COVID-19 will stimulate a new coronavirus research breakthrough: a 20-year bibliometric analysis

Zhengbo Tao[1#], Siming Zhou[1#], Renqi Yao[2,3#], Kaicheng Wen[1], Wacili Da[1], Yan Meng[1], Keda Yang[1], Hang Liu[4], Lin Tao[1,5]

[1]Department of Orthopaedics, First Hospital of China Medical University, Shenyang 110001, China; [2]Trauma Research Center, Fourth Medical Center of the Chinese PLA General Hospital, Beijing 100048, China; [3]Department of Burn Surgery, Changhai Hospital, the Naval Medical University, Shanghai 200433, China; [4]Ragon Institute of Massachusetts General Hospital, MIT and Harvard University, Boston, USA; [5]Institute of Health Sciences of China Medical University, Shenyang 110001, China

Contributions: (I) Conception and design: L Tao, Z Tao; (II) Administrative support: L Tao; (III) Provision of study materials or patients: L Tao; (IV) Collection and assembly of data: Z Tao, S Zhou, R Yao; (V) Data analysis and interpretation: S Zhou, R Yao, K Wen, W Da, Y Meng; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

[#]These authors contributed equally to this work.

Correspondence to: Lin Tao, MD, PhD. Department of Orthopaedics, First Hospital of China Medical University, 155 Nan Jing North Street, Shenyang 110001, China. Email: taolindr@163.com.

Background: COVID-19 is currently rampant in China, causing unpredictable harm to humans. This study aimed to quantitatively and qualitatively investigate the research trends on coronaviruses using bibliometric analysis to identify new prevention strategies.

Methods: All relevant publications on coronaviruses were extracted from 2000–2020 from the Web of Science database. An online analysis platform of literature metrology, bibliographic item co-occurrence matrix builder (BICOMB) and CiteSpace software were used to analyse the publication trends. VOSviewer was used to analyse the keywords and research hotspots and compare COVID-19 information with SARS and MERS information.

Results: We found a total of 9,760 publications related to coronaviruses published from 2000 to 2020. The *Journal of Virology* has been the most popular journal in this field over the past 20 years. The United States maintained a top position worldwide and has provided a pivotal influence, followed by China. Among all the institutions, the University of Hong Kong was regarded as a leader for research collaboration. Moreover, Professors Yuen KY and Peiris JSM made great achievements in coronavirus research. We analysed the keywords and identified 5 coronavirus research hotspot clusters.

Conclusions: We considered the publication information regarding different countries, institutions, authors, journals, etc. by summarizing the literature on coronaviruses over the past 20 years. We analysed the studies on COVID-19 and the SARS and MERS coronaviruses. Notably, COVID-19 must become the research hotspot of coronavirus research, and clinical research on COVID-19 may be the key to defeating this epidemic.

Keywords: Coronavirus; COVID-19; bibliometric analysis; keywords; research hotspots

Submitted Feb 27, 2020. Accepted for publication Mar 18, 2020.

doi: 10.21037/atm.2020.04.26

View this article at: http://dx.doi.org/10.21037/atm.2020.04.26

© Annals of Translational Medicine. All rights reserved.

Introduction

Coronavirus is an enveloped positive-sense single-stranded RNA virus. Its diameter is approximately 80 to 120 nm. It has the largest genetic material among all RNA viruses. It can infect humans, mice, cats, dogs, birds and other vertebrates (1-3). Coronaviruses have proliferated many times throughout the world, causing unimaginable harm to humanity. Words such as SARS and MERS have produced great fear in people's hearts (4,5). There is no doubt that coronavirus has become a problem in the medical profession and even in society. However, with the COVID-19 outbreak in China, coronaviruses have once again become a focus (6).

COVID-19 is a new coronavirus strain that has never been found in humans before, and it is the seventh known coronavirus that can infect humans. It was discovered in a case of Wuhan viral pneumonia in 2019 and was named by the WHO on January 12, 2020 (7,8). As of this study, more than 100,000 people have been diagnosed with infection, and people of all ages can be infected. It has been confirmed that COVID-19 has the characteristics of human-to-human transmission and high concealment (7,9). Additionally, it has multiple transmission routes, including droplets, contact, and even aerosols, and the faecal-oral route may be included (10). Faced with this situation, to defeat the virus, there is still much work to be done by scientists in China and around the world.

In recent years, bibliometric analysis has become popular, which applies literature metrology characteristics to measure the contribution of an area of research, predicts detailed trends of research or hotspots in a certain field, and makes an important contribution to the prevention and treatment of diseases. However, there have been few bibliometric studies on coronaviruses, mainly focusing on MERS, and there is a lack of comprehensive analysis and research hotspot prediction for coronaviruses (11-13). In this article, we applied an integrated analysis of the content and external features of the research literature to summarize past research on coronaviruses and predict future research hotspots. We also provide an in-depth analysis of COVID-19 and summarize all the documented clinical trials to aid clinical treatment and scientific research.

Methods

Data sources and search strategies

Obviously, the Science Citation Index-Expanded and the Social Science Citation Index of Thomson Reuters' Web of Science must be the most appropriate databases to perform bibliometric analysis. We searched Web of Science database comprehensively from 2000 to 2020, and only original articles and reviews were included. The search strategy was presented as follow: TI = (coronavirus) AND Language = English. To avoid bias cursed by frequent database renewal, all the literature retrieval and data download were completed in a single day February 9, 2020.

Data collection

Two reviewers (ZT and SZ) independently performed the primary search and their agreement rate reached 0.90, showing a significant accordance (14). WoSCC data including titles, countries, institutions journals authors and so on, were extracted and imported into the Online Analysis Platform of Literature Metrology (http://bibliometric.com/), CiteSpace V5.5.R1 SE, 64bit (Drexel University, Philadelphia, PA, USA) and VOSviewer (Leiden University, Leiden, the Netherlands) for bibliometric analysis. And the clinic trials data was obtained from ClinicalTrials.gov (https://clinicaltrials.gov/).

Bibliometric analysis

We tried to describe all publication characteristics, including countries, institutions, journals, authors, H index, and so on. We inquired the 2018 version of JCR (Journal Citation Reports) to get the impact factor (IF), which was regarded as an important indicator to measure the scientific value of research (15). In our study, we analyzed the annual publication numbers and growth tendencies of different country/region through Literature Metrology online analysis platform. CiteSpace is an optimal means for collaboration network analysis to connect all kind of publication characteristics. It can also obtain keywords with high citations to predict the research frontiers and emerging trends in this area. CiteSpace can apply "time slicing" function, for example, if you set the "years per slice" to one while the "top N per slice" is set to fifty, and the top fifty papers in a year would be exported into a single file. According to our objective, nodes of different size represented citation counts or publication counts (16,17). In addition, VOSviewer can sort keywords into different clusters based on co-occurrence analysis, and color them at the same time according to time course.

© Annals of Translational Medicine. All rights reserved.

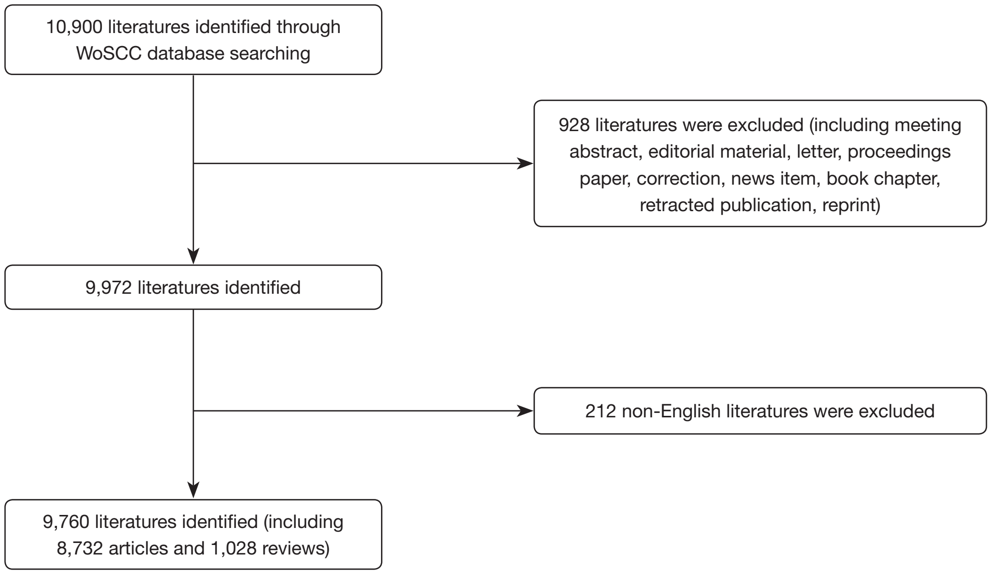


Figure 1 Flow chart of literature filtering included in this study.

Results

Contribution of countries and institutions to global publications

A total of 9,760 studies (8,732 articles and 1,028 reviews) met our inclusion criteria from 2000 to 2020 (*Figure 1*). *Figure 2A* displays a transformative trend in the annual literature numbers related to coronaviruses. All of the incorporated literature on coronaviruses was contributed by at least 114 different countries or regions (*Figure 2B*). The United States (n= 3,452) is the largest contributor to coronavirus research, followed by China (n=2,402), Germany (n=642), England (n=573), and the Netherlands (n=551). Centrality is a major indicator of the importance of nodes in a network, and a higher centrality means that a node is more important in a network, so the results showed that the United States has the most impact on other countries (centrality =0.24), followed by France (0.18) and England (0.15) (*Table 1*). In terms of research institutions, the top 10 include the University of Hong Kong (n=959), Chinese Academy of Sciences (n=469), Chinese University of Hong Kong (n=411), University of North Carolina (n=340), and University of Iowa (n=292) (*Table 1*). The coronavirus research network produces a low-density map (density =0.017) (*Figure 3A*), which means that the research teams are relatively scattered in various institutions, and increased mutual cooperation is needed. Most of the centrality indexes are below 0.15, indicating that the effect of most institutions stays at a low level and that the cooperation between institutions is insufficient. International cooperation analysis shows that the most frequent cooperation occurs in the United States and China (*Figure 3B*).

Journals publishing research on coronaviruses

Recently, 1,323 journals have published research in the coronavirus field. The top 10 popular journals published 2,621 of all 9,760 studies on coronaviruses (26.85%) (*Table 2*). Among them, the top 3 journals are the *Journal of Virology*, *Virology* and *PLoS One*, which account for more than 14.54% of all indexed literature. The highest IF belongs to *Emerging Infectious Diseases* (7.185), followed by the *Journal of Virology* (4.324) and *Viruses-Basel* (3.811). According to the JCR 2018 standards, 5 journals are classified as Q1, 2 journals as Q2 and 3 journals as Q3. An analysis of highly cited papers showed that the *New England Journal of Medicine* and *Science* have an incredible scientific impact on all scholars, and 6 of the top 10 highly cited papers were published in these two journals (*Table 3*).

Contribution of authors to coronavirus research

The ten authors that published the most papers, among all 29,515 authors, on this subject include Yuen KY, Baric RS, Perlman S, Drosten C, and Woo PCY (*Table 4*). Among

© Annals of Translational Medicine. All rights reserved.

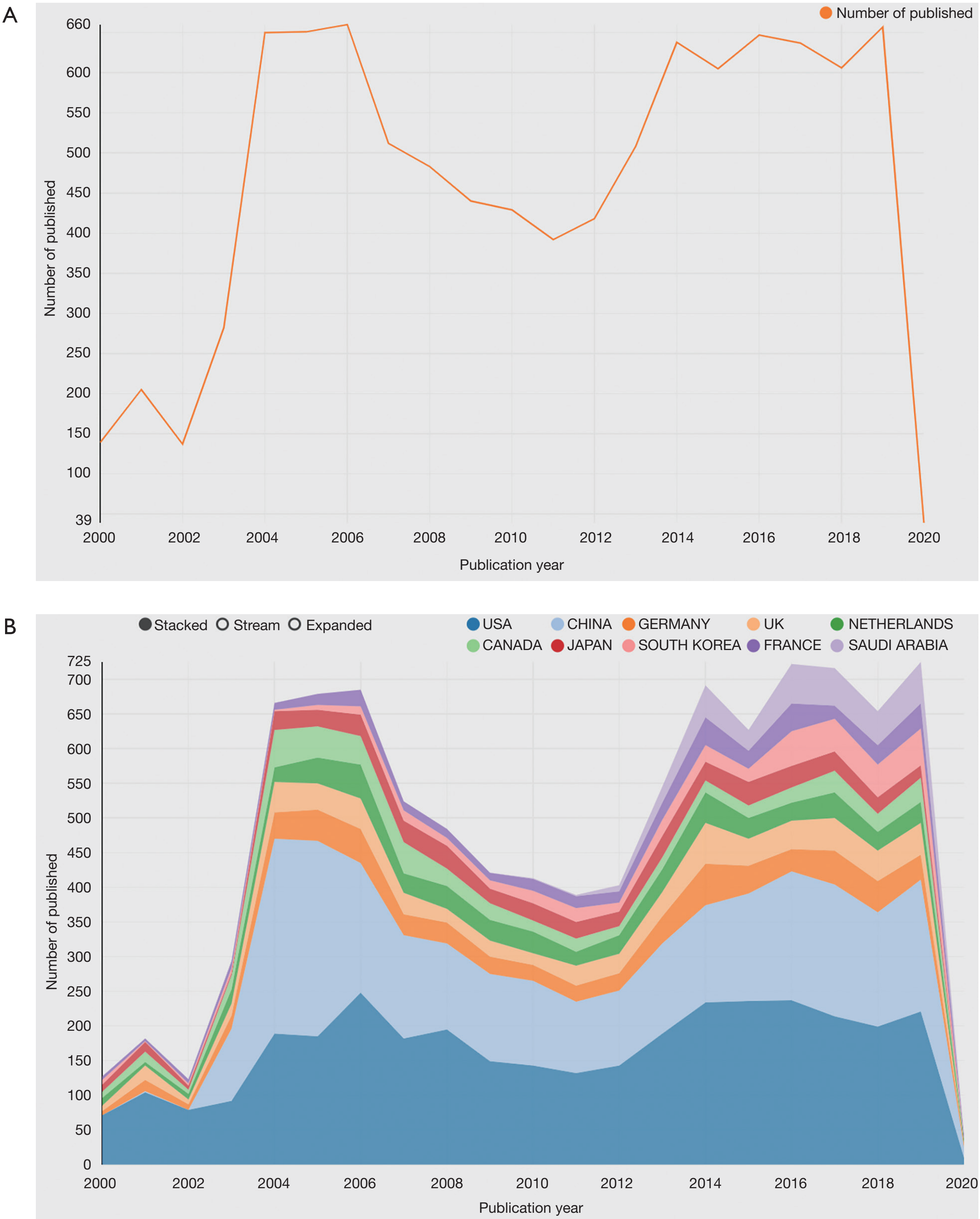


Figure 2 Output of related literature. The number of annual publications (A) and growth trends of the top 10 countries/regions (B) in coronavirus from 2000 to 2020.

© Annals of Translational Medicine. All rights reserved.

Table 1 The top 10 countries/regions and institutions contributing to publications in coronavirus research

Rank	Country/region	Article counts	Centrality	Institutions	Article counts	Total number of citations	Average number of citations	Total number of first author	Total number of first author citations	Average number of first author citations
1	USA	3,452	0.24	Univ Hong Kong	959	33,587	35.02	280	11,226	40.09
2	China	2,402	0.14	Chinese Acad Sci	469	9,870	21.04	173	2,588	14.96
3	Germany	642	0.13	Chinese Univ Hong Kong	411	6,874	16.73	133	1,778	13.37
4	England	573	0.15	Univ N Carolina	340	12,039	35.41	86	2,860	33.26
5	Netherlands	551	0.06	Univ Iowa	292	6,331	21.68	101	1,982	19.62
6	Canada	498	0.08	Ctr Dis Control & Prevent	269	13,860	51.52	68	2,474	36.38
7	Japan	465	0.04	Univ Utrecht	259	8,720	33.67	116	3,294	28.4
8	South Korea	392	0.01	Vanderbilt Univ	241	6,648	27.59	58	1,294	22.31
9	France	379	0.18	NIAID	221	7,584	34.32	84	3,242	38.6
10	Taiwan	373	0.01	Seoul Natl Univ	197	1,992	10.11	67	847	12.64

them, Yuen KY, the chair of Infectious Diseases at the University of Hong Kong, ranks first, with 200 studies; Baric RS from the Department of Epidemiology, Program in Infectious Diseases, University of North Carolina at Chapel Hill in the USA is second, with 134 studies. These two scholars have made great achievements and become authorities in coronavirus research. We analysed the citation information of authors (*Figure 4A*) and co-cited authors (*Figure 4B*), visualizing it in a network produced by CiteSpace. Peiris JSM, with 1,759 co-citations, ranks first among the top ten co-cited authors, followed by Drosten C (n=1,751), Ksiazek TG (n=1,431), and Rota PA (n=1,258) (*Table 4*).

Analysis of coronavirus research hotspots

Keywords were extracted from 9,760 publications and analysed by VOSviewer. In *Figure 5A*, 216 keywords that appeared more than 200 times were included and classified into 5 clusters in the map: cluster 1 (clinical research, in red); cluster 2 (pathogenesis research, in green); cluster 3 (virological research, in blue); cluster 4 (treatment, in yellow) and cluster 5 (origin and transmission research, in purple). Circles with a large size represent the keywords that appeared at a high frequency. Within cluster 1, the following keywords frequently occurred: study (4,070 times), infection (4,057 times), disease (2,462 times), sample (1,672 times) and patient (1,641 times). In cluster 2, relevant keywords included protein (2,653 times), cell (2,381 times), role (1,575 times) and activity (1,393 times). In cluster 3, the primary keywords were virus (4,810 times), coronavirus (3,715 times), analysis (2,194 times), gene (1,624 times) and strain (1,562 times). Similarly, in cluster 4, the main keywords were antibody (1,207 times), assay (1,165 times), specificity (477 times), sensitivity (449 times) and evaluation (383 times). In cluster 5, they were human (920 times), species (719 times), identification (703 times), approach (659 times) and host (620 times). Detailed consequences of keywords are provided in *Table S1*. In *Figure 5B*, all keywords were coloured according to the average time of word appearance, from blue to yellow, representing early to recent appearances, respectively. We analysed the temporal trend of research hotspot shifts according to the top 25 keywords with the strongest citation bursts from 2000 to 2020 (*Figure 6*).

Discussion

Our statistical and quantitative analysis showed that the research output on coronavirus has fluctuated in the last 20 years. In *Figure 2A,B*, it can be seen that there was an explosion of research in this area during 2003–2006, with

© Annals of Translational Medicine. All rights reserved.

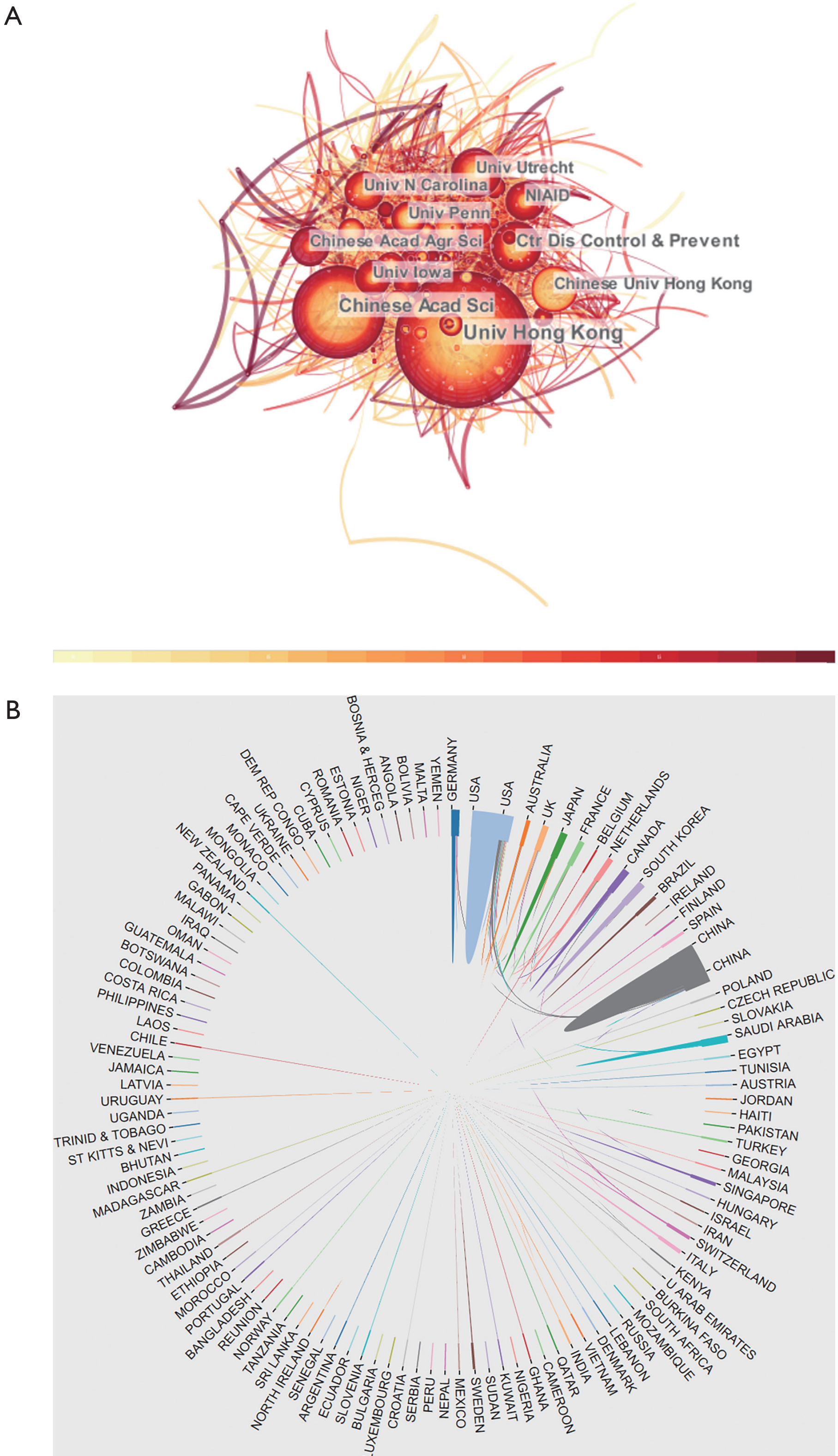


Figure 3 The distribution of countries/regions and institutions. The network map of institutions that involved in coronavirus research (A) and the cooperation of countries/regions (B).

© Annals of Translational Medicine. All rights reserved.

Table 2 The top 10 most active journals that published articles in coronavirus research (sorted by count)

Rank	Journal title	Percentage (N/9,760), %	IF [2018]	Quartile in category [2018]	H-index	Article counts	Total number of citations	Average number of citations
1	*Journal of Virology*	9.07	4.324	Q1	271	885	26,285	29.7
2	*Virology*	3.03	2.657	Q2	162	296	5,063	17.1
3	*PLoS One*	2.44	2.776	Q1	268	238	1,633	6.86
4	*Emerging Infectious Diseases*	2.09	7.185	Q1	202	204	5,612	27.51
5	*Journal of General Virology*	1.99	2.809	Q2	152	194	4,033	20.79
6	*Virus Research*	1.97	2.736	Q2	104	192	2,534	13.2
7	*Viruses-Basel*	1.70	3.811	Q1	59	166	999	6.02
8	*Archives of Virology*	1.59	2.261	Q3	102	155	1,529	9.86
9	*Journal of Virological Methods*	1.52	1.746	Q3	91	148	1,141	7.71
10	*Veterinary Microbiology*	1.47	2.791	Q1	114	143	1,441	10.08

Table 3 The top 10 high-cited papers in coronavirus research during 2000 to 2020

Rank	Title	Journal	Corresponding authors	Publication year	Total citations
1	A novel coronavirus associated with severe acute respiratory syndrome	*New England Journal of Medicine*	Perlman S	2003	1,827
2	Identification of a novel coronavirus in patients with severe acute respiratory syndrome	*New England Journal of Medicine*	Drosten C	2003	1,734
3	Characterization of a novel coronavirus associated with severe acute respiratory syndrome	*Science*	Rota PA	2003	1,488
4	Coronavirus as a possible cause of severe acute respiratory syndrome	*Lancet*	Yuen KY	2003	1,437
5	Isolation of a Novel Coronavirus from a Man with Pneumonia in Saudi Arabia	*New England Journal of Medicine*	Chan KH	2012	1,276
6	The genome sequence of the SARS-associated coronavirus	*Science*	Marra MA	2003	1,274
7	Cloning of a human parvovirus by molecular screening of respiratory tract samples	*Proceedings of The National Academy of Sciences of The United States of America*	Allander T	2005	1,012
8	Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus	*Nature*	Choe H & Farzan M	2003	968
9	Isolation and characterization of viruses related to the SARS coronavirus from animals in Southern China	*Science*	Guan Y	2003	882
10	Bats are natural reservoirs of SARS-like coronaviruses	*Science*	Shi ZL & Zhang SY & Wang LF	2005	841

© Annals of Translational Medicine. All rights reserved.

Table 4 The top 10 most productive authors and co-cited authors contributed to publications in coronavirus research

Rank	Author	Article Counts	Total number of citations	Average number of citations	First author counts	First author citations counts	Average first author citations counts	Corresponding author counts	Corresponding author citation counts	Co-cited author	Citation counts
1	Yuen KY	200	10,405	52.02	0	0	0	85	5,988	Peiris JSM	1,759
2	Baric RS	134	4,044	30.18	4	71	17.75	53	2,154	Drosten C	1,751
3	Perlman S	133	2,873	21.6	6	339	56.5	71	1,760	Ksiazek TG	1,431
4	Drosten C	128	7,944	62.06	7	1,866	266.57	37	2,849	Rota PA	1,258
5	Woo PCY	117	4,671	39.92	38	2,114	55.63	32	664	Woo PCY	1,232
6	Enjuanes L	115	3,352	29.15	7	145	20.71	60	2,050	Marra MA	1,060
7	Chan KH	113	7,414	65.61	5	141	28.2	1	2	Zaki AM	978
8	Lau SKP	109	4,467	40.98	36	1,494	41.5	3	102	Lau SKP	926
9	Snijder EJ	89	3,898	43.8	4	767	191.75	34	1,634	Cavanagh D	902
10	Peiris JSM	88	6,973	79.24	5	2,192	438.4	15	1,057	Li WH	863

China and the United States contributing the most. There is no doubt that this increase is attributable to SARS in 2003. During that disaster, more than 5,000 people were infected with SARS coronavirus, including many medical staff, which caused massive panic worldwide. At that time, many scientists performed a multitude of research in this field, but after that, research on coronaviruses gradually decreased until 2012, when the outbreak of MERS caused research on coronaviruses to reach its peak again.

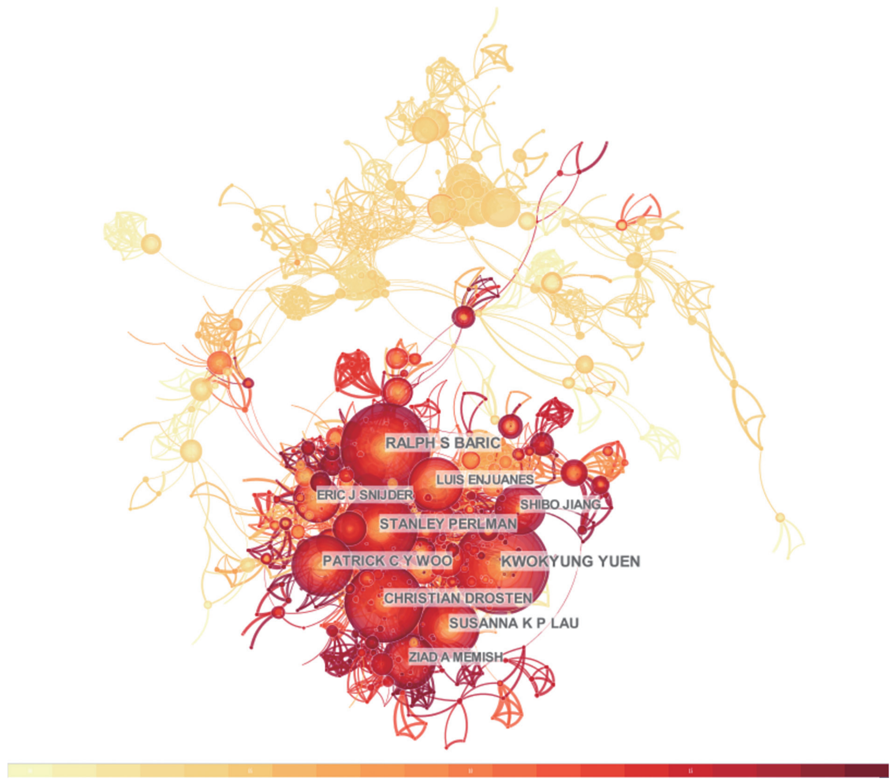
Regarding the contributions of countries and institutions, both the United States and China have played an important role in coronavirus research, and their total numbers of studies rank first and second, respectively. The United States seems to have superior conditions for basic medical research or clinical trials, which include adequate funding, advanced equipment and professional researchers. All the characteristics also show that the United States is leading the field. However, three institutions from China (the University of Hong Kong, Chinese Academy of Sciences and the Chinese University of Hong Kong) are ahead of scientific agencies in other regions. This phenomenon is partly because China was the main place where SARS occurred, and it also shows that the strength of scientific research from China has continuously increasing in recent years. The largest current problem is insufficient cooperation between various countries and institutions, which greatly reduces the efficiency of research. If there is improved communication and cooperation between institutions in various countries, I believe that research on viruses and diseases will achieve an enormous breakthrough.

Notably, the *Journal of Virology* published 885 studies in this area, far ahead of other journals. Other journals, including *Virology*, *PLoS One* and *Emerging Infectious Diseases*, were the primary journals containing coronavirus publications. In addition, the *New England Journal of Medicine* and *Science* focused on coronavirus research, and many highly cited papers were published in them. Thus, these findings imply that future developments in the field may be published in the aforementioned journals. Additionally, authors such as Yuen KY, Baric RS, Perlman S, and Drosten C not only published the largest numbers of papers in this field but also published their own highly cited representative papers in top magazines. Obviously, this publication record demonstrates that they have become an influential core group in the coronavirus field, having carried out substantial research to lay a solid foundation for future development.

We identified five keyword clusters to analyse research hotspots on coronaviruses. We found that the study of coronaviruses is relatively comprehensive, including clinical research, pathogenesis research, virological research, origin and transmission research and disease treatment method research. For research on coronavirus, we first need to understand its infectious disease characteristics, including its origin, susceptible population and transmission route, and then analyse its pathogenic mechanism and

© Annals of Translational Medicine. All rights reserved.

A



B

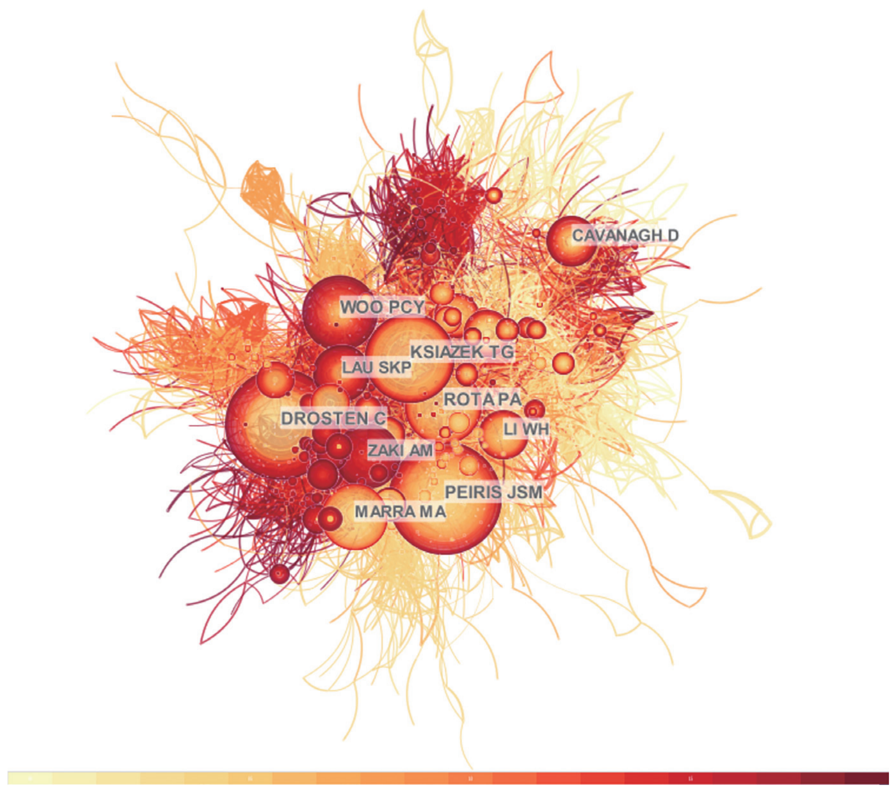


Figure 4 The distribution of authors engaged in coronavirus research. The network map of productive authors (A) and the network map of co-cited authors (B).

© Annals of Translational Medicine. All rights reserved.

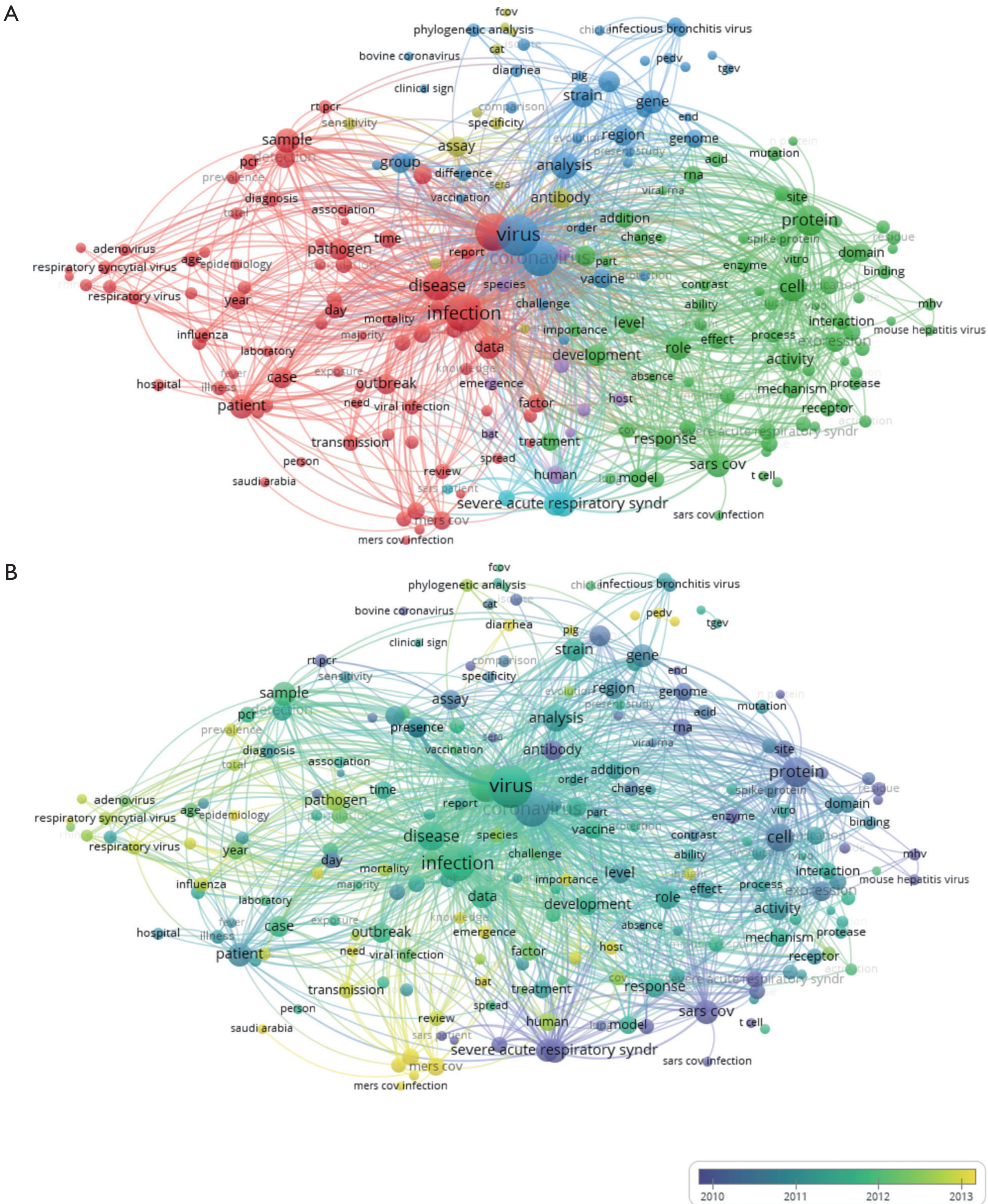


Figure 5 The analysis of keywords in publications of coronavirus research. Mapping of the keywords in the area of coronavirus (A). Distribution of keywords was presented according to the appearance for the average time (B).

Top 25 Keywords with the Strongest Citation Bursts

Keywords	Strength	Begin	End	2000–2020
Acute respiratory syndrome	157.4016	2003	2009	
Hong Kong	108.228	2003	2006	
Saudi arabia	98.3245	2014	2020	
Mers-cov	82.6629	2014	2020	
Severe acute respiratory syndrome	56.625	2003	2006	
Dromedary camel	55.3531	2014	2020	
Mouse hepatitis virus	52.0487	2000	2003	
Respiratory syndrome coronavirus	51.5933	2014	2018	
Mers cov	50.8162	2014	2020	
Outbreak	49.013	2003	2004	
Sar	48.9519	2003	2005	
Sars coronavirus	48.6472	2005	2009	
Genome sequence	48.4025	2003	2005	
Angiotensin converting enzyme 2	47.0054	2005	2009	
Sars associated coronavirus	38.5945	2004	2008	
Mer	35.7769	2015	2020	
Sars-cov	32.443	2005	2009	
Murine coronavirus	32.1864	2000	2003	
Transmission	30.9364	2016	2020	
Middle east	28.1641	2014	2020	
East respiratory syndrome	28.0234	2016	2020	
Identification	27.0364	2003	2005	
United states	27.0076	2015	2018	
Mouse model	26.1527	2015	2020	
Monoclonal antibody	25.3786	2000	2003	

Figure 6 The top 25 keywords with the strongest citation bursts during 2000 to 2020.

viral gene sequence to further find effective treatments and start clinical trials. In addition, the temporal trend of research hotspot shifts showed that the research in this field transferred mainly from early SARS to later MERS, which suggests that the increase in these studies was accompanied by the emergence of research hotspot events. The research increase had a very obvious time lag, which made us unprepared to deal with the emergencies. Therefore, we need to pay constant attention to various coronaviruses and their variants to prevent the emergence of large-scale infectious diseases.

For COVID-19, although there are still few articles, we still summarized many of its bibliometric characteristics and compared them with those of SARS and MERS after our collection and analysis (*Table 5*). COVID-19 has similarity in gene sequence with SARS, and they have a common origin, bats, and a common intracellular receptor, ACE2 (18). Thus, the symptoms of COVID-19 are also similar to those of SARS, often manifesting as fever, cough, shortness of breath, or breathing difficulty, and in severe cases, pneumonia or even death may occur (19,20). However, COVID-19 is more concealed and more contagious than SARS. In the latest study from Guan *et al.*, only 43.1% of patients had a fever when they were admitted to the hospital, and more patients had a fever during their hospital stay (21). For SARS or MERS, both of which are coronavirus induced, almost all patients have fever symptoms when diagnosed, and only 1–2% do not have a fever (22). This presentation means that if the screening of suspected cases relies only on measuring body temperature during epidemic prevention and control, then a large number of infected persons with no fever may be missed. After the Chinese Spring Festival, it is difficult to predict whether a second outbreak will occur as a large number of people return to work all over the country. Thus, this outbreak will be more difficult to address than SARS in 2003 (23).

Many scientists published a large number of articles after

© Annals of Translational Medicine. All rights reserved.

Table 5 The general and bibliometric information about SARS, MERS and COVID-19

	SARS	MERS	COVID-19
Appear time	2003	2012	2019–2020
Appear place	China	Saudi Arabia	China
Origin	Bat, masked palm civet	Bat, dromedary	Bat
Receptor	ACE2	DPP4	ACE2
Hotspots	Structure, origin, pathogenic mechanism, clinic research	Origin, antibody, clinic research	CT manifestations, genome, case series, clinical characteristics
Journals (the most counts)	*Virology*	*Virology*	*Lancet*
Representative articles	1. "A novel coronavirus associated with severe acute respiratory syndrome" *New England Journal of Medicine*	1. "Hospital Outbreak of Middle East Respiratory Syndrome Coronavirus" *New England Journal of Medicine*	1. "Clinical Characteristics of 138 Hospitalized Patients with 2019 Novel Coronavirus-Infected Pneumonia in Wuhan, China" *JAMA*
	2. "Identification of a novel coronavirus in patients with severe acute respiratory syndrome" *New England Journal of Medicine*	2. "Epidemiological, demographic, and clinical characteristics of 47 cases of Middle East respiratory syndrome coronavirus disease from Saudi Arabia: a descriptive study" *Lancet infectious Diseases*	2. "Updated understanding of the outbreak of 2019 novel coronavirus (2019-nCoV) in Wuhan, China" *Journal of Medical Virology*
	3. "Characterization of a novel coronavirus associated with severe acute respiratory syndrome" *Science*	3. "Middle East respiratory syndrome coronavirus neutralising serum antibodies in dromedary camels: a comparative serological study" *Lancet infectious Diseases*	3. "Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding" *Lancet*
Representative author	Perlman S, Drosten C	Memish ZA, Drosten C	Drosten C, Zhong NS

SARS occurred in 2003. Their main focus has been virus structure, origin, and pathogenic mechanism and clinical research. The journal *Virology* published the most literature, and there are some great studies published in top journals, such as the *New England Journal of Medicine* and *Science*. A similar situation occurred after the emergence of MERS. For COVID-19, although the number of articles is small, in a short period of time, many studies have been accepted by different top journals, such as *JAMA* and *Lancet*. Notably, we found that Drosten C has achieved great success in SARS, MERS and COVID-19 research (5,10). There is no doubt that he has become an authority in the field of coronavirus research. At this stage, CT manifestations, genomic sequence, and clinical characteristics are regarded as research hotspots in COVID-19 research, which is of great significance for the further prevention and control of disease. In addition, effective blockade of infectious pathways is also very important for disease prevention, and people now pay much attention to blocking droplets and air, but commonly touched objects cannot be ignored, such as elevators and shoes. Finally, we searched ClinicalTrials.gov and found 18 documented clinical trials (*Table 6*). Oxygen therapy, mechanical ventilation, empirical use of antibiotics and oseltamivir antivirals are the main methods currently used. Remdesivir and chloroquine, which have high potential, have been used in the clinic; although symptoms have improved, the therapeutic effects and side effects need further clinical trial verification. Immunoglobulin infusion, ECMO and other methods also have a certain effect on severe patients, and Chinese medicine should also be a good consideration for patients with different durations of infection. However, what is currently important is the development of new therapeutic drugs and vaccines, which may play a decisive role in defeating COVID-19. I believe

© Annals of Translational Medicine. All rights reserved.

Table 6 The documented clinical trials about COVID-19 (18 items)

	Study title	Conditions	Interventions
1	Mild/Moderate 2019-nCoV Remdesivir RCT	2019-nCoV	Drug: remdesivir; drug: remdesivir placebo
2	The Efficacy of Intravenous Immunoglobulin Therapy for Severe 2019-nCoV Infected Pneumonia	2019-nCoV	Drug: intravenous immunoglobulin; other: standard care
3	A Prospective, Randomized Controlled Clinical Study of Antiviral Therapy in the 2019-nCoV Pneumonia	2019-nCoV	Drug: abidol hydrochloride; drug: oseltamivir; Drug: lopinavir/ritonavir
4	A Prospective, Randomized Controlled Clinical Study of Interferon Atomization in the 2019-nCoV Pneumonia	2019-nCoV	Drug: abidol hydrochloride; drug: abidol hydrochloride combined with interferon atomization
5	A Randomized, Open, Controlled Clinical Study to Evaluate the Efficacy of ASC09F and Ritonavir for 2019-nCoV Pneumonia	2019-nCoV; pneumonia	Drug: ASC09F+oseltamivir; drug: ritonavir + oseltamivir Drug: oseltamivir
6	Clinical Study of Arbidol Hydrochloride Tablets in the Treatment of Pneumonia Caused by Novel Coronavirus	2019-nCoV	Drug: arbidol; other: basic treatment
7	Severe 2019-nCoV Remdesivir RCT	2019-nCov; remdesivir	Drug: remdesivir; drug: remdesivir placebo
8	Mesenchymal Stem Cell Treatment for Pneumonia Patients Infected With 2019 Novel Coronavirus	2019 novel coronavirus pneumonia	Biological: MSCs
9	Efficacy of a Self-test and Self-alert Mobile Applet in Detecting Susceptible Infection of 2019-nCoV	Susceptibility to viral and mycobacterial infection	Other: mobile internet survey on self-test
10	Development of a Simple, Fast and Portable Recombinase Aided Amplification Assay for 2019-nCoV	New coronavirus	Diagnostic test: recombinase aided amplification (RAA) assay
11	2019-nCoV Outbreak and Cardiovascular Diseases	Cardiovascular death; major adverse cardiovascular events	
12	Viral Excretion in Contact Subjects at High/Moderate Risk of Coronavirus 2019-nCoV Infection	Coronavirus	Biological: 2019-nCoV PCR
13	Efficacy and Safety of Darunavir and Cobicistat for Treatment of Pneumonia Caused by 2019-nCoV	Pneumonia, pneumocystis; coronavirus	Drug: darunavir and cobicistat
14	Efficacy and Safety of Hydroxychloroquine for Treatment of Pneumonia Caused by 2019-nCoV (HC-nCoV)	Pneumonia, pneumocystis; coronavirus	Drug: hydroxychloroquine
15	Treatment and Prevention of Traditional Chinese Medicines (TCMs) on 2019-nCoV Infection	Pneumonia caused by human coronavirus (disorder)	Drug: conventional medicines (oxygen therapy, alfa interferon via aerosol inhalation, and lopinavir/ritonavir) and traditional Chinese medicines (TCMs) granules; drug: conventional medicines (oxygen therapy, alfa interferon via aerosol inhalation, and lopinavir/ritonavir)

Table 6 (*continued*)

© Annals of Translational Medicine. All rights reserved.

Table 6 (*continued*)

	Study title	Conditions	Interventions
16	A Survey of Psychological Status of Medical Workers and Residents in the Context of 2019 Novel Coronavirus Pneumonia	Virus; pneumonia	
17	Glucocorticoid Therapy for Novel Coronavirus Critically Ill Patients With Severe Acute Respiratory Failure	Coronavirus infections; respiratory infection virus	Drug: methylprednisolone therapy; other: standard care
18	Washed Microbiota Transplantation for Patients With Coronavirus Pneumonia	Virus pneumonia	Other: washed microbiota transplantation; other: placebo

that these clinical trials will provide reliable support for clinical research and have great guiding significance for the formulation of future therapeutic schedules.

Nonetheless, some limitations may be inevitable. The database updates continuously, and we selected only the literature from 2000 to February 9, 2020, without literature published after that day. Therefore, there is a discrepancy between our bibliometric analysis and real publication conditions. The number of coronavirus studies may increase rapidly with the breakthrough of future research.

Conclusions

We assessed the publication information regarding different countries, institutions, authors, journals, etc. and analysed the research hotspots in the coronavirus field over the past 20 years based on these studies. COVID-19 must become the focus of coronavirus research in the near future. In addition, reviewing previous coronavirus studies and determining their similarities and differences with those on COVID-19 will help us to understand this new virus as soon as possible. Finally, clinical research on coronaviruses, especially randomized controlled trials, has great potential to guide the prevention and treatment of coronaviruses in the future. We believe our research can reflect novel directions for coronavirus research and help the Chinese people overcome this epidemic soon.

Acknowledgments

Funding: None.

Footnote

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at http://dx.doi.org/10.21037/atm.2020.04.26). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: https://creativecommons.org/licenses/by-nc-nd/4.0/.

References

1. Ziebuhr J. Molecular biology of severe acute respiratory syndrome coronavirus. Curr Opin Microbiol 2004;7:412-9.
2. Sawicki SG, Sawicki DL, Siddell SG. A contemporary view of coronavirus transcription. J Virol 2007;81:20-9.
3. Hagemeijer MC, Rottier PJ, de Haan CA. Biogenesis and dynamics of the coronavirus replicative structures. Viruses 2012;4:3245-69.
4. Cui J, Li F, Shi ZL. Origin and evolution of pathogenic coronaviruses. Nat Rev Microbiol 2019;17:181-92.
5. Drosten C, Gunther S, Preiser W, et al. Identification of a novel coronavirus in patients with severe acute respiratory syndrome. N Engl J Med 2003;348:1967-76.
6. Wang C, Horby PW, Hayden FG, et al. A novel

© Annals of Translational Medicine. All rights reserved.

coronavirus outbreak of global health concern. Lancet 2020;395:470-3.
7. Zhu N, Zhang D, Wang W, et al. A Novel Coronavirus from Patients with Pneumonia in China, 2019. N Engl J Med 2020;382:727-33.
8. WHO. Global research and innovation forum to mobilize international action in response to the novel coronavirus (2019-nCoV) emergency. 2020.
9. Chan JFW, Yuan S, Kok KH, et al. A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: a study of a family cluster. Lancet 2020;395:514-23.
10. Rothe C, Schunk M, Sothmann P, et al. Transmission of 2019-nCoV Infection from an Asymptomatic Contact in Germany. N Engl J Med 2020;382:970-1.
11. Bonilla-Aldana DK, Quintero-Rada K, Montoya-Posada JP, et al. SARS-CoV, MERS-CoV and now the 2019-novel CoV: Have we investigated enough about coronaviruses? - A bibliometric analysis. Travel Med Infect Dis 2020;33:101566.
12. Wang Z, Chen Y, Cai G, et al. A Bibliometric Analysis of PubMed Literature on Middle East Respiratory Syndrome. Int J Environ Res Public Health 2016;13:583.
13. Zyoud SH. Global research trends of Middle East respiratory syndrome coronavirus: a bibliometric analysis. BMC Infect Dis 2016;16:255.
14. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33:159-74.
15. Eyre-Walker A, Stoletzki N. The assessment of science: the relative merits of post-publication review, the impact factor, and the number of citations. PLoS Biol 2013;11:e1001675.
16. Chen C. CiteSpace II: Detecting and visualizing emerging trends and transient patterns in scientific literature. J Am Soc Inf Sci Technol 2014;57:359-77.
17. Chen C, Ibekwe-Sanjuan F, Hou J, et al. The structure and dynamics of cocitation clusters: A multiple-perspective cocitation analysis. J Am Soc Inf Sci Technol 2010;61:1386-409.
18. Zhou P, Yang XL, Wang XG, et al. Discovery of a novel coronavirus associated with the recent pneumonia outbreak in humans and its potential bat origin. bioRxiv 2020. doi: https://doi.org/10.1101/2020.01.22.914952.
19. Jin YH, Cai L, Cheng ZS, et al. A rapid advice guideline for the diagnosis and treatment of 2019 novel coronavirus (2019-nCoV) infected pneumonia (standard version). Mil Med Res 2020;7:4.
20. Huang C, Wang Y, Li X, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. Lancet 2020;395:497-506.
21. Guan WJ, Ni ZY, Hu Y, et al. Clinical characteristics of 2019 novel coronavirus infection in China. medRxiv 2020. doi: https://doi.org/10.1101/2020.02.06.20020974
22. Zumla A, Hui DS, Perlman S. Middle East respiratory syndrome. Lancet 2015;386:995-1007.
23. Chen J. Pathogenicity and transmissibility of 2019-nCoV-A quick overview and comparison with other emerging viruses. Microbes Infect 2020;22:69-71.

Cite this article as: Tao Z, Zhou S, Yao R, Wen K, Da W, Meng Y, Yang K, Liu H, Tao L. COVID-19 will stimulate a new coronavirus research breakthrough: a 20-year bibliometric analysis. Ann Transl Med 2020;8(8):528. doi: 10.21037/atm.2020.04.26

© Annals of Translational Medicine. All rights reserved.

Table S1 The analytic consequence of 216 keywords with at least 200 occurrence times

Rank	Keywords	Cluster	Links	Occurences	Average appearing years (AAY)	Average citations
1	Ability	2	215	524	2011.2	30.77
2	Absence	2	215	327	2010.5	34.14
3	Acid	2	215	656	2010.5	29.89
4	Activation	2	212	461	2012.0	27.71
5	Activity	2	215	1,393	2011.0	30.34
6	Addition	2	215	972	2011.1	31.47
7	Adenovirus	1	209	420	2012.2	30.62
8	Age	1	213	676	2012.0	32.05
9	Analysis	3	215	2,194	2011.2	28.01
10	Animal	3	215	813	2011.3	29.43
11	Antibody	4	215	1,207	2010.0	27.94
12	Approach	5	215	659	2011.7	30.70
13	Assay	4	215	1,165	2010.5	26.33
14	Association	1	214	442	2011.4	38.84
15	Bat	5	208	366	2014.3	40.95
16	Binding	2	211	613	2010.7	32.41
17	Bovine coronavirus	3	206	265	2010.0	20.14
18	Case	1	215	1,272	2011.8	30.95
19	Cat	4	211	375	2010.7	20.10
20	Causative agent	3	215	229	2010.5	39.18
21	Cause	1	215	525	2011.0	39.95
22	Cell	2	215	2,381	2010.5	32.58
23	Cell culture	3	215	303	2010.3	40.86
24	Challenge	3	215	443	2012.0	27.76
25	Change	2	215	710	2010.4	32.62
26	Characterization	3	215	653	2011.2	34.76
27	Chicken	3	201	320	2012.0	22.06
28	Child	1	208	537	2011.1	39.77
29	China	3	215	561	2012.0	35.67
30	Clinical sign	3	208	274	2011.3	20.23
31	Combination	4	215	325	2011.2	28.62
32	Comparison	3	215	446	2010.6	26.85
33	Compound	2	204	312	2011.1	30.10
34	Contrast	2	215	479	2010.5	30.10
35	Control	1	215	857	2011.3	28.67
36	Coronavirus	3	215	3,715	2010.8	33.71
37	Coronavirus infection	1	215	544	2011.0	26.60
38	Country	1	213	462	2013.3	26.32
39	Cov	2	215	641	2012.5	31.62
40	Data	1	215	1,584	2011.7	31.72
41	Day	1	215	829	2010.3	28.79
42	Death	1	215	460	2011.6	30.02
43	Detection	1	215	1,223	2011.4	27.87
44	Development	2	215	1,368	2011.6	30.57
45	Diagnosis	1	213	695	2010.9	29.44
46	Diarrhea	3	214	451	2013.4	19.54
47	Difference	3	215	677	2011.7	23.87
48	Discovery	5	212	325	2011.9	43.07
49	Disease	1	215	2,462	2011.6	29.10
50	Domain	2	213	942	2010.8	34.01
51	Effect	2	215	1,089	2011.0	24.88
52	Efficacy	3	215	332	2011.7	27.42
53	Elisa	4	213	336	2010.0	15.94
54	Emergence	5	215	510	2013.2	31.91
55	End	3	214	236	2009.6	28.25
56	Entry	2	215	565	2011.6	31.80
57	Enzyme	2	214	679	2009.5	38.82
58	Epidemiology	1	203	404	2013.0	28.84
59	Evaluation	4	215	383	2011.4	22.12
60	Evidence	1	215	805	2011.3	35.71
61	Evolution	3	214	395	2012.5	29.64
62	Exposure	1	215	344	2011.6	32.58
63	Expression	2	215	1,145	2010.5	32.19
64	Factor	1	215	1,012	2012.2	28.24
65	Fcov	4	188	205	2011.9	16.04
66	Feline coronavirus	4	199	274	2011.3	16.72
67	Fever	1	208	314	2010.8	44.81
68	Fip	4	187	229	2011.7	16.34
69	Function	2	215	899	2010.7	34.77
70	Gene	3	215	1,624	2010.7	28.97
71	Genome	3	215	978	2010.3	38.43
72	Group	3	215	1,212	2010.9	33.09
73	Hcov	1	213	355	2011.5	36.34
74	Hospital	1	197	469	2010.8	27.48
75	Host	5	215	620	2013.0	36.52
76	Host cell	2	211	273	2011.9	31.19
77	Human	5	215	920	2012.5	41.15
78	Human coronavirus	1	215	718	2011.8	34.58
79	Human metapneumovirus	1	187	282	2012.3	42.50
80	Ibv	3	205	573	2011.1	21.02
81	Identification	5	215	703	2010.9	38.51
82	Illness	1	215	599	2010.9	35.75
83	Immune response	2	214	695	2011.6	26.38
84	Importance	2	215	531	2012.9	28.00
85	Important role	2	215	299	2011.5	21.78
86	Induction	2	213	334	2011.0	33.65
87	Infected cell	2	209	284	2009.7	31.03
88	Infection	1	215	4,057	2011.7	31.05
89	Infectious bronchitis virus	3	204	591	2010.8	21.57
90	Infectious disease	1	215	488	2011.0	29.48
91	Influenza	1	210	647	2012.7	27.88
92	Influenza virus	1	211	450	2012.9	30.12
93	Information	1	215	537	2012.0	29.08
94	Inhibition	2	213	478	2011.3	29.16
95	Inhibitor	2	213	654	2010.8	33.35
96	Insight	2	215	452	2012.8	28.57
97	Interaction	2	215	901	2011.3	29.58
98	Interferon	2	214	289	2011.8	34.10
99	Isolate	3	214	469	2010.2	36.77
100	Isolation	3	212	330	2010.3	39.68
101	Knowledge	1	215	432	2013.2	24.19
102	Laboratory	1	213	371	2011.9	33.31
103	Lack	1	215	255	2012.2	26.94
104	Level	2	215	1214	2010.9	26.39
105	Lung	2	215	376	2010.2	34.18
106	Majority	1	215	238	2011.7	27.53
107	Mechanism	2	215	1,117	2011.5	31.26
108	Member	2	215	395	2011.0	38.84
109	Membrane	2	213	512	2010.3	33.74
110	Mer	1	211	726	2016.5	23.20
111	Mers cov	1	215	1,019	2016.4	23.60
112	Mers cov infection	1	199	261	2016.3	29.80
113	Mhv	2	200	445	2008.0	31.09
114	Middle east respiratory syndrome	1	207	325	2016.6	18.64
115	Middle east respiratory syndrome coronavirus	1	213	899	2016.3	24.77
116	Model	2	215	1,009	2011.5	27.45
117	Month	1	215	367	2010.7	29.43
118	Mortality	1	215	609	2012.6	27.45
119	Mouse	2	215	816	2010.2	27.98
120	Mouse hepatitis virus	2	198	404	2007.7	30.73
121	Mutation	2	214	597	2011.0	25.83
122	N protein	2	207	279	2009.3	24.11
123	Need	1	215	372	2013.1	25.16
124	Neutralizing antibody	2	212	292	2011.0	39.57
125	None	1	213	203	2010.8	33.89
126	Novel coronavirus	1	211	280	2008.5	71.96
127	Number	1	215	906	2011.4	29.81
128	Order	3	215	422	2011.2	29.35
129	Outbreak	1	215	1,215	2011.8	32.89
130	Parainfluenza virus	1	194	302	2012.7	32.04
131	Part	2	215	386	2010.7	36.67
132	Pathogen	1	215	1,343	2012.6	30.15
133	Pathogenesis	2	215	779	2011.3	33.54
134	Pathway	2	214	547	2012.3	30.14
135	Patient	1	214	1,641	2010.8	36.71
136	Pcr	1	215	776	2011.8	31.50
137	Pedv	3	205	380	2014.9	20.01
138	Peptide	2	213	404	2009.9	29.06
139	Person	1	210	308	2011.7	35.80
140	Phylogenetic analysis	3	213	487	2012.5	34.15
141	Pig	3	210	333	2012.9	23.48
142	Piglet	3	198	305	2014.6	15.93
143	Pneumonia	1	211	464	2011.1	46.78
144	Population	1	215	739	2012.2	32.82
145	Porcine epidemic diarrhea virus	3	203	334	2014.8	20.63
146	Presence	1	215	925	2010.9	31.61
147	Present study	3	215	400	2011.6	18.39
148	Prevalence	1	211	530	2012.8	25.51
149	Prevention	1	214	336	2012.7	22.25
150	Process	2	215	583	2011.2	32.60
151	Production	2	215	599	2011.0	23.95
152	Protease	2	210	492	2011.3	31.66
153	Protection	3	214	397	2011.5	25.94
154	Protein	2	215	2,653	2010.2	32.26
155	Receptor	2	214	799	2010.9	35.81
156	Region	3	215	1,329	2010.8	27.03
157	Replication	2	215	955	2011.1	35.05
158	Report	1	215	539	2011.1	31.17
159	Research	1	215	401	2012.5	23.15
160	Residue	2	213	557	2009.9	30.78
161	Respiratory syncytial virus	1	202	511	2012.6	33.47
162	Respiratory virus	1	210	602	2012.8	31.24
163	Response	2	215	1,137	2011.3	30.89
164	Review	1	214	641	2012.8	32.34
165	Rhinovirus	1	200	445	2012.5	37.18
166	Risk	1	214	527	2012.8	21.89
167	Rna	2	215	848	2010.1	33.27
168	Rna virus	2	215	375	2011.3	38.95
169	Role	2	215	1,575	2011.2	33.08
170	Rsv	1	185	227	2012.9	29.26
171	Rt pcr	1	214	546	2010.1	33.35
172	S protein	2	212	421	2010.1	31.35
173	Sample	1	215	1,672	2011.9	25.97
174	Sar	6	214	783	2007.4	40.60
175	Sars	6	215	1,575	2007.4	42.70
176	Sars coronavirus	2	213	683	2007.8	41.28
177	Sars cov	2	215	1,676	2008.7	37.27
178	Sars cov infection	2	210	298	2008.0	35.79
179	Sars patient	6	198	324	2005.8	27.59
180	Saudi arabia	1	207	300	2016.2	30.43
181	Sensitivity	4	214	449	2010.9	24.58
182	Sequence	3	215	1,342	2010.3	36.93
183	Sera	4	214	322	2010.0	23.50
184	Severe acute respiratory syndrome	6	215	1657	2007.2	45.17
185	Severe acute respiratory syndrome coronavirus	2	215	848	2009.7	37.27
186	Site	2	215	869	2010.2	32.60
187	Species	5	215	719	2012.4	37.87
188	Specificity	4	215	477	2010.7	26.24
189	Spike	2	212	683	2010.6	31.89
190	Spike protein	2	213	593	2010.6	33.81
191	Spread	1	214	400	2011.8	26.37
192	Strain	3	215	1,562	2011.3	23.96
193	Structure	2	215	902	2010.7	32.96
194	Study	1	215	4,070	2011.9	25.54
195	Symptom	1	215	571	2011.4	37.32
196	T cell	2	211	310	2009.6	27.75
197	Tgev	3	206	303	2011.3	21.11
198	Time	1	215	993	2011.4	28.94
199	Total	1	213	434	2012.7	22.01
200	Transmissible gastroenteritis virus	3	202	271	2011.5	22.80
201	Transmission	1	215	811	2012.8	32.03
202	Treatment	2	215	969	2011.1	26.43
203	Type	2	215	1,201	2011.5	30.94
204	Understanding	5	215	525	2012.9	27.34
205	Use	1	215	727	2010.7	33.34
206	Vaccination	3	211	330	2012.0	21.73
207	Vaccine	3	215	1,064	2011.7	26.96
208	Viral infection	1	215	684	2012.1	27.35
209	Viral replication	2	211	433	2011.4	34.50
210	Viral rna	2	215	275	2010.4	43.71
211	Virus	3	215	4,810	2011.5	32.43
212	Virus replication	2	210	283	2011.2	30.79
213	Vitro	2	215	557	2011.2	25.64
214	Vivo	2	210	227	2011.4	28.10
215	Week	1	215	320	2010.2	29.21
216	Year	1	215	961	2012.2	33.96