 57,747 articles was found

in the total medical literature publications in the PubMed and Lilacs databases, showing a 7% increase in publications during the COVID-19 period. An evaluation of the COVID-19 articles in the two periods shows that the articles related to this topic in the pre-COVID-19 period were 0% (the disease was not yet described), but in the COVID-19 period the publications on the disease represented 0.06 of all the published articles; as a percentage, they represent 6% of the total publications for the period from December 2019 to June 30, 2020.

Table 2 shows how the articles of the PubMed and Scopus databases have been concentrated by medical specialty; this classification was not performed for Lilacs due to the nature of the articles. Table 3 shows the main objectives or intentions of the articles on COVID-19.

Articles were also distributed by country, and these results are found in Table 4. Using the VOSviewer program, the articles' key words having at least two occurrences were correlated; the articles' keyword correlations were extracted in Figures 1 and 2, and the correlation of terms reported in the abstract is reported. Thus, the terms or concepts on which COVID-19 research is focused are visually reported. Furthermore, the journals in which the articles have been published were correlated, with a minimum of two events, with the findings seen in Figure 3.

The distribution by country was performed on the three data bases, with 19,192 articles on PubMed which were not able to be classified. The countries found to have the highest publication volume were United States with 28.9% and England with 17.2%, followed by China with 12.7% and Italy with 10.1%. In the Lilacs Latin American database, the countries with the greatest participation were Brazil 43.5%, Colombia 28.7% and Peru 10.3%.

## Discussion

The present bibliometric review shows an incremental growth in publications related to COVID-19 during the first semester of 2020, accounting for 6% of all medical publications for three of the main global and Latin American journal databases, but leading to a comparative growth of 7%, which is significant for a single disease. In a previously unpublished

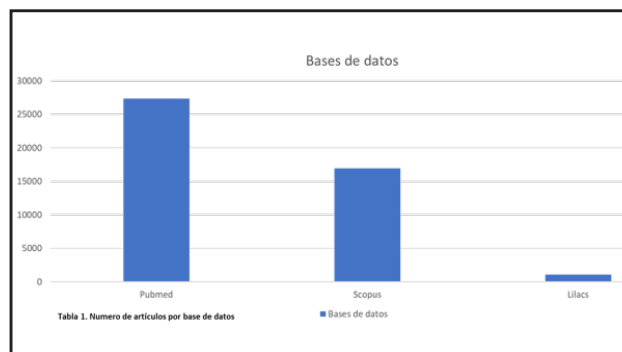


Table 1. Number of articles by database.

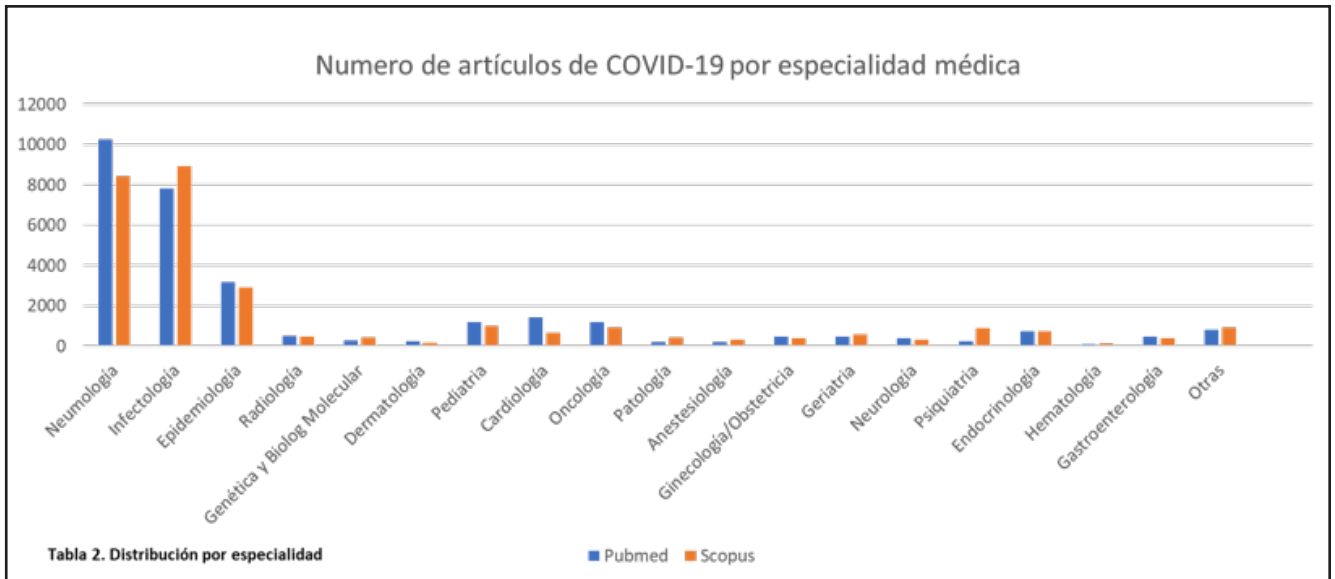


Table 2. Distribution by specialty.

report by our group, with a cut-off date of March 31, 2020, 0.005 of the publications corresponded to this disease, that is, 0.5% compared to 6% as of June 30, 2020. That is, three months later, we saw a 1,200% growth which corresponds to 12 times the first trimester cut-off for the same year.

With regard to the specific topics, the three on which the greatest number of articles have been published have been pulmonology, infectious disease and epidemiology. These are correlated with the main search terms which have been humans, infection control, and retroviral agents. Likewise, there is a heterogeneous distribution of journals or publication sites of the articles.

With regard to the contribution of publications by country, the lead clearly belongs to the United States, where research on the virus is concentrated, along with China where the first case reports were made, being the country of origin of the outbreak. In Latin America, on the other hand, Brazil and Colombia lead the publications, but these have mainly been disease protocols and treatment guidelines, as well as epidemiologic modeling.

The SARS-CoV-2 virus and its disease, COVID-19, have had a significant impact on medical publications, which may possibly be greater in the second trimester of 2020. This shows the response capacity of the scientific community in the face of this public health threat, but at the same time allows us to conclude that this first phase was marked mainly by articles related to the pathophysiology of the virus and statistical modeling, with a significant weight in all the medical literature. In subsequent phases, data more related to treatment, global health and epidemiology may be published. As this is a constantly evolving topic, it is important to continue with bibliometric studies to understand the

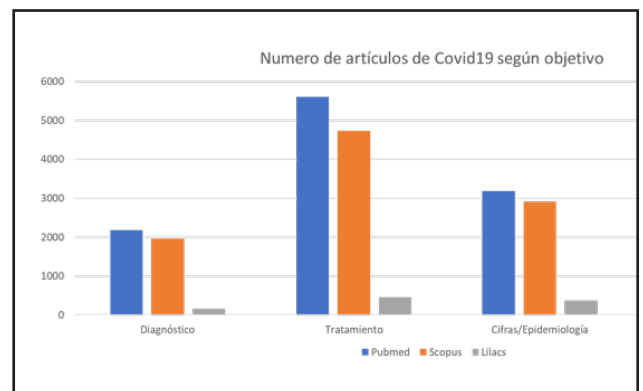


Table 3. Distribution by objective.

publication tendencies, since the high impact journals are being deluged with publication requests and some of them have opted to accept only 16-20% of the articles received (data obtained by the group from an international journal).

### References

1. Zhu N, Zhang D, Wang W, Li X, Yang B, Song J, et al. A novel coronavirus from patients with pneumonia in China, 2019. *N Engl J Med.* 2020;382(8):727–33.
2. World Health Organization (WHO). Novel Coronavirus (2019-nCoV) Situation Report - 1 21 January 2020. *WHO Bull.* 2020; (January): 1–7.
3. Lu R, Zhao X, Li J, Niu P, Yang B, Wu H, et al. Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. *Lancet [Internet].* 2020; 395(10224): 565–74.
4. Rosenbaum L. Facing Covid-19 in Italy — Ethics, Logistics, and Therapeutics on the Epidemic’s Front Line. *N Engl J Med.* 2020;1–3.
5. Chan-yeung M, Xu R. SARS: Epidemiology. *Respirology.* 2003; 8: S9–14.
6. Hellewell J, Abbott S, Gimma A, Bosse NI, Jarvis CI, Russell TW, et al. Feasibility of controlling COVID-19 outbreaks by isolation of cases and contacts. *Lancet Glob Heal.* 2020; 8(4): e488–96.



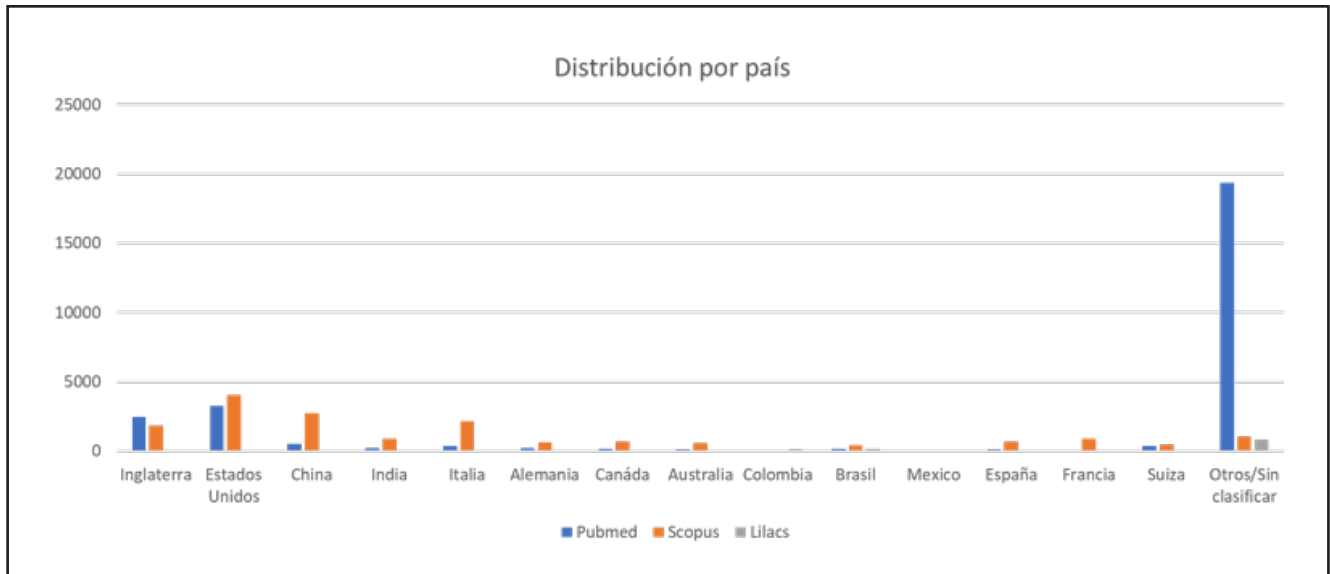


Table 4. Distribution by country.

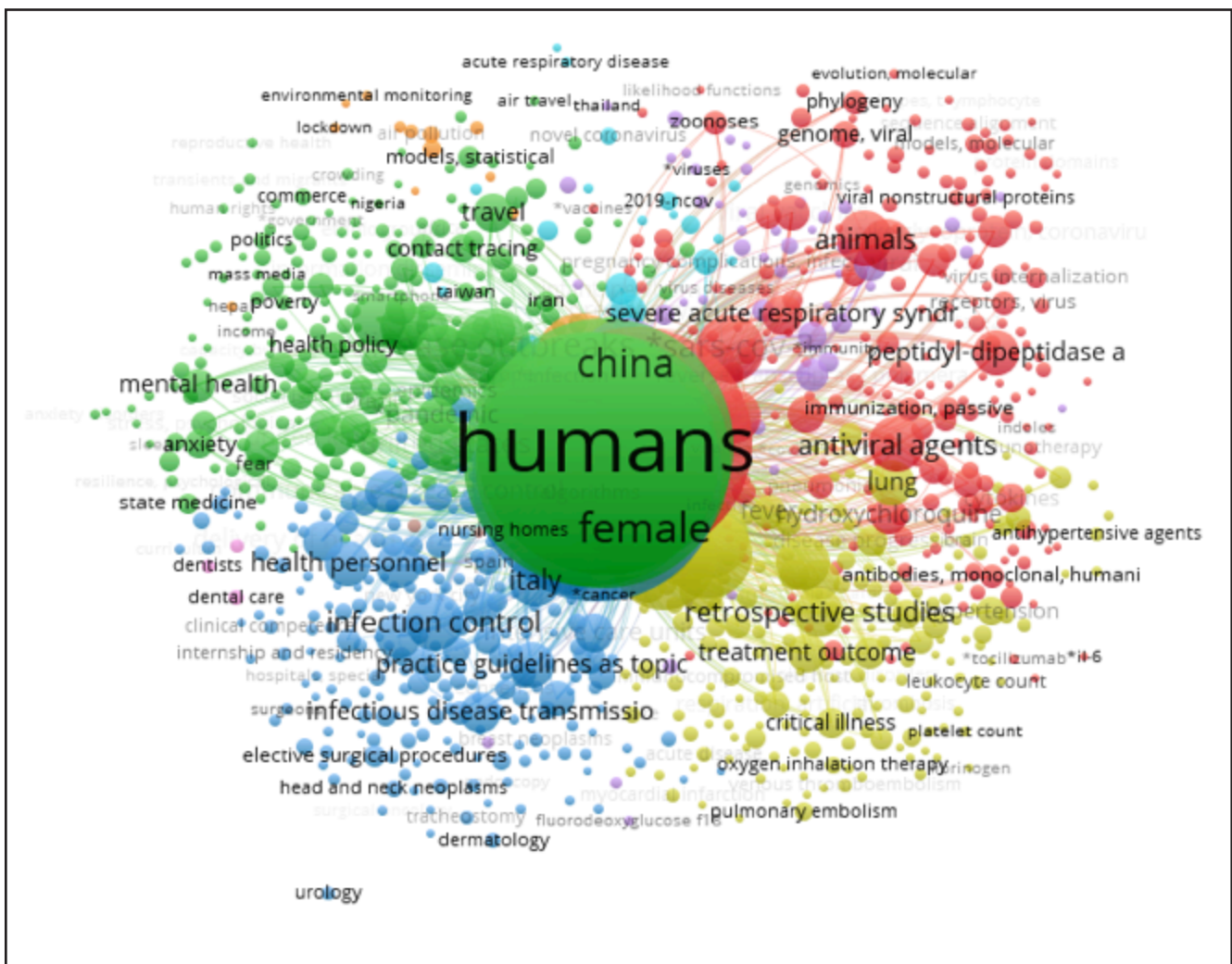


Figure 1. Keyword correlation. The lines which link circles represent the 100 most repeated correlations..

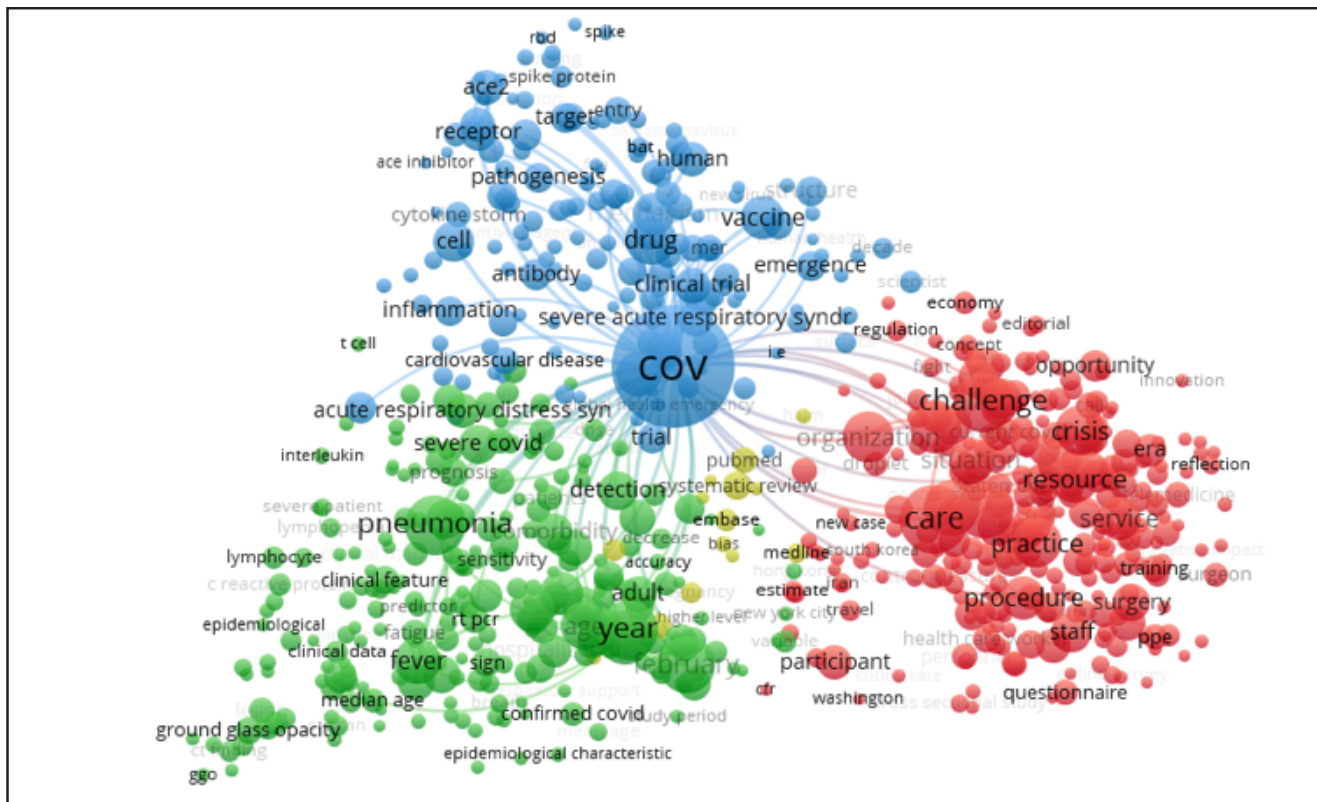


Figure 2. Correlation of terms in the abstract.. The lines which link circles represent the 100 most repeated correlations.

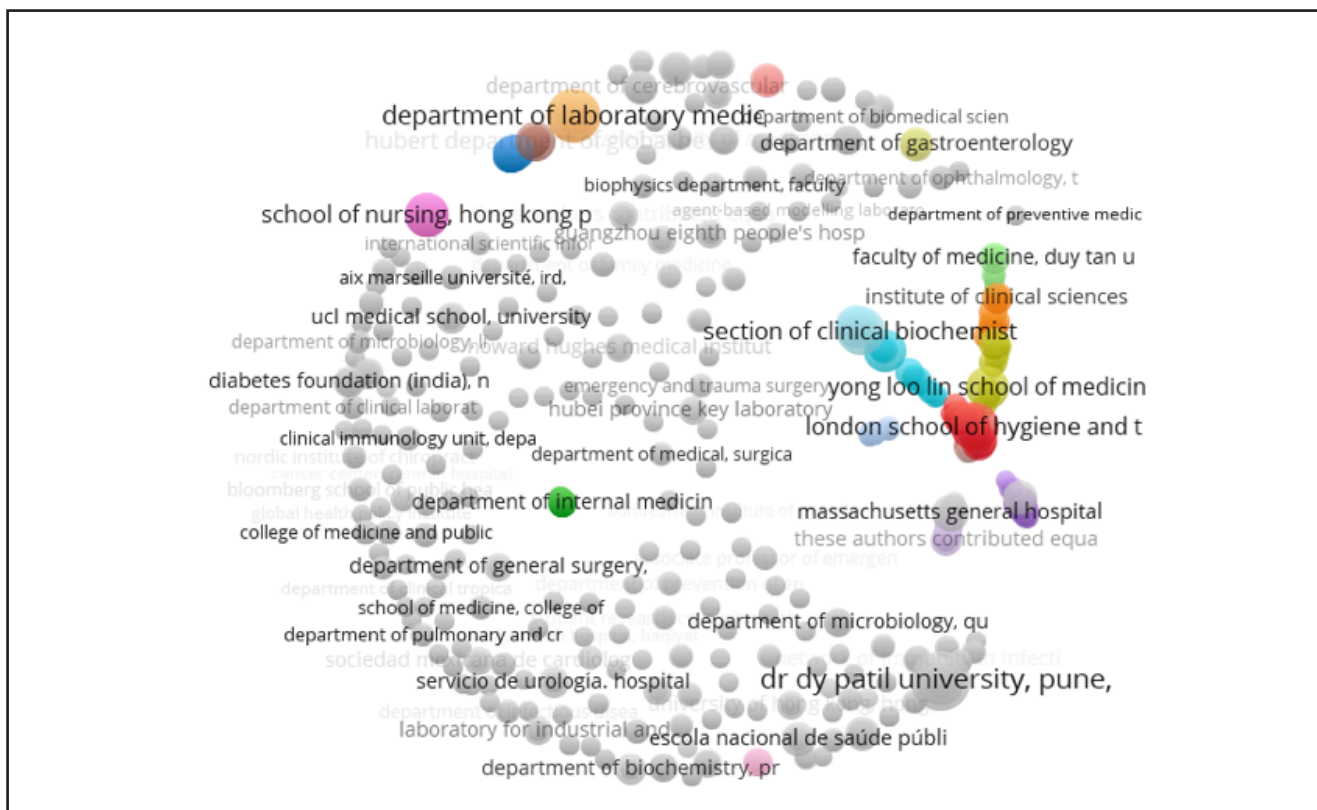


Figure 3. Publication journal correlation. The figure shows the journals with a minimum of two occurrences in the literature.