

Lung Cancer in the News, 2019-2026: Attention, Geography and Tone of Global Lung Cancer Coverage, 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 lung cancer was covered in a global news index from January 2019 to July 2026, using 30,856 de-duplicated lung cancer news articles, and set that coverage beside 154,532 research papers and 6,434 patent families on lung cancer. Six results stand out:

1. **Falling share of attention.** Lung cancer took 2.38 of every 10,000 news articles in 2019 and 1.79 in 2025, a 24.8% fall, although the number of articles rose from 2,816 to 4,824. The peak month was February 2021, at 3.95 articles per 10,000.
2. **Awareness calendar.** February is the busiest month of a typical year (10.16% of coverage), and February, August and November together hold 28.66%.
3. **Geography.** The United States holds 60.4% of located lung cancer news, 1.25 times its share of all news. India holds 5.71%, a lift of 0.43, and Africa as a whole has a lift of 0.21. Within the United States, Massachusetts (3.46), Illinois (2.42) and Missouri (1.9) lead.
4. **Tone and urgency.** Only 27.8% of lung cancer articles are negative in tone (all news: 33.6%), yet 86.1% are rated critical priority (all news: 42.6%).
5. **Measurement.** 72.05% of raw matches for lung cancer vocabulary were removed by rules for scholarly journal abstracts, market notices and boilerplate. Counting them would multiply coverage and distort its timing.
6. **Research and invention.** Lung cancer papers doubled their share of the literature (14.14 to 28.86 per 10,000 papers, 2015-2025). Patent families grew 2.21-fold from 2015 to 2024, and China's office held 61.3% of office filings.

Findings stated for citation

- **F1.** Lung cancer's share of news fell 24.8% from 2019 to 2025 (2.38 to 1.79 per 10,000 articles).
- **F2.** Its share fell 19.3% from 2019 to 2020.
- **F3.** February 2021 was the peak month, at 3.95 articles per 10,000.
- **F4.** February, August and November held 28.66% of a typical year's coverage.
- **F5.** 72.05% of raw lung cancer matches were removed by rules for journal abstracts, market notices and boilerplate.
- **F6.** The United States held 60.4% of located coverage (lift 1.25); India 5.71% (lift 0.43).
- **F7.** By continent, North America's lift was 1.23 and Africa's 0.21.
- **F8.** Massachusetts (3.46), Illinois (2.42) and Missouri (1.9) had the most over-represented US coverage.

- **F9.** India's lung cancer coverage per 10,000 of its news rose 1.44-fold from 2023 to 2025.
- **F10.** 27.8% of lung cancer articles were negative in tone, 0.83 times the all-news share.
- **F11.** 86.1% were rated critical priority, 2.02 times the all-news share.
- **F12.** The theme "lung cancer early detection methods" appeared in 48.97% of themed lung cancer articles.
- **F13.** Lung cancer papers rose 2.04-fold as a share of all papers between 2015 and 2025.
- **F14.** Lung cancer patent families published per year grew 2.21-fold between 2015 and 2024.
- **F15.** China's office held 61.3% of office filings of lung cancer patent families (2015–2024).

1. Introduction

Lung cancer is the most commonly diagnosed cancer worldwide and the leading cause of cancer death. News coverage shapes what the public knows about its causes, screening and treatment, and which patients, drugs and policies receive attention. Yet coverage is rarely measured as data in its own right. Cancer registries record cases. Bibliometric studies record research. Few sources record attention: how much of the news lung cancer takes, and when, where 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 (7,625 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, an editorial 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 lung cancer when its opening text contains *lung cancer(s)*, *lung carcinoma*, *lung adenocarcinoma*, *NSCLC*, *small cell lung* or *lung tumour/tumor* as whole words. The exact phrase "lung cancer" appears in 92.35% of matched articles. Three families of matches are excluded:

- scholarly journal abstracts, which the index also carries. They are recognised by their layout (a paper title joined directly to the first sentence of its summary, with scientific vocabulary in the title) or by section labels such as "background:" and "methods:";
- market-forecast notices ("market size", "CAGR", "forecast period");
- a recurring footer from one patient-organisation website.

Together these removed 79,536 of 110,392 raw matches (72.05%). Many abstracts carry a default date of 1 January, so without this step January would appear to be the peak month of

every year. One reviewer read 40 randomly drawn matched openings after the rules were final. 39 were news, press or feature items about lung cancer (precision 97.5%), and in 20 (50%) lung cancer was the main subject. The one miss was an abstract without a title line. An earlier draw under looser rules had found 8 of 40 abstracts, which is how the exclusions were tuned.

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. Seasonality uses each year's within-year month shares, averaged over the complete years 2019–2025. Geography, tone, priority and themes use January 2023 – July 2026. Geography is expressed as lift: a place's share of lung cancer articles divided by its share of all located articles. Place tags derived from ordinary words (such as "Congress", "Earth", "Winter" or "Hurricane"; 714 words) were treated as unlocated in every location field, which removed 15.3% of city-tagged articles.

Research and patents. Research papers are de-duplicated articles and reviews (2000–2025) with *lung cancer*, *lung carcinoma*, *lung adenocarcinoma*, *non-small cell lung*, *small cell lung* or *NSCLC* in the title. Repository deposits are excluded. Patent families (2006–2024) use the same title rule and are counted once per filing office.

3. Results

3.1 How much and when

The index holds 30,856 lung cancer news articles from January 2019 to July 2026. Raw counts rose from 2,816 in 2019 to 4,824 in 2025, but the share of all news did not (Table 1). It was 2.38 per 10,000 in 2019, fell 19.3% to 1.92 in 2020, recovered to 2.39 in 2023 and fell to 1.79 in 2025, 24.8% below 2019. The highest monthly share was February 2021, at 3.95 per 10,000, the month in which the death of the broadcaster Rush Limbaugh from lung cancer was reported (Figure 1). Other high months were June 2023 (3.27) and November 2023 (3.19).

Year	Lung cancer articles	All articles in index	Per 10,000
2019	2,816	11,807,177	2.38
2020	3,117	16,230,014	1.92
2021	3,693	17,281,220	2.14
2022	4,076	19,115,515	2.13
2023	4,886	20,484,506	2.39
2024	4,967	23,311,896	2.13
2025	4,824	26,990,006	1.79

Table 1. Lung cancer news articles per year and as a share of the index. 2026 (January–July) is partial and omitted.

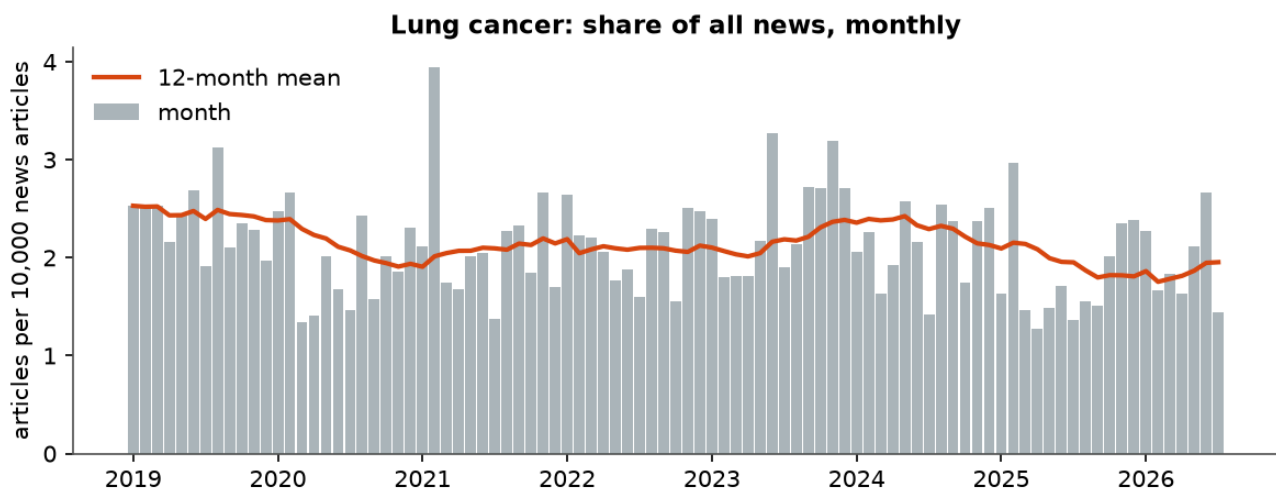


Figure 1. Lung cancer coverage as articles per 10,000 news articles in the index, monthly, January 2019 – July 2026. Bars are months; the line is the 12-month mean.

Coverage follows the awareness calendar more than the seasons (Figure 2). February holds the largest share of a typical year (10.16%), followed by January (9.7%), November (9.29%) and August (9.21%). July is the quietest month (6.11%). February, August and November together hold 28.66%; these are the months of World Cancer Day, World Lung Cancer Day and lung cancer awareness month. The pattern is mild: an even spread would give those three months 25%.

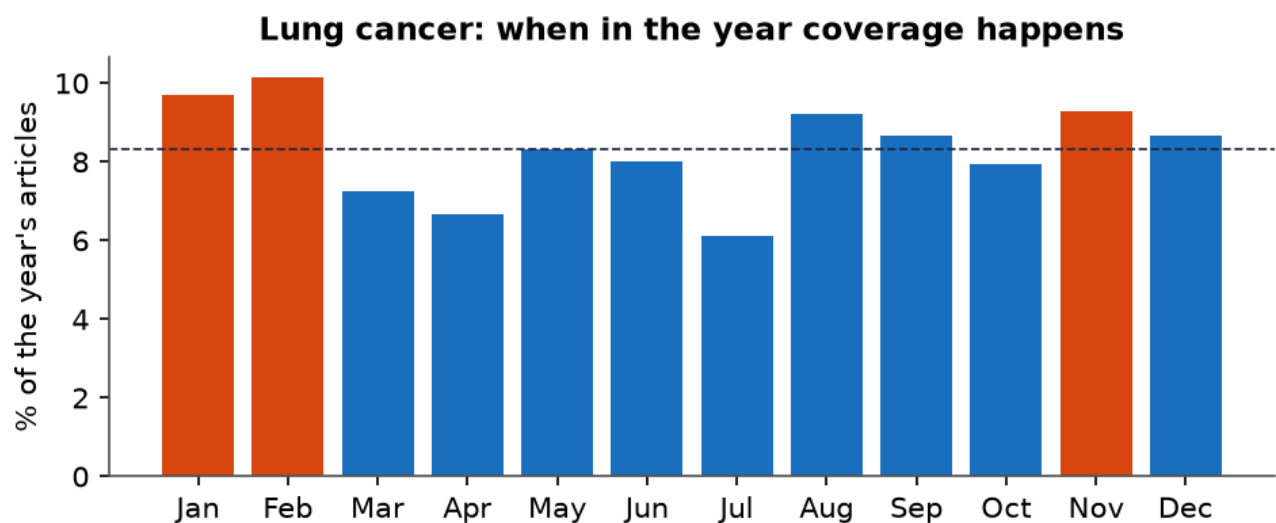


Figure 2. Share of each year's lung cancer articles published in each month, averaged over 2019-2025. The dashed line is an even spread.

3.2 Where

Between January 2023 and July 2026, 7,237 lung cancer articles carried a primary location (42.19% of the window). The United States holds 60.4% of them, with a lift of 1.25. China (1.96), Spain (1.57) and the United Kingdom (1.31) are also over-represented. India holds 5.71% with a lift of only 0.43, and Canada (0.8), Australia (0.78) and France (0.71) are below parity (Table 2, Figure 3). Only 8 countries reach the 100-article threshold for a lift, so the country picture is narrow.

Country	Located articles	Share %	Lift vs all news
United States	4,371	60.4	1.25
United Kingdom	554	7.66	1.31
India	413	5.71	0.43
China	373	5.15	1.96
Canada	124	1.71	0.8
Australia	121	1.67	0.78
Spain	111	1.53	1.57
France	101	1.4	0.71

Table 2. Countries with at least 100 located lung cancer articles, January 2023 – July 2026, with lift against all located news.

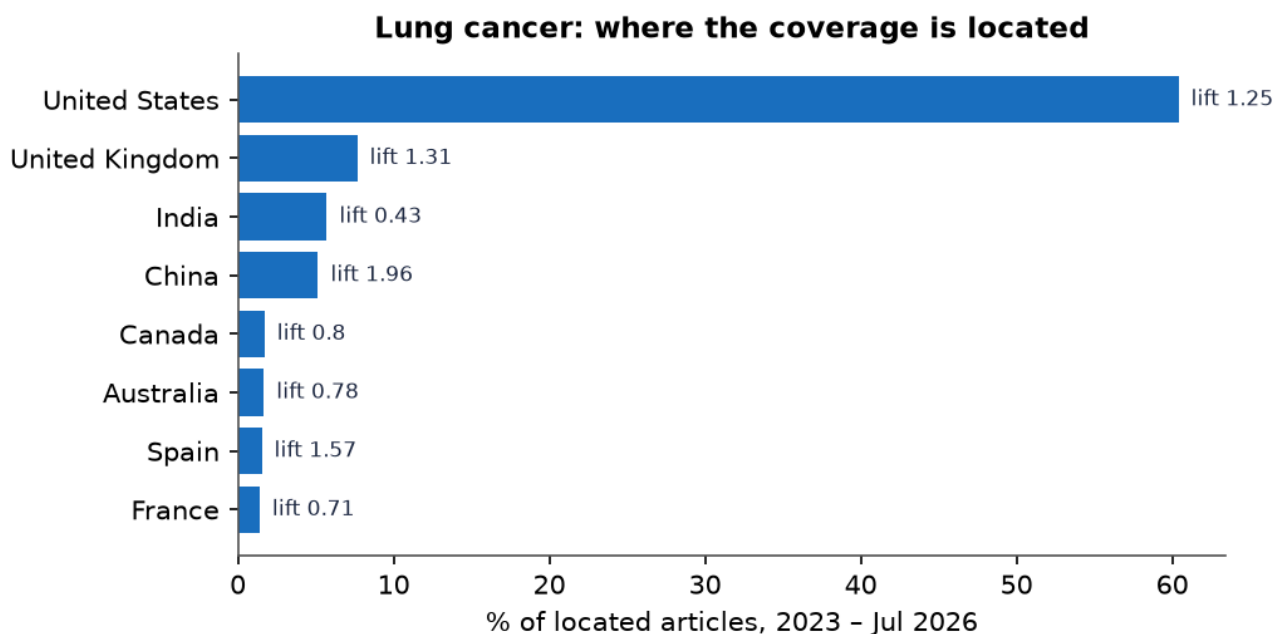


Figure 3. Leading countries by share of located lung cancer articles, 2023 – Jul 2026, labelled with lift against all located news.

3.3 Geo intelligence: continents, states and shifting attention

The world map (Figure 4) shows the same pattern at a glance. Lung cancer is over-represented in the news of China, Spain, the United Kingdom and the United States, and under-represented in India, Canada, Australia and France; most other countries have too few located articles to colour. By continent, North America (1.23) is over-represented, Europe (1.04) and Oceania (0.96) are close to parity, and Asia (0.67), South America (0.66) and Africa (0.21) are under-represented (Figure 6, left). This is consistent with the weight of company and trial announcements in the coverage, but the index does not test that link.

Lung cancer: which countries over- and under-cover the topic

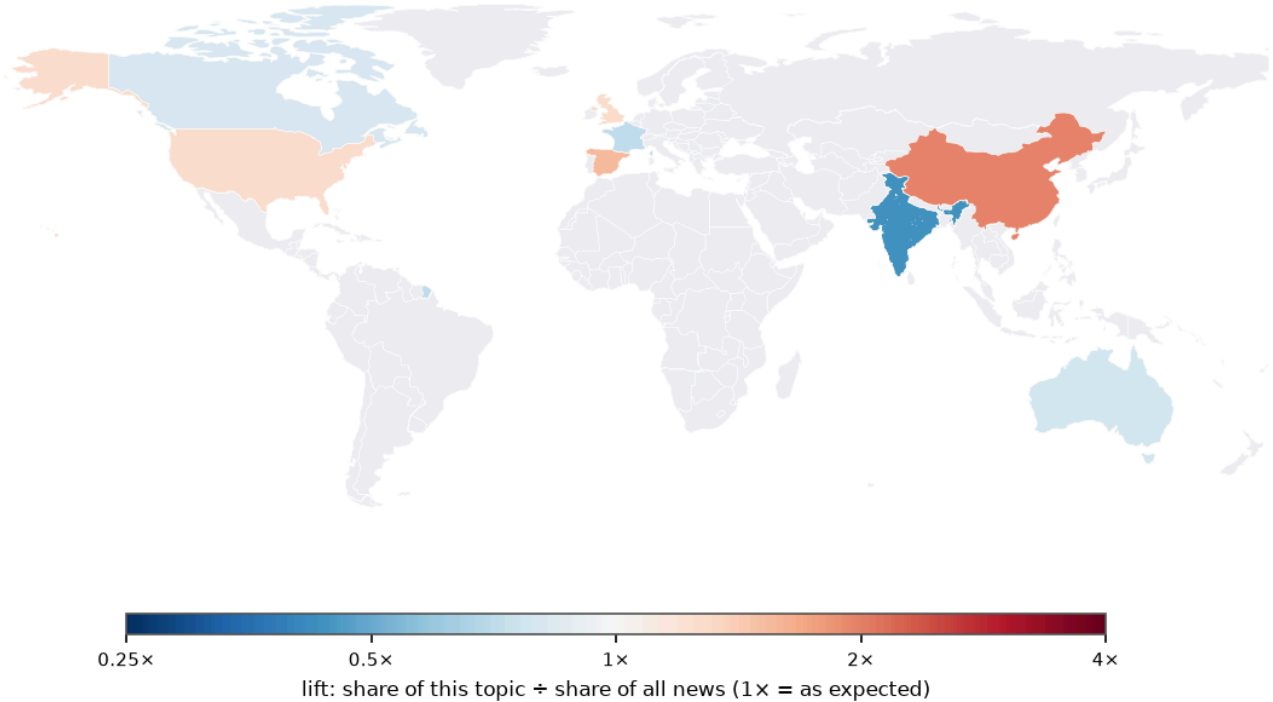


Figure 4. Lift of each country's share of located lung cancer articles, January 2023 – July 2026 (red = over-represented, blue = under-represented, grey = fewer than 100 articles). India is drawn with its official outline.

Within the United States the picture is sharper (Table 3, Figure 5). Measured against each state's share of American news, and counting only the states with at least 50 of the 3,446 located US articles, Massachusetts (3.46), Illinois (2.42), Missouri (1.9) and California (1.78) are the most over-represented. Minnesota is left out: most of its sample rows are tagged with the town of Orr, which is the trial abbreviation ORR (objective response rate) read as a place name. These states host large cancer centres and biotechnology clusters, which may explain part of the pattern. New Jersey (0.21) and Pennsylvania (0.4) are far below their weight in American news, and Florida (0.66), Washington (0.73) and Michigan (0.76) are also below parity. Within India, only 280 articles carry a state and only Delhi passes 50 (66 articles, lift 1.82), so no Indian state map is drawn and no state comparison is made.

Highest lift: state	Articles	Lift	Lowest lift: state	Articles	Lift
Massachusetts	283	3.46	New Jersey	134	0.21
Illinois	220	2.42	Pennsylvania	83	0.4
Missouri	65	1.9	Florida	122	0.66
California	613	1.78	Washington	66	0.73
New York	540	1.34	Michigan	61	0.76
North Carolina	72	1.31	Texas	220	1.08

Table 3. US states with the highest and lowest lift for lung cancer coverage, January 2023 – July 2026, states with at least 50 located articles (Minnesota excluded, see text; Georgia, 72 articles and lift 1.12, falls between the two columns). Lift = the state's share of the country's located lung cancer articles ÷ its share of all located news from that country.

Lung cancer: states of United States by lift

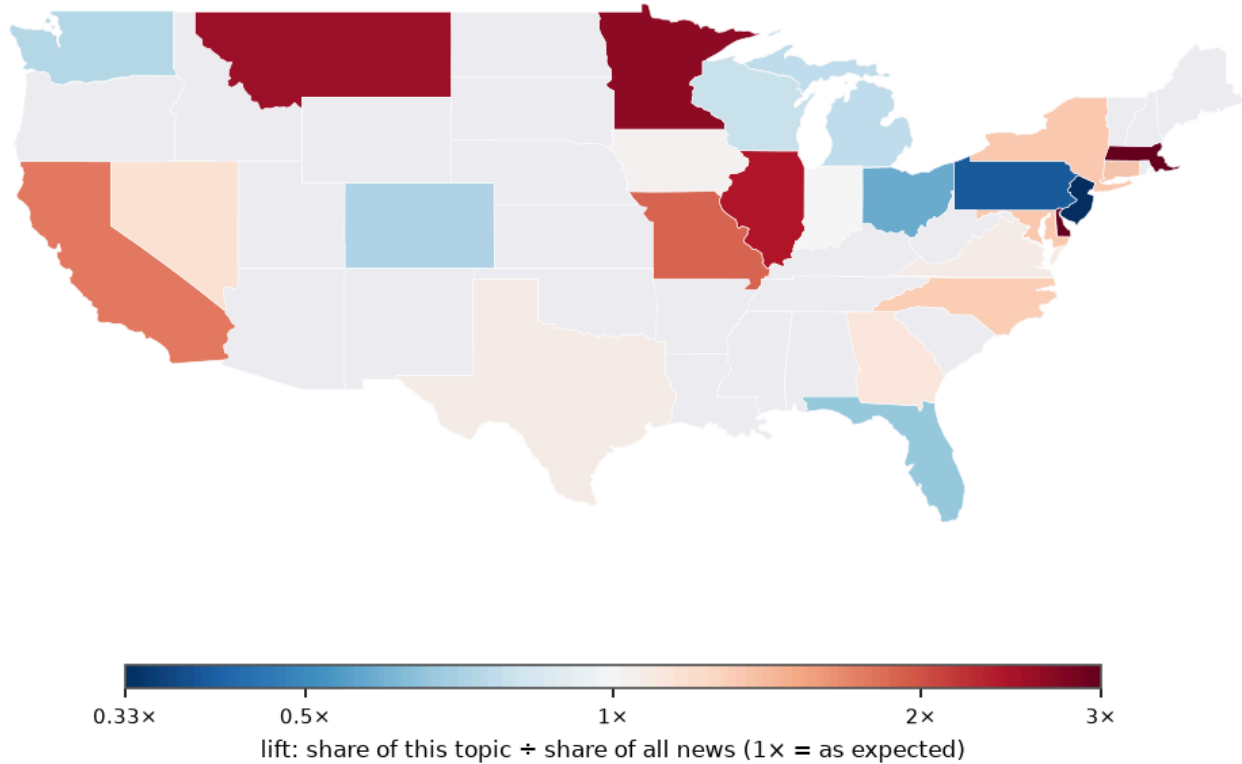


Figure 5. Lift of each US state's share of located lung cancer articles against its share of all located US news, January 2023 – July 2026.

Attention also moves between countries from year to year (Table 4, Figure 6, right). Measured as lung cancer articles per 10,000 of each country's located news, India rose from 0.62 in 2023 to 0.89 in 2025 (1.44 times) and China was level at 3.81 and 3.83 (1.01 times). The United States (0.75 times) and the United Kingdom (0.66 times) fell over the same years. China's share had risen from 1.12 in 2019 to a high of 5.39 in 2024.

Tone also differs by country when each is set against its own news. In China, 13.1% of lung cancer articles were negative against 27.2% of all Chinese news, the most positive coverage relative to its baseline. In the United States (28.9% against 33.1%) lung cancer coverage is a little less negative than the national news, and in the United Kingdom (36.2% against 36.9%) it sits at the baseline. India is the one country where it is more negative than the national news (45.0% against 41.2%).

Country	Per 10,000 of country's news, 2023	2025	Change	Negative % (topic)	Negative % (all news)
India	0.62	0.89	1.44	45.0	41.2
China	3.81	3.83	1.01	13.1	27.2
United States	2.65	1.99	0.75	28.9	33.1
United Kingdom	2.95	1.96	0.66	36.2	36.9

Table 4. Shifting attention and tone by country. Change = 2025 rate ÷ 2023 rate. Tone uses January 2023 – July 2026.

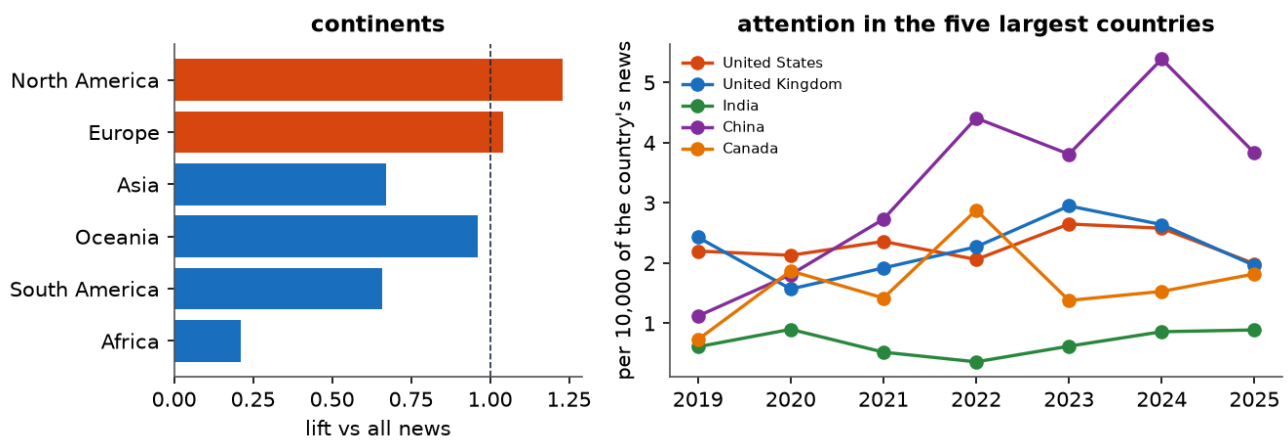


Figure 6. Left: lift by continent, January 2023 – July 2026. Right: lung cancer articles per 10,000 of each country's located news, 2019–2025, for the five countries with the most articles.

3.4 Tone, urgency and themes

Lung cancer coverage is less negative than news in general but rated far more urgent (Figure 7). Among articles with a tone label, 27.8% are negative, against 33.6% across the index (0.83 times), and 70.3% are positive (1.24 times the all-news share). Drug approvals, trial results and screening programmes are labelled positive. The machine-assigned editorial priority rating puts 86.1% of lung cancer articles at critical level (7 or higher), against 42.6% of all news (2.02 times).

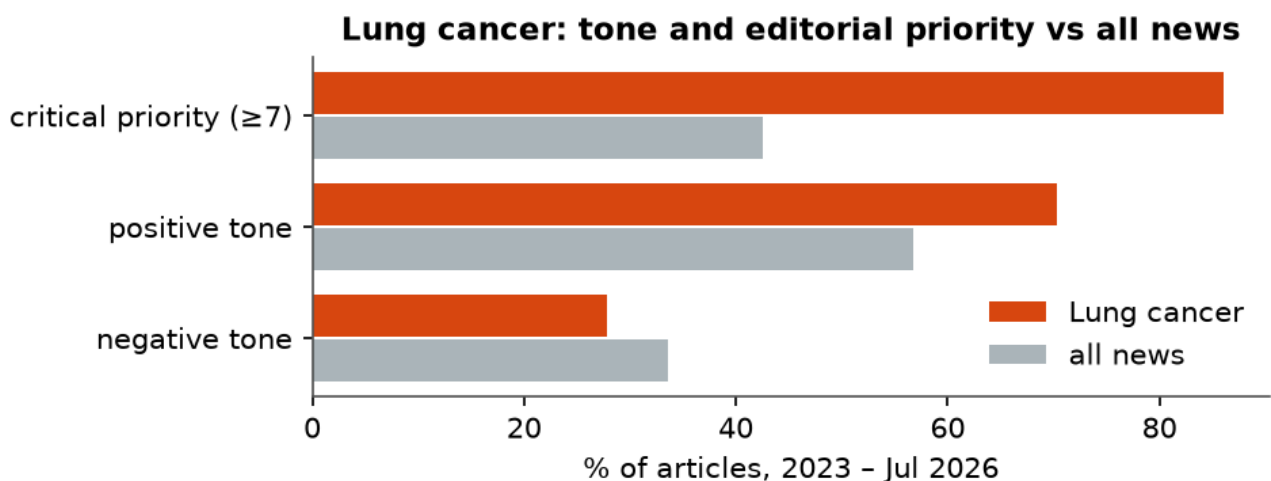


Figure 7. Tone and the share rated critical priority, lung cancer articles vs all news, 2023 – Jul 2026.

Themes show how the stories are framed (Table 5). The broad label "lung cancer early detection methods" is the most frequent theme, carried by 48.97% of themed articles, 492.49 times its rate in all news. Targeted therapy advances (19.76%) and liquid biopsy (6.4%) follow. Smaller themes on the link between smoking and lung cancer are the most specific to the topic. Obituaries (3.96%) and tributes to celebrities (2.64%) form a distinct strand: part of the coverage is about known people who have died of lung cancer.

Theme	Articles	Share %	Lift
lung cancer & smoking correlation	272	2.38	1752.14
health risks of tobacco smoking	227	1.98	1378.23
lung cancer screening and therapies	139	1.21	1253.56
smoking-related respiratory diseases	163	1.42	1035.57
lung cancer early detection methods	5,608	48.97	492.49
oncology and targeted therapy advancements	2,263	19.76	107.34
clinical trials and regulatory approvals	685	5.98	75.02
liquid biopsy for early detection	733	6.4	73.27

Table 5. Themes most specific to lung cancer coverage (lift against all themed news), January 2023 – July 2026. The organisations named most often are drug companies and cancer bodies: AstraZeneca (708 articles), the American Cancer Society (508), Merck (324), the National Cancer Institute (288) and Roche (269). A press-release wire and a professional social network also rank high, which reflects how much of this coverage begins as company announcements.

3.5 Research attention and invention

Research on lung cancer grew much faster than the scientific literature. Papers with lung cancer in the title rose from 14.14 per 10,000 papers in 2015 to 28.86 in 2025, a 2.04-fold increase in share, while the annual number of papers in the comparison corpus was about level (a factor of 1.0). Between 2015 and 2025, 154,532 such papers were published. The most-cited are the immunotherapy trials that changed treatment of non-small-cell lung cancer [1–4], early palliative care [5] and reviews of the disease's biology [6].

News attention does not track research attention (Figure 8). Research share almost doubled, while news share drifted down.

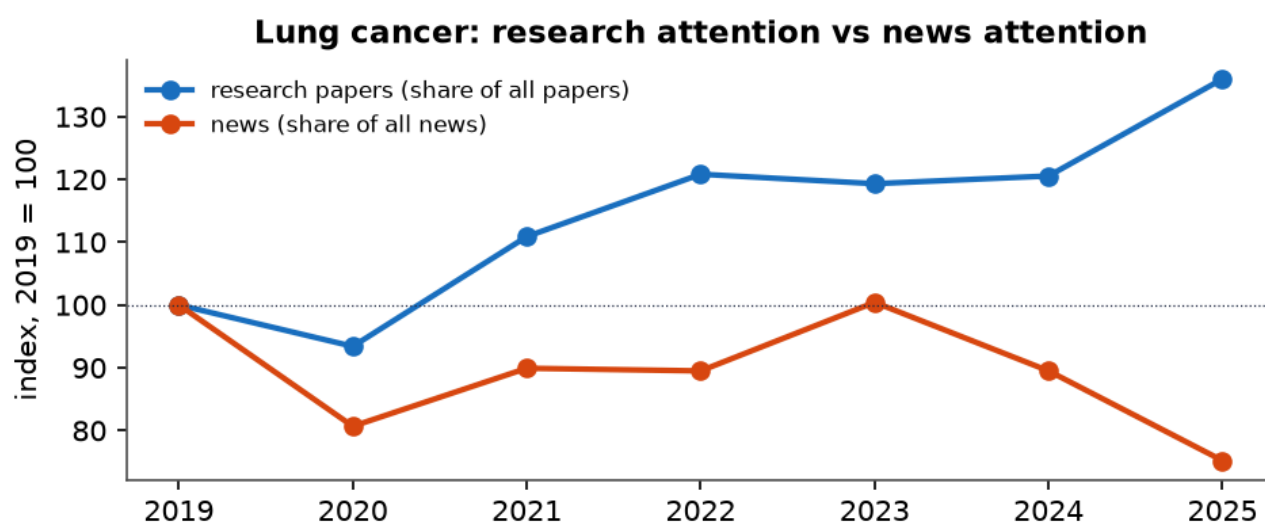


Figure 8. Research papers with lung cancer in the title (share of all papers) and lung cancer news (share of all news), each indexed to 2019 = 100.

Patenting also rose: 428 lung cancer patent families were published in 2015 and 948 in 2024 (2.21 times). China's office holds 61.3% of office filings (family-office pairs) from 2015 to 2024, followed by international applications (8.5%) and the United States (8.2%) (Figure 9). The leading technology classes are medical preparations (A61K, 1,023 families), nucleic-acid testing (C12Q, 937) and analytical methods (G01N, 650): diagnosis is nearly as large a patent field as treatment.

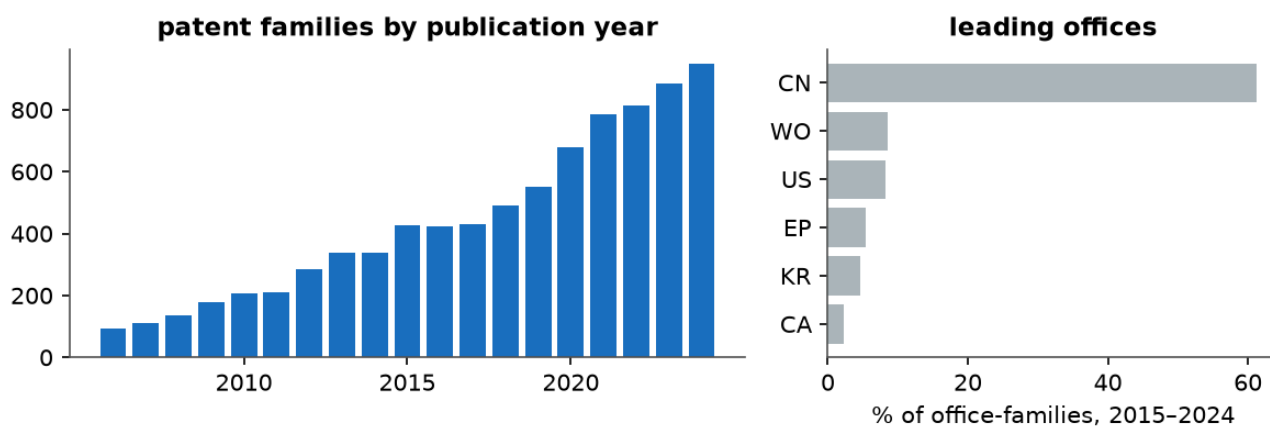


Figure 9. Lung cancer patent families by publication year (left) and the leading filing offices, 2015-2024 (right).

4. Discussion

Three points matter for anyone who uses news as a record of lung cancer. First, coverage measures attention, not burden. Lung cancer's share of news fell between 2019 and 2025 while research on it doubled its share of the literature. The highest month was the month of a well-known broadcaster's death, and the busiest months of the year are those of awareness days; the data do not test either link, and they say nothing about incidence. Second, raw shares mislead in a different way here than for most topics. The United States is over-represented, not just large, which is consistent with the weight of company and trial news in this coverage. India takes less than half its expected share, although its attention rose from 2023 to 2025. Canada, which looked over-represented before place tags derived from ordinary words were removed, is below parity once they are. Third, most raw keyword matches in this news index are removed by the abstract and notice rules. Without excluding them, the topic would look more than three times larger, its peak month would be January, and its geography and tone would describe journal publishing rather than journalism.

Research, invention and news run on different clocks. Research share grows steadily. Patent filings grow and are concentrated in one office, with diagnostics close behind therapeutics. News follows announcements and the awareness calendar. The dataset lets researchers link the three, for example to test whether drug approvals reported in the news are followed by changes in research or patent activity.

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.
- **Coverage after May 2026.** The index thins after May 2026, so 2026 is partial and is excluded from year comparisons.
- **Matching and exclusion.** Matching uses the article opening. Articles that mention lung cancer only later are missed. The abstract exclusion works on wording: a few abstracts remain (1 of 40 in the hand check), and some news items written in the same layout are lost.
- **Obituaries.** Many matched articles are obituaries that name lung cancer as the cause of death. They are kept as coverage, but only 50% of matched articles are mainly about lung cancer.

- **Machine labels.** Location, tone, priority and themes are machine-labelled. Fewer than half of the articles carry a country. The city field is noisy and is not analysed. Place tags derived from ordinary words were treated as unlocated (15.3% of city-tagged articles), so location counts are lower than the tagged totals and a few true places with common-word names are lost. Some abbreviations still pass as places (ORR, a trial measure, is read as Orr, Minnesota), so Minnesota is left out of the state ranking.
- **Outlet mix.** The index over-represents English-language outlets and press-release wires. This shapes both geography and tone.
- **Research and patent counts.** These match titles only. Office shares reflect filing practice.

Data availability

The dataset (article sample, monthly and country tables, research and patent tables, all metrics) is released under CC BY-NC 4.0 with this paper. The sample holds 7,625 of 30,853 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 "lung cancer" --since 2019-01-01 --until 2026-07-31 --granularity monthly`.

References

1. Martin Reck et al. (2016). Pembrolizumab versus Chemotherapy for PD-L1-Positive Non-Small-Cell Lung Cancer. *New England Journal of Medicine*. <https://doi.org/10.1056/nejmoa1606774>
2. H. Borghaei et al. (2015). Nivolumab versus Docetaxel in Advanced Nonsquamous Non-Small-Cell Lung Cancer. *New England Journal of Medicine*. <https://doi.org/10.1056/nejmoa1507643>
3. Julie R. Brahmer et al. (2015). Nivolumab versus Docetaxel in Advanced Squamous-Cell Non-Small-Cell Lung Cancer. *New England Journal of Medicine*. <https://doi.org/10.1056/nejmoa1504627>
4. Naiyer A. Rizvi et al. (2015). Mutational landscape determines sensitivity to PD-1 blockade in non-small cell lung cancer. *Science*. <https://doi.org/10.1126/science.aaa1348>
5. Jennifer S. Temel et al. (2010). Early Palliative Care for Patients with Metastatic Non-Small-Cell Lung Cancer. *New England Journal of Medicine*. <https://doi.org/10.1056/nejmoa1000678>
6. Roy S. Herbst et al. (2018). The biology and management of non-small cell lung cancer. *Nature*. <https://doi.org/10.1038/nature25183>
7. Fliar BI (2026). *Lung Cancer in the News 2019–2026: article-level dataset and coverage tables*. Prism News API. <https://github.com/fliarbi/prism-news-api>