

COVID-19 in the News, 2019-2026: The Rise and Fall of Pandemic Coverage, Its Geography and Tone, and How It Compares with Research and Invention

Fliar BI · Prism Research Layer · Working paper, October 2026 · Dataset and code: github.com/fliarbi/prism-news-api · Licence CC BY-NC 4.0

Abstract

We measure how COVID-19 was covered in a global news index from January 2019 to July 2026, using 4,670,267 de-duplicated COVID-19 articles, and set that coverage beside 782,075 research papers (2020-2025) and 12,953 patent families (2020-2024) on the disease and its virus. Six results stand out:

1. **Peak.** In April 2020 COVID-19 took 3,150.18 of every 10,000 news articles, or 31.5% of all news that month. 49.91% of all COVID-19 articles in the window appeared in 2020.
2. **Decline.** The topic's share of news fell from 1,436.12 per 10,000 in 2020 to 801.31 in 2021, 278.34 in 2022, 83.39 in 2023 and 38.05 in 2024: a 37.74-fold fall in four years.
3. **Tone and urgency.** In January 2023 – July 2026, 53.7% of COVID-19 articles were negative in tone (all news: 33.6%), and 74.5% were rated critical priority (all news: 42.6%).
4. **Geography.** The United States holds 53.85% of located COVID-19 news (lift 1.11). China is the most over-represented large country (2.43); India is under-represented (0.75). Within India, Kerala (2.97), Delhi (1.67) and Goa (1.61) lead. US state lifts are dominated by an archive back-fill and are not interpreted.
5. **Shifting attention.** As a share of each country's own news, COVID-19 coverage fell between 2023 and 2025 in China (370.64 to 78.73 per 10,000) and India (93.65 to 25.85).
6. **Research and invention.** Research papers peaked a year after news, in 2021 (215,544 papers, 3.44% of all papers). Patent families peaked in 2022 (3,418), 2.9 times the 2020 count.

Findings stated for citation

- **F1.** April 2020 was the peak month: 3,150.18 COVID-19 articles per 10,000 news articles (31.5%).
- **F2.** COVID-19's share of news fell 37.74-fold between 2020 (1,436.12 per 10,000) and 2024 (38.05).
- **F3.** 49.91% of the 4,670,267 COVID-19 articles from January 2019 to July 2026 appeared in 2020.
- **F4.** 53.7% of COVID-19 articles were negative in tone, 1.6 times the all-news share.
- **F5.** 74.5% of COVID-19 articles carried a machine-assigned critical priority rating, 1.75 times the all-news share.

- **F6.** The United States held 53.85% of located COVID-19 news (lift 1.11); China's lift was 2.43 and India's 0.75.
- **F7.** Vaccine development and efficacy trials were the most frequent theme (15.78% of themed articles).
- **F8.** "Covid" appeared in 1.66 times as many article openings as "coronavirus".
- **F9.** Research papers peaked in 2021 at 3.44% of all papers; their share fell 5.41-fold by 2025.
- **F10.** Patent families rose 2.9-fold from 2020 to 2022; China's office held 37.5% of office filings.
- **F11.** Within India, Kerala (2.97), Delhi (1.67), Goa (1.61) and Maharashtra (1.52) had the most over-represented coverage; Bihar (0.27) and Jammu and Kashmir (0.32) the least.
- **F12.** COVID-19's share of China's news fell from 370.64 per 10,000 in 2023, the highest of 25 tracked countries, to 85.15 in 2024 and 78.73 in 2025.
- **F13.** Relative to its own news, coverage was most negative in China (65.6% against 27.2%) and Canada (60.7% against 27.1%).
- **F14.** By continent, COVID-19 news was over-represented in Oceania (1.33) and under-represented in Africa (0.62).

1. Introduction

COVID-19 is the disease caused by the SARS-CoV-2 coronavirus. Its spread from early 2020 produced the largest single news story of the period covered here. News was the main public record of the pandemic as it happened: case counts, lockdowns, vaccines and their effects on work, schools and health systems. Coverage also shapes what the public treats as urgent, and when it stops treating it so.

Most measures of the pandemic count cases, deaths or vaccinations. Few measure attention: how much of the news the pandemic took, when that share rose and fell, where it was concentrated and in what tone. This paper and its dataset do that. They combine three records for one topic: a global news index, research papers and patents. Each comparison uses a denominator, so that growth in the index or in publishing is not mistaken for growth in interest. The companion dataset publishes an article-level sample (97,067 rows) and full coverage tables, so every figure here can be re-derived. The [Prism News API](#) serves the full history.

2. Data and methods

News. The Prism news index holds articles from outlets worldwide; non-English articles are indexed in English translation. Each article carries its opening text (the first few sentences), a primary location, a graded tone label, a machine-assigned priority rating (1–9), themes, keywords and named organisations. We count unique article identifiers; repeated rows for the same article are removed. When this release was built, 92.88% of the index was readable, so all counts are lower bounds.

Matching. An article is about COVID-19 when its opening text contains *covid*, *covid19*, *coronavirus*, *coronaviruses*, *SARS-CoV-2*, *2019-nCoV*, *Omicron*, *corona virus*, *corona pandemic* or *corona crisis* as whole words. The word *pandemic* alone was tested and dropped: it is used for other diseases and figuratively, and often only as background in business news. Three families of matches are excluded:

- research abstracts filed as news items, recognised by an opening of the form "title.Background:" or "title.Objective:";
- market-size reports and laboratory assay listings;
- job advertisements ("recruitment 2026").

Together these removed 138,293 of 4,808,560 raw matches (2.88%). One reviewer read 40 randomly drawn matched openings. 38 referred to the COVID-19 pandemic, the disease or the virus (precision 95%), and in 28 (70%) COVID-19 was the main subject. The two false positives were a newsletter sign-up line above an unrelated story and an abstract on another respiratory virus. Some research abstracts with other opening formats remain; they concern COVID-19 and are counted.

Measures. Coverage over time is expressed per 10,000 articles in the index that month or year, because the index grew from 11,807,177 articles in 2019 to 26,990,006 in 2025. Geography, tone, priority and themes use January 2023 – July 2026, the window the index's location and label fields cover consistently. That window is the late phase of the pandemic, not its peak. Geography is expressed as lift: a country's share of COVID-19 articles divided by its share of all located articles. Place tags derived from ordinary words ("Congress", "Earth", "Winter", "Cloud", "Mobile", "Hurricane" and 714 such words in all) were treated as unlocated in every location field, which removed 15.3% of city-tagged articles. Year comparisons use 2020–2024. 2019 is not a usable base (the outbreak was reported only at the end of that year), 2025 is affected by an archive back-fill described in §3.1, and 2026 is partial.

Research and patents. Research papers are de-duplicated articles and reviews with *COVID*, *COVID19*, *coronavirus disease 2019*, *SARS-CoV-2*, *2019-nCoV* or *2019 novel coronavirus* in the title. Repository deposits and animal coronaviruses (porcine, feline, bovine, avian and others) and MERS are excluded. Patent families (2005–2024) use a similar title rule with *coronavirus* and *novel coronavirus*, are counted once per filing office, and exclude the same animal and MERS senses.

3. Results

3.1 How much and when

The index holds 4,670,267 COVID-19 articles from January 2019 to July 2026. Almost half of them (49.91%, or 2,330,831) appeared in 2020. The topic's share of all news was 1,436.12 per 10,000 in 2020, that is 14.36% of everything the index collected that year. It then fell every year to 2024 (Table 1). By 2023 the share was 5.81% of its 2020 level.

Year	COVID-19 articles	All articles in index	Per 10,000
2019	1,941	11,807,177	1.64
2020	2,330,831	16,230,014	1,436.12
2021	1,384,755	17,281,220	801.31
2022	532,067	19,115,515	278.34
2023	170,814	20,484,506	83.39
2024	88,694	23,311,896	38.05
2025	138,010	26,990,006	51.13

Table 1. COVID-19 articles per year and as a share of the index. 2026 (January–July, 17.64 per 10,000) is partial and omitted. The 2025 row includes 2020–2021 archive stories added with August 2025 dates.

The monthly series (Figure 1) rises from 160.5 per 10,000 in January 2020 to 459.0 in February, 2,891.16 in March and a peak of 3,150.18 in April 2020, when 412,557 articles matched. From May 2020 the share declines in steps, with later rises in December 2020 (1,169.15 per 10,000) and December 2021 (820.38). The second of these, with January 2022 (784.97), covers the months of the Omicron wave.

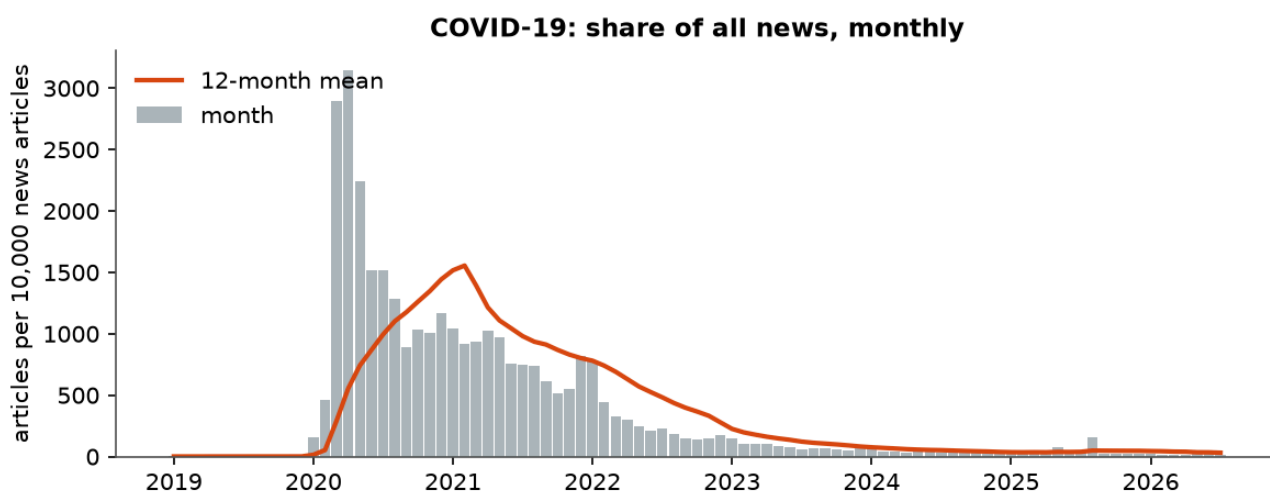


Figure 1. COVID-19 coverage as articles per 10,000 news articles, monthly, January 2019 – July 2026. Bars are months; the line is the 12-month mean.

One month breaks the pattern. August 2025 holds 47,574 matched articles (157.85 per 10,000), several times the months around it (50.83 in July and 24.8 in September). Reading its openings shows that most are stories from 2020 and 2021, such as reopening plans and early outbreak reports, which entered the index with 2025 dates. We treat this as an ingest artefact. It is the reason 2025 (51.13 per 10,000) appears higher than 2024 (38.05), and we make no claim of a 2025 rebound.

Averaged within-year month shares (Figure 2) do not show a stable season. January (12.3%), August (11.53%) and March (10.34%) are the largest shares, and October (5.68%) and November (5.62%) the smallest. These reflect individual waves and the August 2025 back-fill, not a yearly cycle.

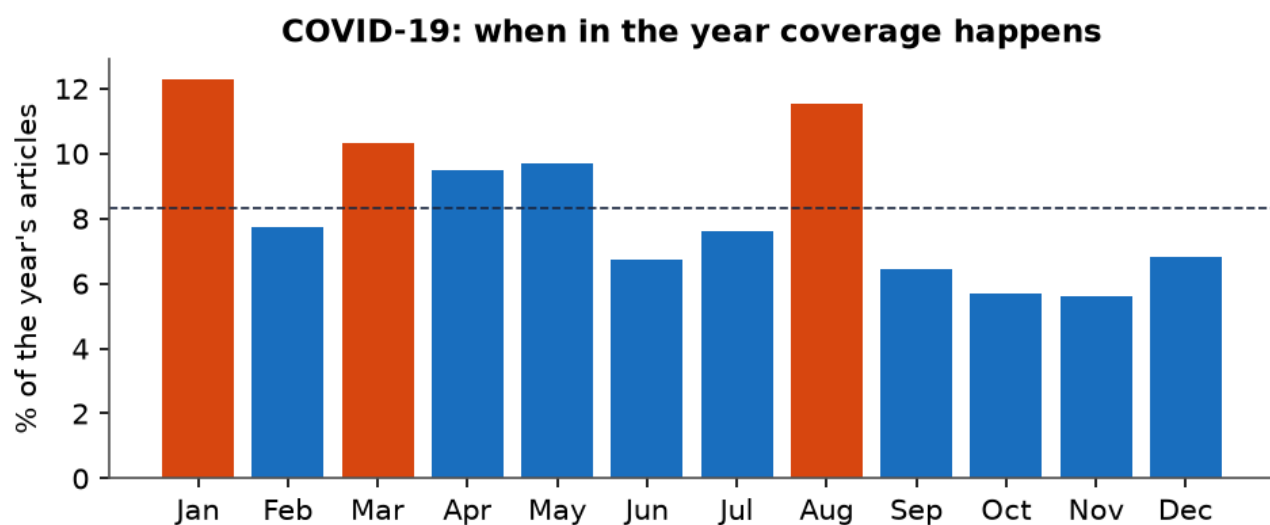


Figure 2. Share of each year's COVID-19 articles published in each month, averaged over 2019-2025. Dashed line = an even spread (8.3% per month).

3.2 Where

Between January 2023 and July 2026, 202,999 COVID-19 articles carried a primary location. The United States holds the largest raw share (53.85%) with a lift of 1.11. China's share (6.41%) comes with a lift of 2.43, the highest of any country in Table 2. India holds 9.87% of located articles but its lift is 0.75 (Table 2, Figure 3).

Country	Located COVID-19 articles	Share %	Lift vs all news
United States	109,320	53.85	1.11
India	20,032	9.87	0.75
China	13,007	6.41	2.43
United Kingdom	12,318	6.07	1.04
Australia	4,889	2.41	1.12
Canada	4,128	2.03	0.95
Brazil	2,979	1.47	1.7
Singapore	2,621	1.29	1.45
New Zealand	2,540	1.25	1.93
France	2,164	1.07	0.54
Japan	1,895	0.93	0.88
Germany	1,646	0.81	0.69
Nigeria	989	0.49	0.36
Pakistan	214	0.11	0.23

Table 2. Leading countries in located COVID-19 news, January 2023 – July 2026, with lift against all located news, and the least-covered country with at least 100 articles (Pakistan).

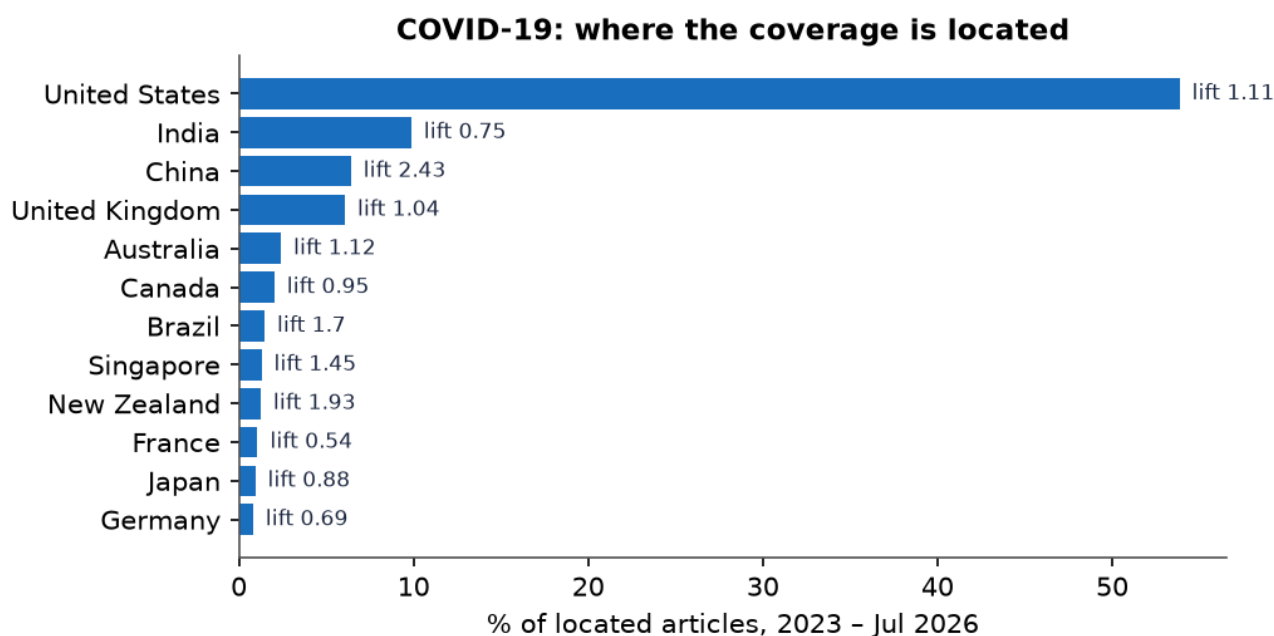


Figure 3. Leading countries by share of located COVID-19 articles, 2023 – Jul 2026, labelled with lift against all located news.

3.3 Geo intelligence: continents, states and shifting attention

The world map (Figure 4) shows the pattern at a glance. COVID-19 is over-represented in the news of China, New Zealand, Brazil and, more mildly, the United States and Australia. It is under-represented across most of Europe, the Middle East, South Asia and Africa. Among

countries with at least 100 located articles, China (2.43), New Zealand (1.93), Brazil (1.7), North Korea (1.57) and Singapore (1.45) have the highest lift, and Pakistan (0.23), Iran (0.32), Nigeria (0.36), Israel (0.39) and Turkey (0.4) the lowest. By continent, Oceania (1.33), North America (1.08) and South America (1.08) are over-represented, while Europe (0.94), Asia (0.9) and Africa (0.62) are not (Figure 7, left).

COVID-19: which countries over- and under-cover the topic

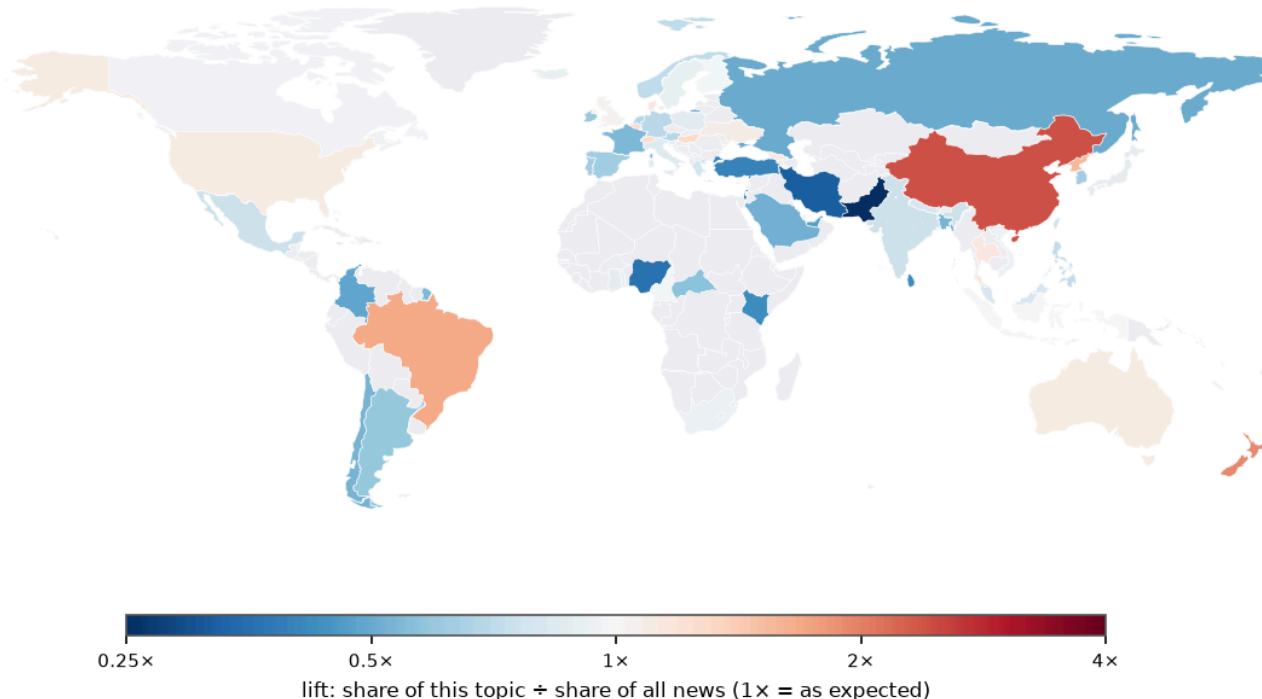


Figure 4. Lift of each country's share of located COVID-19 articles, January 2023 – July 2026 (red = over-represented, blue = under-represented, grey = fewer than 100 articles). India is drawn with its official outline.

Within countries the differences are larger (Table 3). In India, 13,975 COVID-19 articles carried a state. Measured against each state's share of Indian news, Kerala (2.97), Delhi (1.67), Goa (1.61), Maharashtra (1.52) and Karnataka (1.27) are the most over-represented. Bihar (0.27), Jammu and Kashmir (0.32), Jharkhand (0.37), Andhra Pradesh (0.45), Assam (0.48) and Uttarakhand (0.48) are under-represented, as is Uttar Pradesh (0.51) despite 902 articles (Figure 5). Few Indian articles fall in the August 2025 back-fill, so the Indian ranking describes late-phase coverage. In the United States, 85,318 articles carried a state. Kentucky (3.49), Vermont (2.69), Delaware (2.5) and Rhode Island (2.29) lead by lift. New Jersey (0.27) and Pennsylvania (0.49) are far below their weight in American news, partly because their local outlets are numerous in the index (Figure 6). The US ranking is not interpreted. The window includes the August 2025 back-fill of old stories, most of which are located in the United States, and most sampled rows for the leading states belong to it; Kentucky reached first place only once place tags derived from ordinary words were treated as unlocated. The back-fill also raises the US share of located news and the North American continent lift. States with fewer than 50 articles are not discussed.

India: state	Articles	Lift	United States: state	Articles	Lift
Kerala	1,010	2.97	Kentucky	1,032	3.49
Delhi	3,022	1.67	Vermont	242	2.69
Goa	124	1.61	Delaware	907	2.5
Maharashtra	3,382	1.52	Rhode Island	302	2.29
Karnataka	985	1.27	New Mexico	617	2.06
Tamil Nadu	758	1.08	Tennessee	3,526	1.82
<i>Lowest lift</i>					
Uttar Pradesh	902	0.51	Idaho	127	0.8
Assam	109	0.48	Wyoming	164	0.8
Uttarakhand	149	0.48	North Dakota	315	0.79
Andhra Pradesh	143	0.45	Pennsylvania	2,506	0.49
Jharkhand	126	0.37	New Jersey	4,332	0.27
Jammu and Kashmir	105	0.32			
Bihar	183	0.27			

Table 3. States with the highest and lowest lift in COVID-19 coverage, January 2023 – July 2026 (states with at least 100 articles). US lifts are dominated by the August 2025 back-fill. Lift = the state's share of the country's located COVID-19 articles ÷ its share of all located news from that country.

COVID-19: states of India by lift

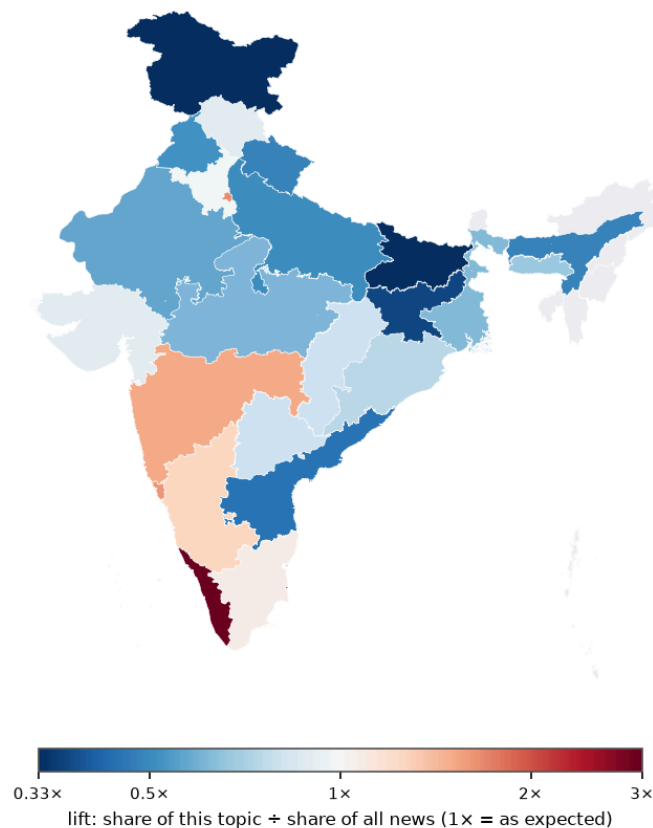


Figure 5. Lift of each Indian state's share of located COVID-19 articles against its share of all located Indian news, January 2023 – July 2026. Official outline; Ladakh is drawn within Jammu and Kashmir in the boundary file used.

COVID-19: states of United States by lift

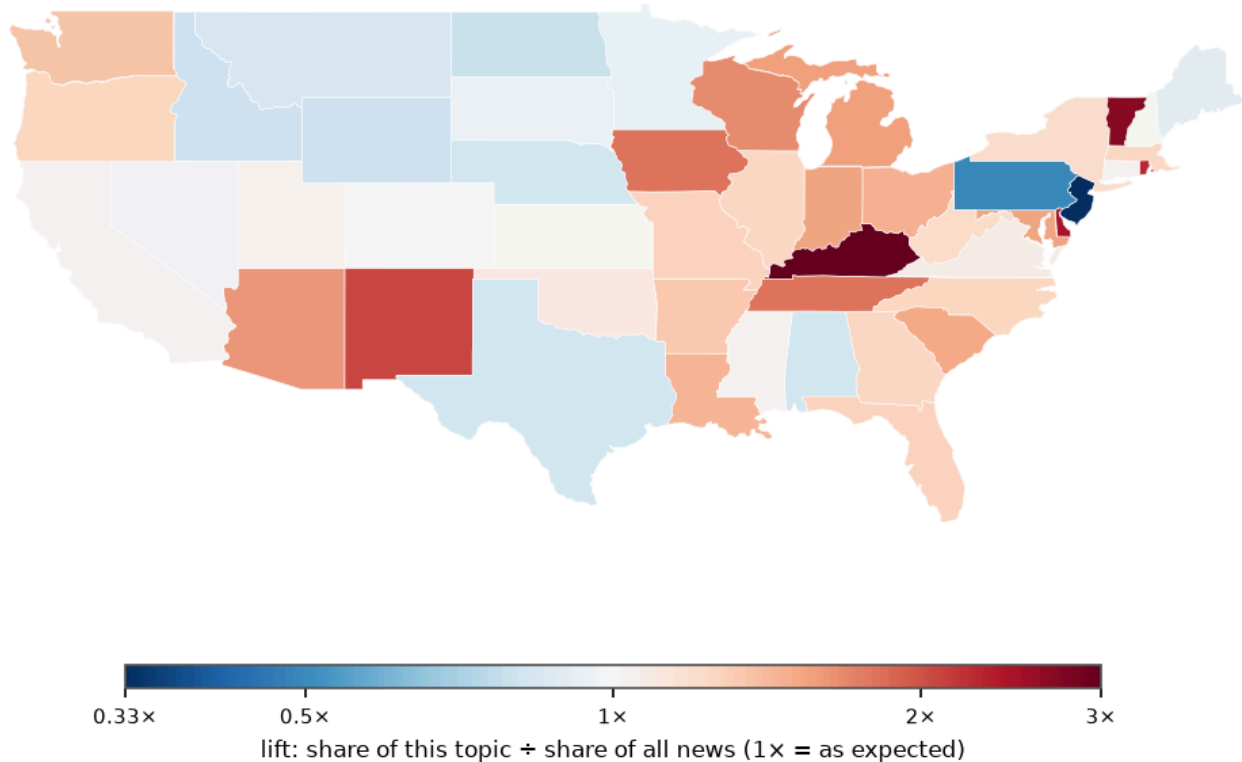


Figure 6. Lift of each US state's share of located COVID-19 articles against its share of all located US news, January 2023 – July 2026.

Attention also moved between countries (Figure 7, right). In 2020 COVID-19 took 3,412.59 of every 10,000 located Chinese articles, 2,068.99 Indian, 1,946.29 British and 1,259.35 American. Measured as COVID-19 articles per 10,000 of each country's located news, the steepest falls from 2023 to 2025 were in Ireland (80.28 to 10.76), Indonesia (123.0 to 22.27), China (370.64 to 78.73), Ukraine (106.85 to 26.87) and India (93.65 to 25.85). China's share had already fallen to 85.15 in 2024. The countries whose 2025 share rose (Brazil, Belgium, the United States, Canada) are not reported as rising: the US rise sits in the August 2025 back-fill, which is mostly located in the United States, and Brazil's sits almost entirely in the single month of May 2025.

Tone is negative in every large country, but its distance from each country's baseline varies. The gap is largest in China (65.6% negative against 27.2% for all Chinese news), Canada (60.7% against 27.1%) and the United States (64.2% against 33.1%). In Ukraine, where news overall is very negative (72.4%), COVID-19 coverage was less negative than the baseline (65.7%), as it was in Israel (52.4% against 73.2%) and Russia (61.2% against 72.1%). In France the two were close (51.6% against 47.6%).

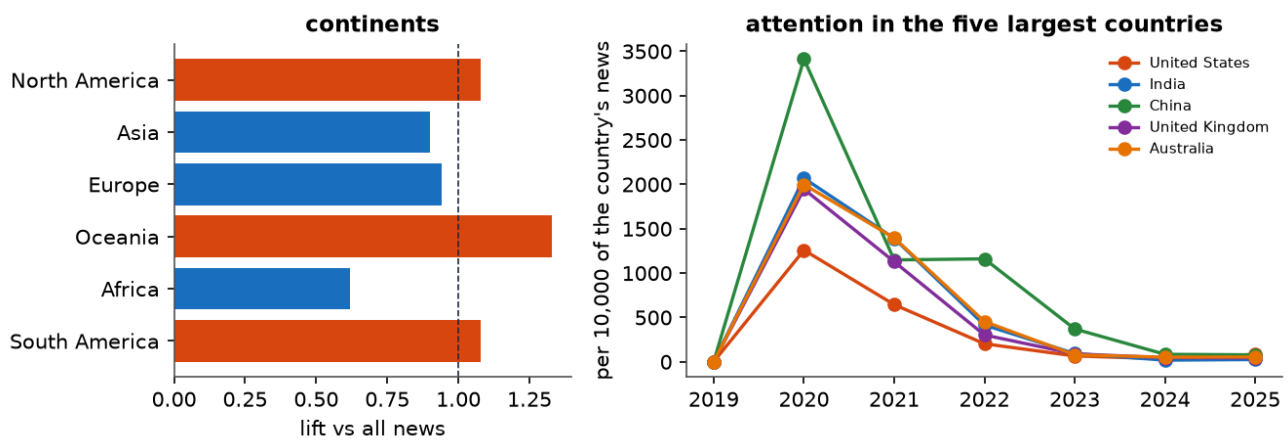


Figure 7. Left: lift by continent, January 2023 – July 2026. Right: COVID-19 articles per 10,000 of each country's located news, 2019–2025, for the five countries with the most articles.

3.4 Tone, urgency and themes

COVID-19 coverage is more negative and more urgent than news in general (Figure 8). Among articles with a tone label, 53.7% are negative, against 33.6% across the index (1.6 times), and 42.6% are positive, against 56.8%. The index's machine-assigned priority rating puts 74.5% of COVID-19 articles at critical level (7 or higher), against 42.6% of all news (1.75 times).

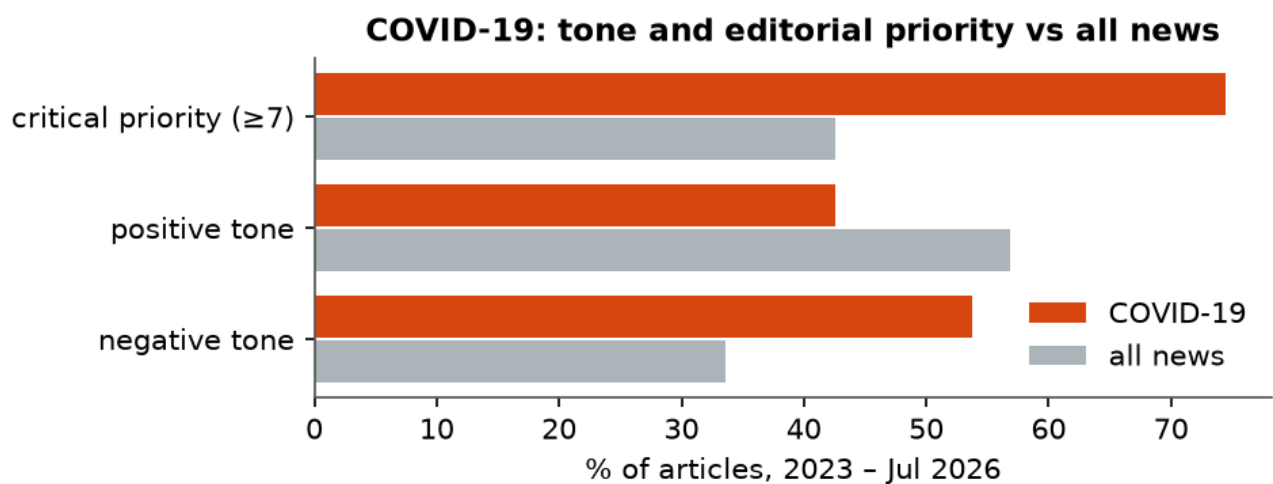


Figure 8. Tone and the share rated critical priority, COVID-19 articles vs all news, 2023 – Jul 2026.

Themes show what the late-phase coverage was about (Table 4). Vaccine development and efficacy trials are the most common (15.78% of themed articles), followed by public health policies and lockdown measures (10.3%) and vaccination campaigns (9.35%). Mental health effects (7.5%), variants (7.27%) and long-term effects (6.27%) each carry a substantial share. The themes most specific to the topic by lift are booster doses and long-term immunity (145.59) and pandemic vaccines (141.8).

Theme	Articles	Share %	Lift
covid-19 vaccine development and efficacy trials	40,353	15.78	95.22
covid-19 public health policies and lockdown measures	26,341	10.3	83.87
covid-19 vaccination campaigns and rollout strategies	23,916	9.35	91.67
covid-19 epidemiology and transmission modeling	22,399	8.76	98.94
covid-19 mental health impact and pandemic-related stress	19,178	7.5	80.33
covid-19 variants (alpha, delta, omicron) and mutations	18,575	7.27	105.88
covid-19 antiviral treatments and therapeutics	16,733	6.55	92.56
covid-19 long-term effects (long covid) and complications	16,027	6.27	91.98

Table 4. The most frequent themes in COVID-19 coverage, with lift against all themed news, January 2023 – July 2026.

The organisations named most often include the Centers for Disease Control and Prevention (16,527 articles), Pfizer (12,839), Moderna (10,118) and the National Institutes of Health (3,905). The list also contains entity-linking errors: the most frequent "organisation" (23,735 articles) is an unrelated group to which the abbreviation CDC was wrongly linked. We therefore draw no finding from the organisation list.

3.5 Research attention and invention

Research on COVID-19 grew and declined on a slower clock than news. Papers with the topic in the title numbered 156,619 in 2020 (239.6 per 10,000 papers), peaked at 215,544 in 2021 (344.45 per 10,000, or 3.44% of all papers) and then fell to 172,512 in 2022, 120,179 in 2023, 72,453 in 2024 and 44,768 in 2025 (63.64 per 10,000). From 2020 to 2025, 782,075 such papers were published. The most-cited are early clinical descriptions of patients in Wuhan and the description of how the virus enters cells [1–6].

News and research peaked a year apart (Figure 9). News share was highest in 2020; research share was highest in 2021. By 2025 research share was 5.41 times below its peak, while news share had already fallen 37.74-fold from 2020 to 2024. Because the topic barely existed in 2019, both series are tiny in that base year, so the index values in Figure 9 are very large; the figure is useful for the shape and timing of the two curves, not their level.

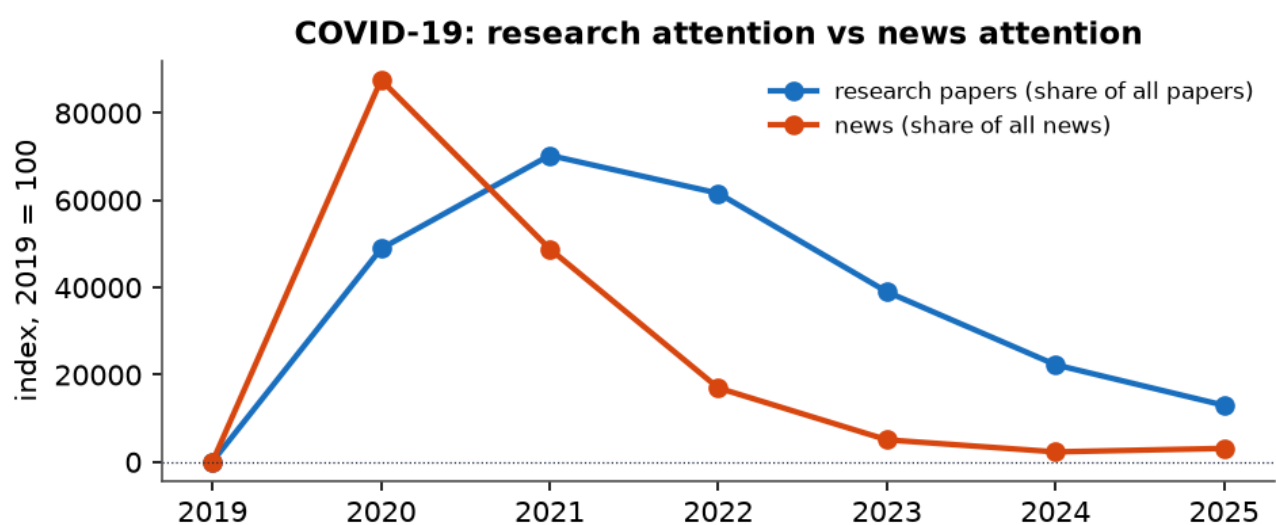


Figure 9. Research papers with COVID-19 terms in the title (share of all papers) and COVID-19 news (share of all news), each indexed to 2019 = 100. Both series are close to zero in 2019, so only the shape and timing are meaningful.

Patenting peaked later still. 1,178 COVID-19 patent families were published in 2020, 2,897 in 2021, 3,418 in 2022, 3,272 in 2023 and 2,188 in 2024. China's office holds 37.5% of office filings of these families from 2015 to 2024, followed by international applications (20.1%), the United States (12.8%) and the European Patent Office (8.0%) (Figure 10). The leading technology classes are medicinal preparations (A61K, 2,560 families), peptides and antibodies (C07K, 1,676), and testing and analysis (G01N, 1,080), matching the vaccine, therapeutic and diagnostic sides of the response.

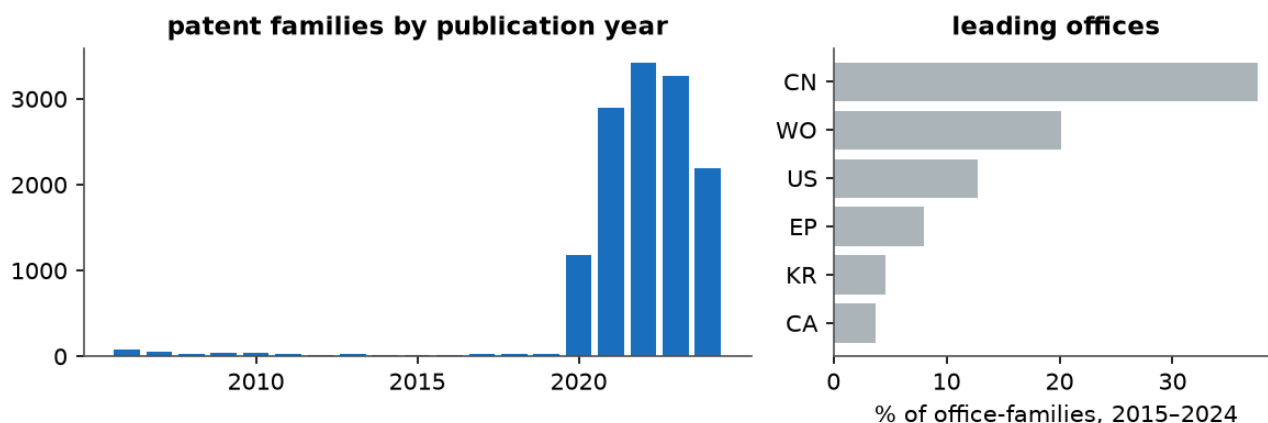


Figure 10. COVID-19 patent families by publication year (left) and the leading filing offices, 2015-2024 (right).

4. Discussion

Three points matter for anyone who uses news as a record of the pandemic. First, coverage measures attention, not the course of the disease, and attention fell quickly. Within four years of its peak the topic's share of news was 37.74 times smaller. Second, raw shares and lifts tell different stories. The United States produces most located COVID-19 news, but China, New Zealand and Brazil are where the topic took the largest share relative to their weight in the index in 2023-2026, and India is under-represented. Within India, the over-representation of Kerala, Delhi, Goa and Maharashtra, and the under-representation of Bihar, Jammu and Kashmir and Jharkhand, describe where the late-phase coverage was concentrated. US state lifts, by contrast, mostly reflect which states the August 2025 back-fill was tagged with. Third, the index itself has artefacts that a keyword count would mistake for events: a back-fill of old stories in August 2025, misdated items in 2019 and research abstracts filed as news.

News, research and invention run on different clocks. News peaked in April 2020 and declined steadily. Research peaked in 2021 and patent families in 2022, each following the previous record by about a year. The dataset lets researchers link the three, for example to test whether the countries with the most COVID-19 coverage are those whose research output on the disease persisted longest.

5. Limitations

- **Index availability.** 92.88% of the index was readable at build time, so counts are lower bounds. Shares and lifts are less affected, because the missing part is spread across topics.
- **Back-filled and misdated items.** Old stories entered the index with August 2025 dates, and 1,941 articles dated 2019 are mostly misdated. 2025 is not used as an end point for trends.

- **Coverage after May 2026.** The index thins after May 2026, so 2026 is partial and is excluded from year comparisons.
- **Window for labels.** Geography, tone, priority and themes cover January 2023 – July 2026, the late phase of the pandemic. They do not describe the 2020–2021 peak, and they include the August 2025 back-fill, which is mostly located in the United States and overstates the US share.
- **Matching.** Matching uses the article opening. Articles that mention COVID-19 only later are missed, and 5% of matches are not about COVID-19. In 30% of matches COVID-19 is context rather than the main subject.
- **Machine labels.** Location, tone, priority, themes and organisations are machine-labelled. The city field is noisy and is not analysed, and the organisation list contains linking errors.
- **Place tags from ordinary words.** Place tags derived from ordinary words were treated as unlocated in every location field, which removed 15.3% of city-tagged articles. Some genuine places that share a name with a common word are lost with them, and remaining tags can still be wrong.
- **Outlet mix.** The index over-represents English-language outlets and the United States.
- **Research and patent counts.** These match titles only. The research corpus contains repository and regional index records that the venue filter does not fully remove. Office shares reflect filing practice.

Data availability

The dataset (article sample, monthly, country and state tables, research and patent tables, all metrics) is released under CC BY-NC 4.0 with this paper. The sample holds 97,067 of 4,667,790 matched articles to 31 July 2026. Full article text, headlines, publishers, later articles and other topics are served by the Prism News API: <https://github.com/fliarbi/prism-news-api> (base URL <https://radar-api.fliarbi.com/news>). A worked query: `prism-news trends "covid" --since 2019-01-01 --until 2026-07-31 --granularity monthly`.

References

1. Chaolin Huang et al. (2020). Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *The Lancet*. [https://doi.org/10.1016/s0140-6736\(20\)30183-5](https://doi.org/10.1016/s0140-6736(20)30183-5)
2. Wei-jie Guan et al. (2020). Clinical Characteristics of Coronavirus Disease 2019 in China. *New England Journal of Medicine*. <https://doi.org/10.1056/nejmoa2002032>
3. Fei Zhou et al. (2020). Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. *The Lancet*. [https://doi.org/10.1016/s0140-6736\(20\)30566-3](https://doi.org/10.1016/s0140-6736(20)30566-3)
4. Nanshan Chen et al. (2020). Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. *The Lancet*. [https://doi.org/10.1016/s0140-6736\(20\)30211-7](https://doi.org/10.1016/s0140-6736(20)30211-7)
5. Markus Hoffmann et al. (2020). SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. *Cell*. <https://doi.org/10.1016/j.cell.2020.02.052>
6. Dawei Wang et al. (2020). Clinical Characteristics of 138 Hospitalized Patients With 2019 Novel Coronavirus-Infected Pneumonia in Wuhan, China. *JAMA*.

<https://doi.org/10.1001/jama.2020.1585>

7. Fliar BI (2026). *COVID-19 in the News 2019-2026: article-level dataset and coverage tables*. Prism News API. <https://github.com/fliarbi/prism-news-api>