

Breast Cancer in the News, 2019-2026: Seasonality, Geography and Tone of Global 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 breast cancer was covered in a global news index from January 2019 to July 2026, using 78,123 de-duplicated breast cancer articles, and set that coverage beside 233,244 research papers (2015–2025) and 7,888 patent families (2015–2024) on the disease. Six results stand out:

1. **An October topic.** On average 22.9% of each year's breast cancer coverage appears in October, Breast Cancer Awareness Month, 2.75 times an even spread. October was the month of highest share in every year from 2019 to 2025.
2. **Attention, not occurrence.** Breast cancer took a 31.1% smaller share of news in 2020 than in 2019 (4.43 against 6.43 articles per 10,000) and has not regained its 2019 level since.
3. **Geography.** The United Kingdom's share of located breast cancer news is 1.94 times its share of all news, and the United States' 1.14 times. Asia is under-represented (0.49), with India and China both at 0.47, although India's share of its own news rose 1.4 times from 2023 to 2025. Within the United States, Connecticut (2.23) and Massachusetts (2.12) lead; within India, Telangana (2.4) and Maharashtra (1.85).
4. **Tone and urgency.** 72.4% of breast cancer articles are positive in tone (all news: 56.8%), yet 72.5% are rated critical priority (all news: 42.6%).
5. **Measurement.** 57.2% of raw matches for breast cancer vocabulary were removed, most of them research-paper abstracts carried by the index. Counting them would more than double the apparent volume of news.
6. **Research and invention.** Breast cancer papers grew from 25.27 to 41.37 per 10,000 papers between 2015 and 2025 (1.64 times). Patent families grew 1.72-fold from 2015 to 2024, and 53.3% of their office filings were in China.

Findings stated for citation

- **F1.** Between 2019 and 2025, an average of 22.9% of each year's breast cancer news appeared in October, against 8.33% for an even spread.
- **F2.** October was the month of highest share in every year 2019–2025, at 11.25 to 17.4 articles per 10,000.
- **F3.** Breast cancer took a 31.1% smaller share of news in 2020 than in 2019 (4.43 vs 6.43 per 10,000).
- **F4.** The United States held 55.36% of located breast cancer news (lift 1.14), the United Kingdom 11.31% (lift 1.94) and India 6.18% (lift 0.47).
- **F5.** Ghana's lift was 2.87, the highest of any country with 100 located articles (Sweden: 2.84); 18.6% of its breast cancer articles were negative, against 43.3% of

all Ghanaian news.

- **F6.** 72.4% of breast cancer articles were positive in tone, 1.27 times the all-news share.
- **F7.** 72.5% of breast cancer articles had a machine-assigned critical priority rating, 1.7 times the all-news share.
- **F8.** Screening guidelines appeared in 13.29% of themed articles, the theme most specific to the topic.
- **F9.** Celebrity health updates appeared in 9.23% of themed articles.
- **F10.** Breast cancer papers rose 1.64-fold as a share of all papers between 2015 and 2025.
- **F11.** 53.3% of office filings of breast cancer patent families (2015–2024) were in China.
- **F12.** Breast cancer patent families published per year grew 1.72-fold between 2015 and 2024.
- **F13.** Within the United States, Connecticut (2.23), Massachusetts (2.12) and North Carolina (1.84) had the most over-represented coverage, New Jersey (0.13) the least.
- **F14.** Breast cancer's share of India's news rose 1.4 times from 2023 to 2025; Telangana (2.4) led Indian states by lift.
- **F15.** The United Kingdom's share rose from 9.44 to 9.96 per 10,000 of its news (2023–2025); Canada's fell from 7.14 to 3.91.

1. Introduction

Breast cancer is the most commonly diagnosed cancer in women in most countries. It is also one of the most visible diseases in public life, with a month of awareness campaigns, a widely recognised symbol and a long history of patient advocacy. News coverage is part of that visibility. It shapes what readers know about screening, treatment and risk, and it records which stories about the disease are told.

Coverage is rarely measured as data in its own right. Cancer registries record cases. Bibliometric studies record research. Few sources record attention, meaning how much of the news breast 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 (19,424 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 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 breast cancer when its opening text contains *breast cancer(s)*, *breast carcinoma(s)*, *breast tumour(s)* or *breast tumor(s)*, *mammogram(s)*,

mammography, *mastectomy* or *mastectomies* as whole words. Four families of matches are excluded:

- research-paper abstracts that the index carries alongside news, recognised by a paper title followed directly by abstract headings ("background:", "purpose:") or by the vocabulary of study reports ("this study aimed", "cell lines", "case report");
- market-report press releases;
- stories that use "mastectomy" for gender-affirming surgery;
- thousands of syndicated copies of three wire stories published in September 2021.

Together these removed 104,249 of 182,372 raw matches (57.2%). One reviewer read 40 randomly drawn matched openings. 35 were news or press items about breast cancer (precision 87.5%), and in 30 (75%) breast cancer was the main subject. Of the five that were not, four were research abstracts that the patterns missed (10% of the check) and one was a soap-opera storyline. The remaining abstracts carry no primary location, so they do not enter the geographic measures.

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 country's share of breast cancer articles divided by its share of all located articles. The same rule gives state lifts within India and the United States, against each country's own located news. 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.

Research and patents. Research papers are de-duplicated articles and reviews (2000–2025) with *breast cancer*, *breast carcinoma*, *breast tumour*, *breast neoplasm*, *mammography*, *mastectomy* or *triple-negative breast* in the title. Repository deposits and veterinary senses (canine and feline mammary tumours, bovine mastitis) are excluded. Patent families (2005–2024) use a similar title rule with *breast lesion* and *mammographic* added, and are counted once per filing office.

3. Results

3.1 How much and when

The index holds 78,123 breast cancer articles from January 2019 to July 2026. Raw counts rose from 7,596 in 2019 to 13,916 in 2025, but the share of all news did not: 6.43 per 10,000 in 2019, 5.7 in 2023 and 5.16 in 2025 (Table 1). The low point was 2020, at 4.43 per 10,000. That is 31.1% below 2019. No cause for the fall is tested here. The highest monthly share was October 2019, at 17.4 per 10,000 (Figure 1).

Year	Breast cancer articles	All articles in index	Per 10,000
2019	7,596	11,807,177	6.43
2020	7,195	16,230,014	4.43
2021	9,069	17,281,220	5.25
2022	10,263	19,115,515	5.37
2023	11,672	20,484,506	5.7
2024	12,620	23,311,896	5.41
2025	13,916	26,990,006	5.16

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

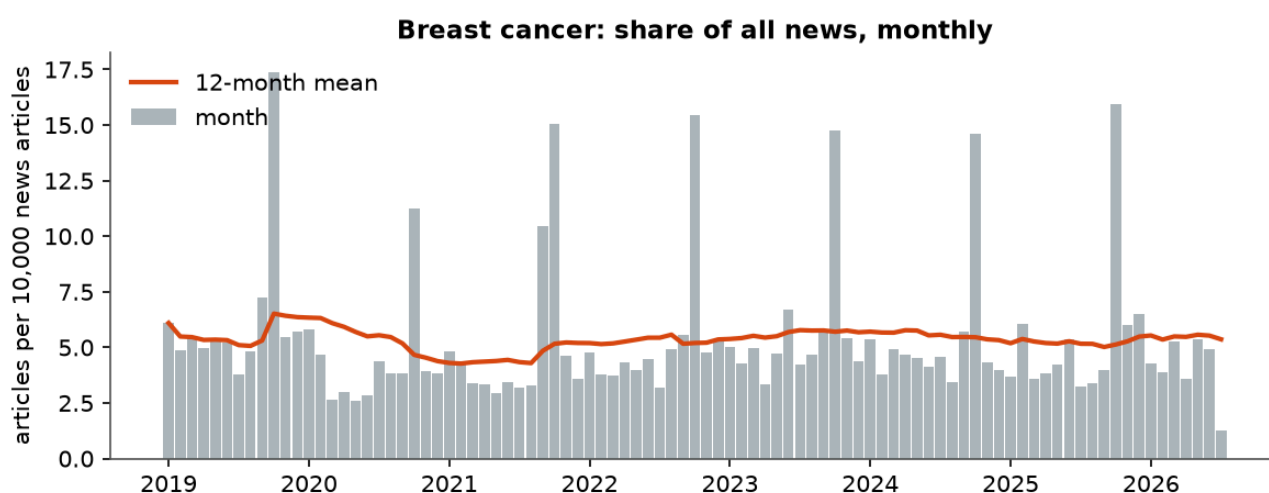


Figure 1. Breast cancer coverage as articles per 10,000 news articles, monthly, January 2019 – July 2026. Bars are months; the line is the 12-month mean.

Coverage is dominated by one month (Figure 2). October holds 22.9% of a typical year's breast cancer articles, 3.9 times the share of April, the quietest month (5.87%). September (9.83%) and January (8.62%) come next. In every year from 2019 to 2025, October had the highest share of any month, between 11.25 and 17.4 articles per 10,000, against annual averages of 4.43 to 6.43. The pattern follows Breast Cancer Awareness Month, when campaigns, fundraising events and screening offers are reported. September 2021 also stands out (10.47 per 10,000), the month of the death of a British singer from the disease; we removed the most repeated wire copies of that story but do not treat the month as a finding.

Breast cancer: when in the year coverage happens

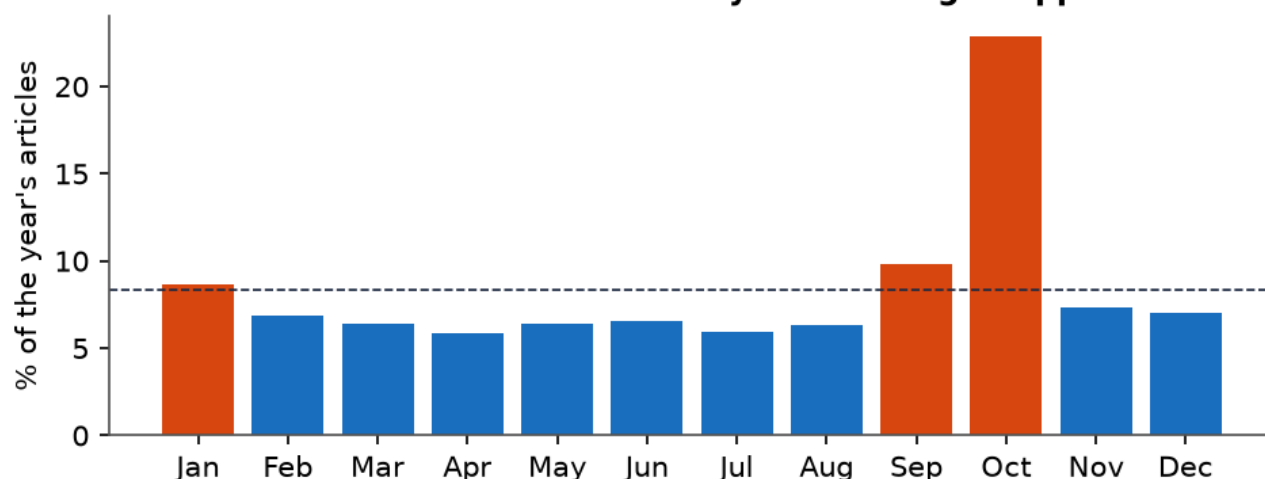


Figure 2. Share of each year's breast cancer articles published in each month, averaged over 2019–2025. Dashed line = an even spread (8.3% per month).

3.2 Where

Between January 2023 and July 2026, 17,362 breast cancer articles carried a primary location. The United States holds the largest raw share (55.36%), with a lift of 1.14. The United Kingdom's share (11.31%) comes with a lift of 1.94. India, the third-largest source of located articles (6.18%), is strongly under-represented at 0.47, as is China (0.47); France (0.61) is also well below parity. Among the 16 countries with at least 100 located articles, Ghana (2.87) and Sweden (2.84) are the most over-represented, followed by the United Kingdom (Table 2, Figure 3).

Country	Located breast cancer articles	Share %	Lift vs all news
United States	9,612	55.36	1.14
United Kingdom	1,963	11.31	1.94
India	1,073	6.18	0.47
Canada	499	2.87	1.34
Australia	484	2.79	1.3
Ghana	240	1.38	2.87
China	215	1.24	0.47
France	209	1.2	0.61
Brazil	189	1.09	1.26
Germany	177	1.02	0.87
Nigeria	176	1.01	0.74
Spain	166	0.96	0.98
Ireland	144	0.83	1.15
Italy	133	0.77	0.79
New Zealand	120	0.69	1.07
Sweden	108	0.62	2.84

Table 2. The 16 countries with at least 100 located breast cancer articles, January 2023 – July 2026, with lift against all located news.

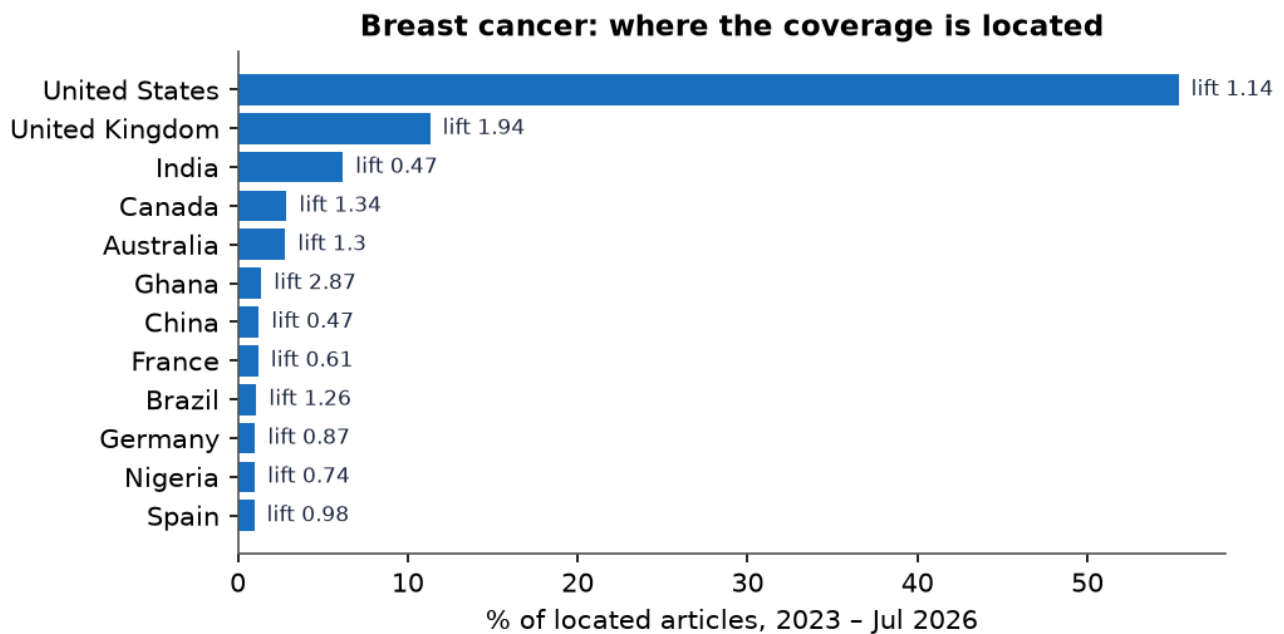


Figure 3. Leading countries by share of located breast cancer articles, 2023 – Jul 2026, labelled with lift against all located news.

3.3 Geo intelligence: continents, states and shifting attention

The world map (Figure 4) shows the pattern at a glance. Breast cancer is over-represented in the news of Ghana, Sweden, the United Kingdom, Canada, Australia, Brazil, Ireland, the United States and New Zealand, and under-represented in India, China and much of continental Europe, notably France (0.61) and Italy (0.79). By continent, Oceania (1.31), North America (1.18) and Europe (1.17) are over-represented. Africa (0.97) is close to parity, while South America (0.83) and Asia (0.49) are under-represented (Figure 6, left).

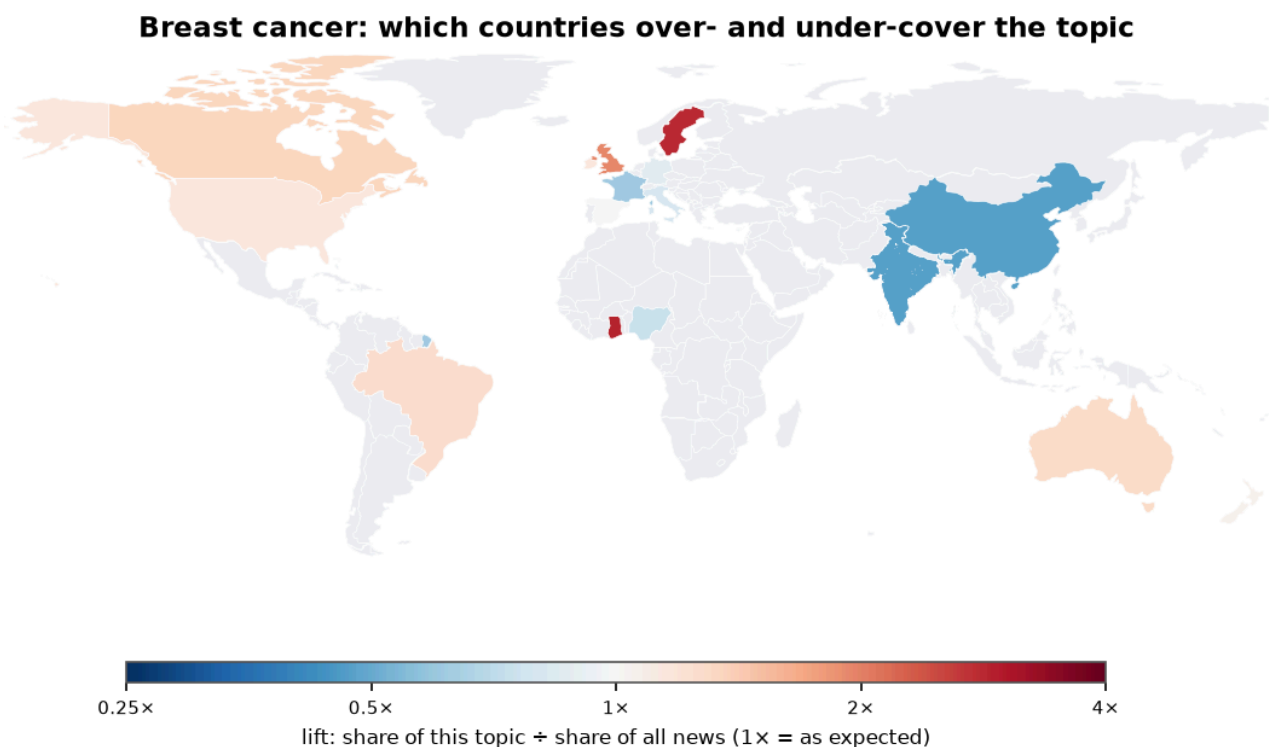


Figure 4. Lift of each country's share of located breast cancer articles, January 2023 – July 2026 (red = over-represented, blue = under-represented, grey = fewer than 100 articles). India is drawn with its official outline.

Within countries, the picture is sharper (Table 3). In the United States, 7,731 articles carried a state. Measured against each state's share of American news, breast cancer coverage is most over-represented in Connecticut (2.23), Massachusetts (2.12), North Carolina (1.84), Maryland (1.8) and Missouri (1.77). New Jersey (0.13) and Pennsylvania (0.61) are far below their weight in American news (Figure 5). In India, only 718 articles carried a state, and only five states reach 50 articles. Among them, Telangana (2.4), Maharashtra (1.85), Karnataka (1.71) and Tamil Nadu (1.61) are over-represented, and Delhi (1.1) is close to parity. Only six Indian states reach 30 articles, so no Indian state map is drawn for this topic.

Group	State	Articles	Lift
United States, highest	Connecticut	121	2.23
United States, highest	Massachusetts	389	2.12
United States, highest	North Carolina	226	1.84
United States, highest	Maryland	141	1.8
United States, highest	Missouri	136	1.77
United States, lowest	Oklahoma	64	0.68
United States, lowest	Pennsylvania	285	0.61
United States, lowest	New Jersey	191	0.13
India	Telangana	60	2.4
India	Maharashtra	211	1.85
India	Karnataka	68	1.71
India	Tamil Nadu	58	1.61
India	Delhi	103	1.1

Table 3. States with the highest and lowest lift for breast cancer coverage, January 2023 – July 2026, states with at least 50 located articles. Lift = the state's share of the country's located breast cancer articles ÷ its share of all located news from that country.

Breast cancer: states of United States by lift

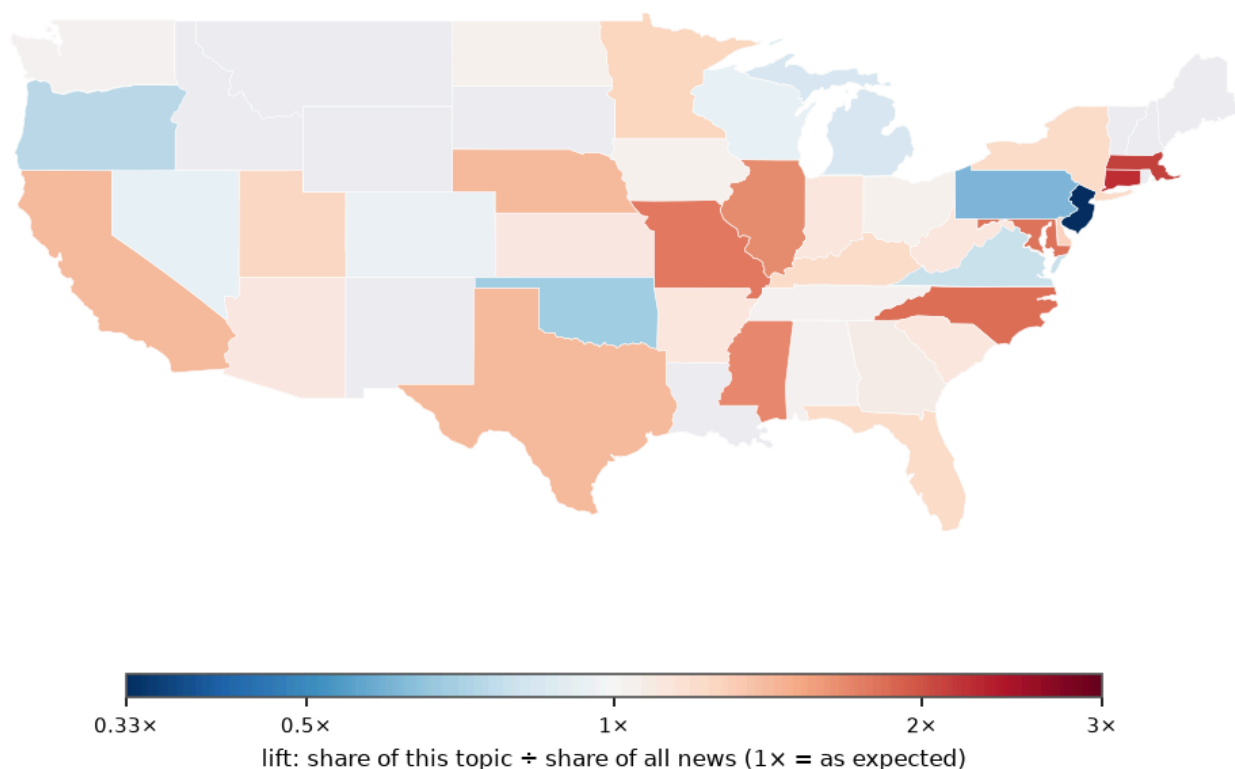


Figure 5. Lift of each US state's share of located breast cancer articles against its share of all located US news, January 2023 – July 2026.

Attention also moves between countries from year to year (Table 4, Figure 6 right). Measured as breast cancer articles per 10,000 of each country's located news, India rose from 1.88 in 2023 to 2.64 in 2025 (1.4 times) and the United Kingdom from 9.44 to 9.96 (1.06 times). Australia fell from 6.54 to 5.82 (0.89 times), the United States from 5.56 to 4.48 (0.81 times) and Canada from 7.14 to 3.91 (0.55 times). Tone differs by country too. Measured against each country's own news, breast cancer coverage is far less negative in India (18.9% negative against 41.2% for all Indian news), Ghana (18.6% against 43.3%), France (31.1% against 47.6%) and China (17.2% against 27.2%). It is slightly more negative than the country's news overall in the United Kingdom (38.0% against 36.9%) and more negative in Canada (33.3% against 27.1%).

Country	Per 10,000, 2023	Per 10,000, 2025	Change	Negative %, topic	Negative %, all news
United States	5.56	4.48	0.81	28.0	33.1
United Kingdom	9.44	9.96	1.06	38.0	36.9
India	1.88	2.64	1.4	18.9	41.2
Canada	7.14	3.91	0.55	33.3	27.1
Australia	6.54	5.82	0.89	33.5	38.4

Table 4. Attention and tone in the five countries with the most located breast cancer articles: articles per 10,000 of the country's located news in 2023 and 2025, and the negative share of breast cancer articles against the country's news overall (January 2023 – July 2026).

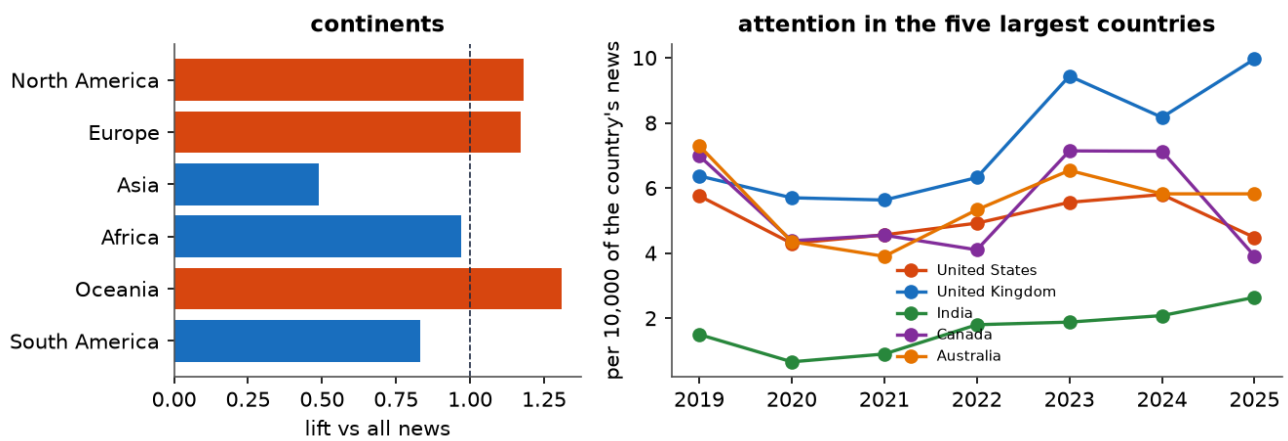


Figure 6. Left: lift by continent, January 2023 – July 2026. Right: breast cancer articles per 10,000 of each country's located news, 2019–2025, for the five countries with the most articles.

3.4 Tone, urgency and themes

Breast cancer coverage is more positive than news in general, and also more urgent (Figure 7). Among articles with a tone label, 72.4% are positive, against 56.8% across the index (1.27 times), and 26.0% are negative, against 33.6%. Survivor stories, fundraising events, screening offers and treatment advances make up much of the coverage. The machine-assigned priority rating puts 72.5% of breast cancer articles at critical level (7 or higher), against 42.6% of all news (1.7 times).

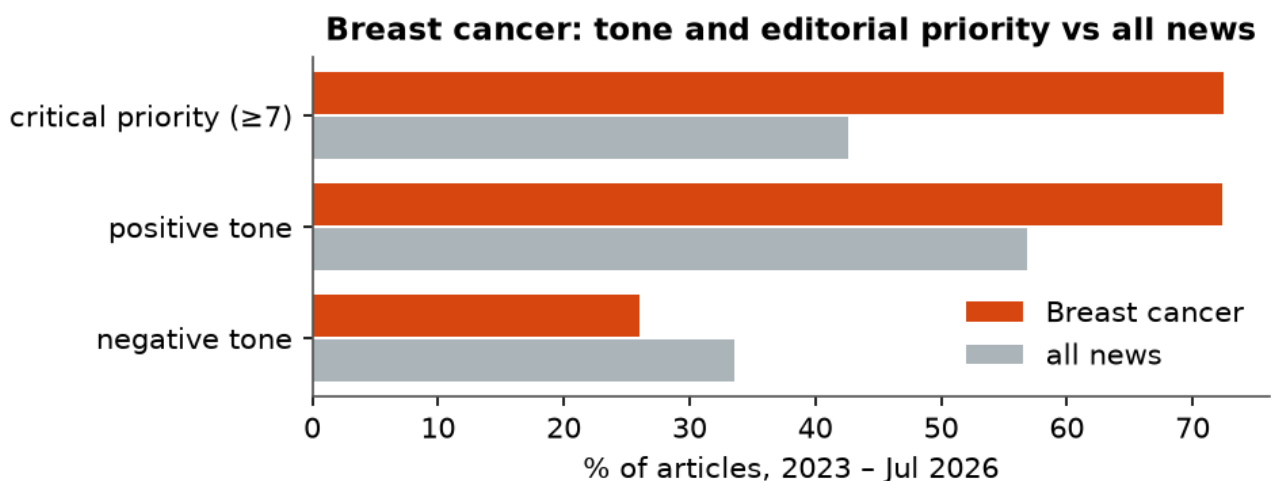


Figure 7. Tone and the share rated critical priority, breast cancer articles vs all news, 2023 – Jul 2026.

Themes show what breast cancer stories are about (Table 5). Treatment options are the most common (53.94% of themed articles), followed by hormone therapy (38.13%) and survivorship (26.7%). Screening guidelines appear in 13.29%, with the highest lift of any theme (1025.13). Celebrity health updates appear in 9.23%. A social-media platform is the organisation named most often (2,440 articles), usually because a public figure announced a diagnosis there.

Theme	Articles	Share %	Lift
breast cancer screening guidelines	4,050	13.29	1025.13
breast cancer treatment options	16,442	53.94	422.36
hormone therapy for breast cancer	11,622	38.13	420.37
genetic counseling for hereditary cancers	4,166	13.67	295.69
cancer survivorship lifestyle adjustments	8,139	26.7	134.11
general cancer prevention strategies	4,972	16.31	118.74
psychosocial support in cancer care	5,486	18.0	118.72
celebrity health and wellness updates	2,813	9.23	38.97

Table 5. Themes of breast cancer coverage with at least 2,000 articles, ranked by lift against all themed news, January 2023 – July 2026.

The other organisations named most often are cancer charities, research institutes and drug makers: the American Cancer Society (1,701 articles), the National Cancer Institute (831), AstraZeneca (557), Susan G. Komen for the Cure (487) and the Mayo Clinic (429).

3.5 Research attention and invention

Research on breast cancer grew faster than the scientific literature. Papers with breast cancer terms in the title rose from 25.27 per 10,000 papers in 2015 to 41.37 in 2025, a 1.64-fold increase in share, while the total literature in the corpus stayed about the same size (a factor of 1.0) over the same years. Between 2015 and 2025, 233,244 such papers were published. The most-cited include the molecular classification of breast tumours, global burden statistics and trials of new treatments for triple-negative and BRCA-associated disease [1-7].

News attention does not track research attention (Figure 8). Research share rose every year from 2015 to 2025, while news share fell in 2020 and has stayed below its 2019 level since.

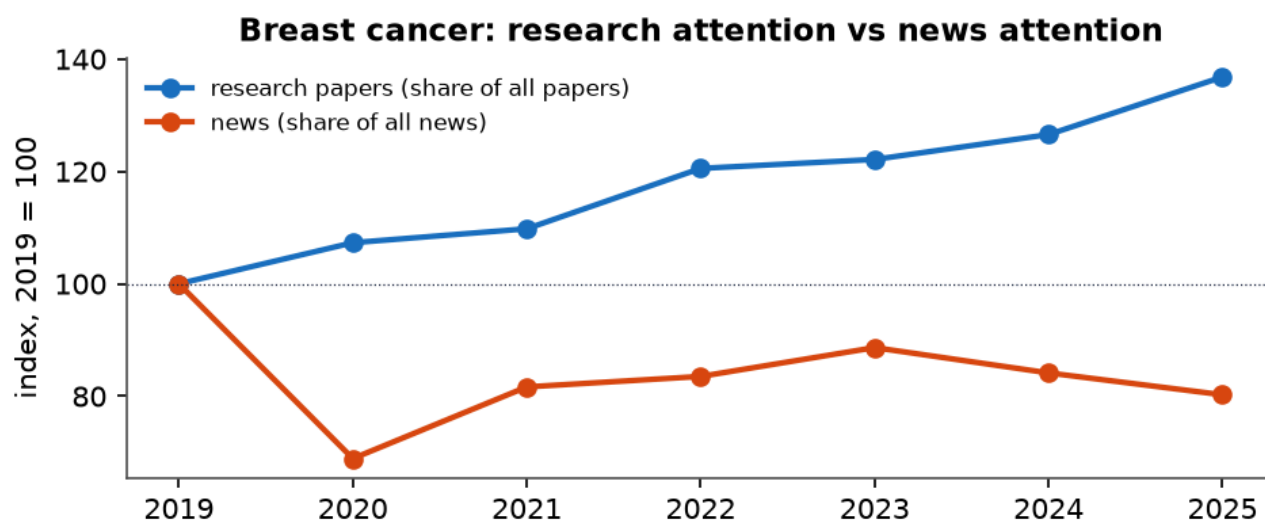


Figure 8. Research papers with breast cancer in the title (share of all papers) and breast cancer news (share of all news), each indexed to 2019 = 100.

Patenting rose too: 576 breast cancer patent families were published in 2015 and 988 in 2024 (1.72 times), with most of the growth by 2021 (1,023 families). China's office holds 53.3% of office filings of these families from 2015 to 2024, followed by the United States (11.7%), international applications (8.5%) and the European office (7.3%) (Figure 9). The leading technology classes are drug preparations (A61K, 1,076 families), diagnostic assays (C12Q, 809), diagnostic devices (A61B, 515) and analytical methods (G01N, 484).

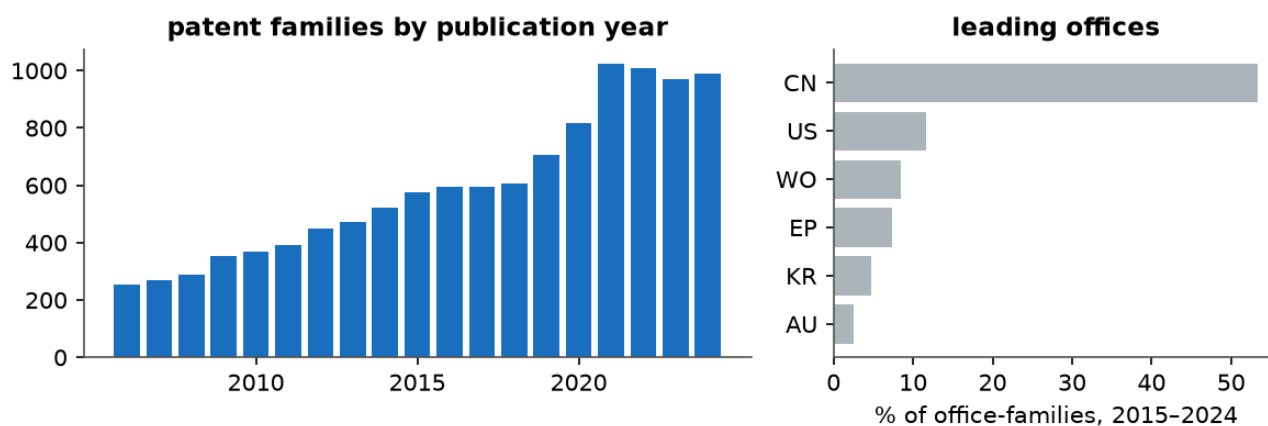


Figure 9. Breast cancer patent families by publication year (left) and the leading filing offices, 2015-2024 (right).

4. Discussion

Three points matter for anyone who uses news as a record of breast cancer. First, coverage is shaped by the calendar more than by events. Nearly a quarter of each year's articles appear in October, and no other month comes close. Any count of breast cancer news needs to compare like months, or it will measure the awareness campaign rather than the disease. Second, raw shares mislead. The United States leads in raw terms, but the United Kingdom is more over-represented, and Ghana most of all. India, where breast cancer is the most common cancer among women, is strongly under-represented in the index's breast cancer news, though its share is rising. Third, a large share of what keyword searches return as "breast cancer news" is not journalism. Research-paper abstracts made up most of the raw matches, and a tenth of what remained after filtering.

The tone of coverage is a finding in itself. Breast cancer news is more positive than news overall, while also being rated urgent. This matches the survivor-centred framing of breast cancer culture, in which diagnosis stories are told as stories of recovery, fundraising and early detection.

Research, invention and news run on different clocks. Research share grows steadily. Patent publications rose until 2021 and then levelled. News follows the October campaign and the diagnoses of public figures. The dataset lets researchers link the three, for example to test whether screening guideline changes are followed by shifts in coverage in the countries concerned.

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 if the missing part is spread evenly across topics.
- **Coverage after May 2026.** The index thins after May 2026, so 2026 (January-July) is partial and is excluded from year comparisons.
- **Matching.** Matching uses the article opening. Articles that mention breast cancer only later are missed. The abstract filters also remove some genuine news that is written like a study report.
- **Research abstracts.** About 10% of matched items are research abstracts that the filters missed. They carry no location, so geography is unaffected, but they add to the monthly series and the tone shares.

- **Machine labels.** Location, tone, priority and themes are machine-labelled. The city field is noisy and is not analysed.
- **Place tags from ordinary words.** Place tags derived from ordinary words ("Congress", "Winter", "Mobile" and similar) were treated as unlocated, which removed 15.3% of city-tagged articles; located counts are therefore smaller, and some genuine places that share a name with such a word are lost.
- **Outlet mix.** The index over-represents English-language outlets. This shapes both seasonality and geography, and may explain part of the low lifts for India and China.
- **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 19,424 of 78,107 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 "breast cancer" --since 2019-01-01 --until 2026-07-31 --granularity monthly`.

References

1. The Cancer Genome Atlas Network (2012). Comprehensive molecular portraits of human breast tumours. *Nature*. <https://doi.org/10.1038/nature11412>
2. Carol DeSantis et al. (2019). Breast cancer statistics, 2019. *CA A Cancer Journal for Clinicians*. <https://doi.org/10.3322/caac.21583>
3. Melina Arnold et al. (2022). Current and future burden of breast cancer: Global statistics for 2020 and 2040. *The Breast*. <https://doi.org/10.1016/j.breast.2022.08.010>
4. Adrienne G. Waks et al. (2019). Breast Cancer Treatment. *JAMA*. <https://doi.org/10.1001/jama.2018.19323>
5. Peter Schmid et al. (2020). Pembrolizumab for Early Triple-Negative Breast Cancer. *New England Journal of Medicine*. <https://doi.org/10.1056/nejmoa1910549>
6. Mark E. Robson et al. (2017). Olaparib for Metastatic Breast Cancer in Patients with a Germline BRCA Mutation. *New England Journal of Medicine*. <https://doi.org/10.1056/nejmoa1706450>
7. Babak Ehteshami Bejnordi et al. (2017). Diagnostic Assessment of Deep Learning Algorithms for Detection of Lymph Node Metastases in Women With Breast Cancer. *JAMA*. <https://doi.org/10.1001/jama.2017.14585>
8. Fliar BI (2026). *Breast Cancer in the News 2019–2026: article-level dataset and coverage tables*. Prism News API. <https://github.com/fliarbi/prism-news-api>