Collaboration in the Time of COVID: A Scientometric Analysis of Multidisciplinary SARS-CoV-2 Research

ABSTRACT

The novel coronavirus SARS-CoV-2 and the COVID-19 illness it causes have inspired unprecedented levels of multidisciplinary research in an effort to address a generational public health challenge. In this work we conduct a scientometric analysis of COVID-19 research, paying particular attention to the nature of collaboration that this pandemic has fostered among different disciplines. Increased multidisciplinary collaboration has been shown to produce greater scientific impact, albeit with higher co-ordination costs. As such, we consider a collection of over 166,000 COVID-19-related articles to assess the scale and diversity of collaboration in COVID-19 research, which we compare to non-COVID-19 controls before and during the pandemic. We show that COVID-19 research teams are not only significantly smaller than their non-COVID-19 counterparts, but they are also more diverse. Furthermore, we find that COVID-19 research has increased the multidisciplinarity of authors across most scientific fields of study, indicating that COVID-19 has helped to remove some of the barriers that usually exist between disparate disciplines. Finally, we highlight a number of interesting areas of multidisciplinary research during COVID-19, and propose methodologies for visualising the nature of multidisciplinary collaboration, which may have application beyond this pandemic.

Supplemental Materials

	Articles	Articles (explicit)	Authors per paper
medicine	149,111	93,350	6.2
biology	131,408	12,454	6.2
chemistry	54,808	5,357	7.1
economics	50,389	1,357	5.2
psychology	49,871	10,184	5.6
computer science	47,275	9,004	5.4
political science	45,146	8,739	5.0
engineering	38,521	295	5.6
sociology	32,721	2,556	5.2
mathematics	27,679	1,321	5.7
physics	26,898	576	6.1
philosophy	22,442	403	4.6
geography	14,922	7,071	5.1
business	14,718	6,983	4.2
history	11,972	2,225	4.7
art	11,519	271	4.8
materials science	8,123	1,229	6.0
geology	6,404	95	5.4
environmental science	5,513	1,245	5.1

Table 1. Summary of *2020-COVID-related research* analysed in this work, in terms of the number of research articles and mean number of authors per paper for each discipline. In the case of each discipline, we include any articles that are assigned to fields of study that are within the discipline, i.e., the discipline is a parent of any of the smaller fields of study to which the paper is assigned. We also report the number of articles that are explicitly assigned to that discipline, i.e., the discipline is included in the original set of fields of study attributed to the paper.

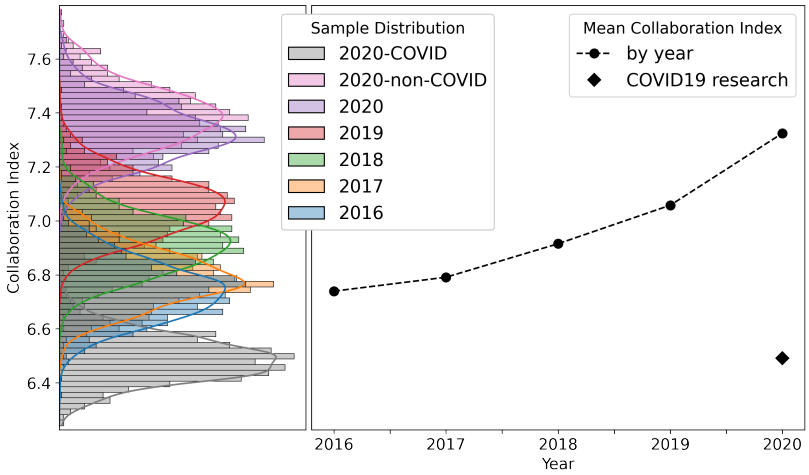
	Articles	Articles (explicit)	Authors per paper
medicine	915,474	581,147	8.1
biology	865,190	105,707	8.3
chemistry	551,393	105,925	8.8
engineering	358,897	2,774	7.3
computer science	328,293	111,965	6.9
psychology	320,474	74,131	7.0
physics	284,391	21,023	8.6
mathematics	276,038	16,153	7.0
political science	229,465	18,415	6.4
economics	226,455	9,727	6.2
materials science	173,926	62,646	7.8
sociology	170,212	12,583	6.5
philosophy	114,323	1,041	6.1
geography	99,992	15,512	7.1
geology	76,014	5,154	7.8
art	74,140	1,328	6.8
business	72,259	27,964	5.1
history	59,372	2,492	6.4
environmental science	48,295	22,300	6.8

Table 2. Summary of *2020-non-COVID research* analysed in this work, in terms of the number of research articles and mean number of authors per paper for each discipline.

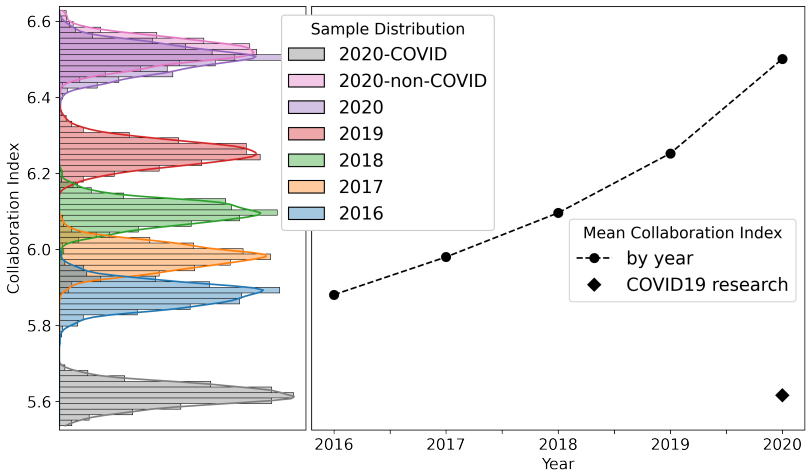
	Articles	Articles (explicit)	Authors per paper
medicine	2,961,762	1,838,989	7.3
biology	2,804,222	434,859	7.5
chemistry	1,833,632	323,882	8.1
engineering	1,203,777	45,147	6.7
psychology	1,035,184	241,721	6.2
computer science	1,033,976	336,604	6.3
physics	960,508	81,679	8.2
mathematics	862,481	68,010	6.3
political science	750,923	72,185	5.6
economics	745,753	43,975	5.4
materials science	590,016	208,028	7.3
sociology	548,850	46,088	5.6
philosophy	419,432	11,681	5.4
geography	332,303	50,735	6.5
art	280,670	12,696	5.8
geology	249,208	19,768	7.3
business	244,328	82,362	4.5
history	200,924	8,940	5.8
environmental science	157,073	64,045	6.1

Table 3. Summary of *pre-2020 research* analysed in this work, in terms of the number of research articles and mean number of authors per paper for each discipline.

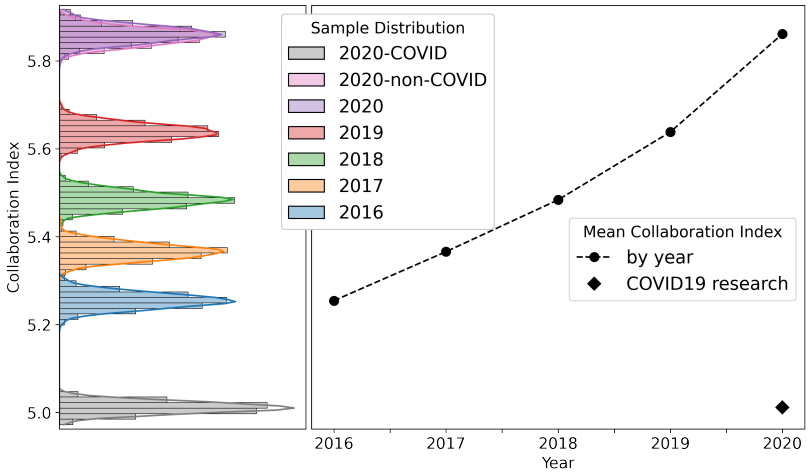
Figure 1. Sample distributions for Collaboration Index scores for different years and portions of the dataset. The experiment is repeated for different sample sizes. The trend towards greater collaboration is evident regardless of sample size and the COVID-19 research is consistently shown to be significantly less collaborative than other research in 2020.



(a) $n = 10,000$



(b) $n = 50,000$



(c) $n = 100,000$

	2017	2018	2019	2020	2020-non-COVID
art	5.01e-06*	2.28e-02*	7.32e-04*	3.27e-60*	1.17e-25*
medicine	6.76e-64*	4.00e-100*	1.48e-205*	0.00e+00*	1.18e-62*
geology	2.21e-02*	1.90e-26*	1.15e-17*	1.82e-222*	7.89e-11*
geography	1.30e-03*	0.73	4.31e-04*	3.29e-85*	7.02e-10*
economics	7.21e-64*	2.39e-211*	4.25e-116*	0.00e+00*	1.14e-18*
computer science	1.41e-26*	1.00	1.01e-07*	0.00e+00*	0.00e+00*
philosophy	0.14	1.97e-02*	1.44e-09*	1.02e-23*	1.01e-10*
biology	1.91e-74*	1.00	6.16e-256*	0.00e+00*	1.45e-271*
business	2.35e-12*	2.24e-33*	8.59e-10*	3.14e-110*	5.21e-23*
political science	0.91	0.13	1.71e-05*	1.00	0.57
history	1.79e-09*	4.49e-03*	4.93e-05*	1.38e-39*	1.62e-27*
psychology	4.82e-06*	1.00	9.23e-18*	0.00e+00*	1.22e-11*
sociology	1.20e-11*	0.95	8.14e-07*	4.59e-06*	0.13
mathematics	1.97e-20*	1.00	3.88e-70*	0.00e+00*	1.75e-172*
engineering	0.64	0.82	3.64e-13*	0.00e+00*	7.82e-155*
physics	1.28e-29*	1.00	3.07e-131*	0.00e+00*	3.14e-139*
materials science	2.89e-02*	1.43e-03*	2.38e-09*	1.53e-212*	3.86e-23*
chemistry	1.42e-35*	0.93	3.78e-81*	0.00e+00*	8.24e-88*
environmental science	1.44e-06*	7.17e-03*	3.92e-06*	1.02e-36*	1.21e-02*
all disciplines	2.62e-199*	1.42e-53*	0.00e+00*	0.00e+00*	4.40e-146*

Table 4. P values for Fisher's Exact test associated with the odds ratios presented in Figure 2. Proportions of external edges in the network are compared using a one-tailed test with the alternative hypothesis that the proportion of external edges in the given year is greater than that of the previous year.

Team size	t	p	mean ± std 2020-COVID	mean ± std pre-2020
2	6.73	8.34e-12*	0.134 ± 0.15	0.128 ± 0.14
3	10.41	1.16e-25*	0.136 ± 0.12	0.128 ± 0.11
4	14.14	1.20e-45*	0.135 ± 0.11	0.124 ± 0.10
5	19.07	2.96e-81*	0.134 ± 0.10	0.118 ± 0.09
6	21.16	1.64e-99*	0.132 ± 0.09	0.113 ± 0.08
7	23.06	1.30e-117*	0.130 ± 0.09	0.108 ± 0.07
8	20.71	2.71e-95*	0.126 ± 0.08	0.104 ± 0.07
9	20.74	2.08e-95*	0.124 ± 0.08	0.100 ± 0.07
10	20.99	1.51e-97*	0.125 ± 0.08	0.098 ± 0.06
11	18.82	7.85e-79*	0.123 ± 0.08	0.096 ± 0.06
12	16.95	2.38e-64*	0.121 ± 0.07	0.094 ± 0.06
13	15.71	1.61e-55*	0.120 ± 0.07	0.093 ± 0.06
14	15.67	3.36e-55*	0.121 ± 0.07	0.091 ± 0.06
15	14.27	4.12e-46*	0.120 ± 0.07	0.089 ± 0.06
16	14.95	3.16e-50*	0.121 ± 0.07	0.087 ± 0.06
17	11.65	2.04e-31*	0.116 ± 0.07	0.088 ± 0.06
18	13.04	1.21e-38*	0.122 ± 0.07	0.086 ± 0.06
19	12.40	3.93e-35*	0.119 ± 0.07	0.083 ± 0.05

(a) Alternative hypothesis: 2020-COVID research teams are more diverse than pre-2020.

Team size	t	p	mean ± std 2020-COVID	mean ± std 2020-non-COVID
2	4.11	1.98e-05*	0.134 ± 0.15	0.131 ± 0.14
3	7.48	3.76e-14*	0.136 ± 0.12	0.130 ± 0.11
4	10.86	8.93e-28*	0.135 ± 0.11	0.126 ± 0.10
5	15.77	2.87e-56*	0.134 ± 0.10	0.121 ± 0.09
6	18.39	9.83e-76*	0.132 ± 0.09	0.115 ± 0.08
7	19.88	4.93e-88*	0.130 ± 0.09	0.110 ± 0.07
8	19.15	9.35e-82*	0.126 ± 0.08	0.106 ± 0.07
9	18.13	1.65e-73*	0.124 ± 0.08	0.102 ± 0.07
10	20.15	4.70e-90*	0.125 ± 0.08	0.099 ± 0.06
11	17.81	8.98e-71*	0.123 ± 0.08	0.097 ± 0.06
12	16.31	1.18e-59*	0.121 ± 0.07	0.094 ± 0.06
13	15.74	1.16e-55*	0.120 ± 0.07	0.092 ± 0.06
14	14.93	3.29e-50*	0.121 ± 0.07	0.092 ± 0.06
15	13.55	9.72e-42*	0.120 ± 0.07	0.091 ± 0.06
16	12.53	6.14e-36*	0.121 ± 0.07	0.090 ± 0.06
17	10.58	3.33e-26*	0.116 ± 0.07	0.089 ± 0.06
18	11.16	7.31e-29*	0.122 ± 0.07	0.088 ± 0.06
19	10.30	7.60e-25*	0.119 ± 0.07	0.087 ± 0.06

(b) Alternative hypothesis: 2020-COVID research teams are more diverse than 2020-non-COVID.

Team size	t	p	mean ± std 2020-non-COVID	mean ± std pre-2020
2	4.82	7.30e-07*	0.131 ± 0.14	0.128 ± 0.14
3	5.72	5.22e-09*	0.130 ± 0.11	0.128 ± 0.11
4	6.73	8.24e-12*	0.126 ± 0.10	0.124 ± 0.10
5	6.79	5.64e-12*	0.121 ± 0.09	0.118 ± 0.09
6	5.54	1.50e-08*	0.115 ± 0.08	0.113 ± 0.08
7	6.19	3.00e-10*	0.110 ± 0.07	0.108 ± 0.07
8	2.91	1.80e-03*	0.106 ± 0.07	0.104 ± 0.07
9	4.59	2.27e-06*	0.102 ± 0.07	0.100 ± 0.07
10	1.36	0.09	0.099 ± 0.06	0.098 ± 0.06
11	1.66	0.05	0.097 ± 0.06	0.096 ± 0.06
12	0.85	0.20	0.094 ± 0.06	0.094 ± 0.06
13	-0.36	0.64	0.092 ± 0.06	0.093 ± 0.06
14	0.78	0.22	0.092 ± 0.06	0.091 ± 0.06
15	1.58	0.06	0.091 ± 0.06	0.089 ± 0.06
16	3.52	2.19e-04*	0.090 ± 0.06	0.087 ± 0.06
17	1.15	0.12	0.089 ± 0.06	0.088 ± 0.06
18	2.45	7.09e-03*	0.088 ± 0.06	0.086 ± 0.06
19	3.26	5.67e-04*	0.087 ± 0.06	0.083 ± 0.05

(c) Alternative hypothesis: 2020-non-COVID research teams are more diverse than pre-2020.

Table 5. Independent *t* tests comparing research team diversities. We compare diversity scores for teams of equal sizes with the null hypothesis that the mean diversity in each distribution is the same. Each test is one-tailed with the alternative hypothesis stated in the sub-caption. Values significant at $p < 0.01$ are marked with a *. COVID research teams of all sizes are shown to be significantly more diverse than non-COVID baselines (see tables 4(a) and 4(b)).

	Title	Citations
1	A Novel Coronavirus from Patients with Pneumonia in China, 2019.	6700
2	A pneumonia outbreak associated with a new coronavirus of probable bat origin	5562
3	Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus-Infected Pneumonia.	5002
4	SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor.	4345
5	A new coronavirus associated with human respiratory disease in China.	2783
6	Aerosol and Surface Stability of SARS-CoV-2 as Compared with SARS-CoV-1.	2499
7	Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation.	2321
8	Remdesivir and chloroquine effectively inhibit the recently emerged novel coronavirus (2019-nCoV) in vitro.	2239

(a) Most cited papers

	Title unigrams	Title bigrams	Abstract	Field of Study
1	covid-19 (11129)	covid-19 pandemic (1543)	virus (4822)	virology (22561)
2	sars-cov-2 (5321)	coronavirus disease (1275)	2 (4595)	coronavirus disease 2019 (18238)
3	coronavirus (3888)	disease 2019 (1058)	sars (4168)	severe acute respiratory syndrome coronavirus 2 (17579)
4	pandemic (2680)	sars-cov-2 infection (748)	cov (4135)	2019 20 coronavirus outbreak (16990)
5	disease (2058)	novel coronavirus (647)	coronavirus (3875)	medicine (15247)
6	infection (2057)	covid-19 patient (469)	19 (3873)	pandemic (8790)
7	patient (1893)	2019 covid-19 (410)	covid (3801)	betacoronavirus (5784)
8	virus (1521)	acute respiratory (384)	infection (3670)	coronavirus (5159)
9	2019 (1376)	respiratory syndrome (365)	disease (3564)	biology (4582)
10	de (1042)	severe acute (354)	respiratory (2746)	coronavirus infections (3210)

(b) Most frequent terms

Table 6. Virology papers. A description of 22,561 Virology research articles. Table (a) lists the titles of the most cited papers in the subset. Table (b) lists terms that appear most frequently in the titles and abstracts of the papers, along with the most common MAG FoS. The number of articles with the respective terms/FoS are included in brackets.

	Title	Citations
1	Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR.	1899
2	Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing.	680
3	COVID-Net: a tailored deep convolutional neural network design for detection of COVID-19 cases from chest X-ray images.	320
4	Diagnosing COVID-19: The Disease and Tools for Detection	257
5	Covid-19: automatic detection from X-ray images utilizing transfer learning with convolutional neural networks.	228
6	A deep learning algorithm using CT images to screen for Corona Virus Disease (COVID-19)	227
7	Automatic Detection of Coronavirus Disease (COVID-19) Using X-ray Images and Deep Convolutional Neural Networks.	204
8	An updated estimation of the risk of transmission of the novel coronavirus (2019-nCov)	201

(a) Most cited papers

	Title unigrams	Title bigrams	Abstract	Field of Study
1	covid-19 (2792)	covid-19 pandemic (348)	model (4908)	computer science (9004)
2	using (1097)	deep learning (268)	19 (3734)	coronavirus disease 2019 (3453)
3	learning (1005)	machine learning (261)	covid (3729)	artificial intelligence (2081)
4	model (959)	neural network (260)	data (3632)	severe acute respiratory syndrome coronavirus 2 (1287)
5	data (722)	contact tracing (179)	paper (2897)	2019 20 coronavirus outbreak (1036)
6	network (702)	chest x-ray (139)	approach (2863)	pandemic (1025)
7	system (569)	x-ray image (127)	system (2830)	machine learning (835)
8	analysis (566)	case study (120)	time (2757)	deep learning (711)
9	based (537)	artificial intelligence (109)	network (2251)	data science (620)
10	detection (533)	model covid-19 (101)	pandemic (2247)	pattern recognition (505)

(b) Most frequent terms

Table 7. Computer Science papers. A description of 9,004 Computer Science research articles. Table (a) lists the titles of the most cited papers in the subset. Table (b) lists terms that appear most frequently in the titles and abstracts of the papers, along with the most common MAG FoS. The number of articles with the respective terms/FoS are included in brackets.

	Title	Citations
1	The airborne lifetime of small speech droplets and their potential importance in SARS-CoV-2 transmission.	204
2	Electrochemical biosensors for pathogen detection	41
3	Household Materials Selection for Homemade Cloth Face Coverings and Their Filtration Efficiency Enhancement with Triboelectric Charging.	28
4	The versatile biomedical applications of bismuth-based nanoparticles and composites: therapeutic, diagnostic, biosensing, and regenerative properties	26
5	DeepCOVIDExplainer: Explainable COVID-19 Predictions Based on Chest X-ray Images	24
6	Green synthesis, biomedical and biotechnological applications of carbon and graphene quantum dots. A review	22
7	Modeling ambient temperature and relative humidity sensitivity of respiratory droplets and their role in Covid-19 outbreaks	21
8	An overview of functional nanoparticles as novel emerging antiviral therapeutic agents	20

(a) Most cited papers

	Title unigrams	Title bigrams	Abstract	Field of Study
1	effect (88)	filtration efficiency (14)	material (404)	materials science (1229)
2	using (86)	additive manufacturing (13)	high (403)	nanotechnology (152)
3	application (86)	mechanical property (10)	surface (402)	chemical engineering (149)
4	covid-19 (83)	solar cell (9)	property (321)	coronavirus disease 2019 (143)
5	property (70)	face mask (9)	application (315)	composite material (135)
6	surface (68)	respiratory droplet (9)	effect (312)	optoelectronics (123)
7	material (62)	recent advance (8)	2 (281)	severe acute respiratory syndrome coronavirus 2 (93)
8	based (52)	covid-19 pandemic (8)	1 (275)	nanoparticle (74)
9	nanoparticles (47)	3d printed (8)	time (269)	biomedical engineering (73)
10	study (44)	quantum dot (8)	different (265)	2019 20 coronavirus outbreak (65)

(b) Most frequent terms

Table 8. Materials Science papers. A description of 1,229 Materials Science research articles. Table (a) lists the titles of the most cited papers in the subset. Table (b) lists terms that appear most frequently in the titles and abstracts of the papers, along with the most common MAG FoS. The number of articles with the respective terms/FoS are included in brackets.

	Title	Citations
1	The socio-economic implications of the coronavirus pandemic (COVID-19): A review.	505
2	COVID-19: towards controlling of a pandemic.	379
3	Return of the Coronavirus: 2019-nCoV.	337
4	Effective containment explains subexponential growth in recent confirmed COVID-19 cases in China.	224
5	The effect of large-scale anti-contagion policies on the COVID-19 pandemic.	185
6	Countries test tactics in 'war' against COVID-19.	134
7	The impact of COVID-19 and strategies for mitigation and suppression in low- and middle-income countries.	116
8	If the world fails to protect the economy, COVID-19 will damage health not just now but also in the future.	101

(a) Most cited papers

	Title unigrams	Title bigrams	Abstract	Field of Study
1	covid-19 (1052)	covid-19 pandemic (247)	covid (1158)	development economics (1564)
2	pandemic (450)	impact covid-19 (63)	19 (1152)	coronavirus disease 2019 (1072)
3	health (172)	response covid-19 (41)	pandemic (1031)	pandemic (997)
4	crisis (157)	covid-19 crisis (41)	country (816)	2019 20 coronavirus outbreak (573)
5	global (149)	social distancing (35)	health (617)	severe acute respiratory syndrome coronavirus 2 (531)
6	impact (148)	covid-19 outbreak (34)	disease (535)	business (507)
7	response (123)	spread covid-19 (29)	economic (533)	political science (447)
8	economic (119)	pandemic covid-19 (29)	impact (511)	geography (282)
9	social (111)	public health (27)	social (480)	public health (248)
10	policy (103)	covid-19 epidemic (25)	policy (456)	outbreak (210)

(b) Most frequent terms

Table 9. Development Economics papers. A description of 1,564 Development Economics research articles. Table (a) lists the titles of the most cited papers in the subset. Table (b) lists terms that appear most frequently in the titles and abstracts of the papers, along with the most common MAG FoS. The number of articles with the respective terms/FoS are included in brackets.