



# Impactful COVID-19 discoveries from China are neglected in the media

Yinxian Zhang<sup>1</sup>

Received: 24 October 2022 / Accepted: 14 June 2023 / Published online: 3 July 2023  
© The Author(s) 2023

## Abstract

Do the media and the scientific community often agree on what research is worth public attention? During the COVID-19 pandemic, I analyzed all COVID-related research articles ( $n=665$ ) published in 13 world-leading medical and science journals. I found that the media attention an article receives is positively associated with its scientific impact—proxied by citation counts. However, given the same level of scientific impact, an article from China, on average, would receive 5.25 times fewer media mentions per additional citation, compared to an article from other parts of the world. The articles being neglected feature timely and impactful discoveries including asymptomatic carriers of SARS-CoV-2, and the origins and transmission dynamics of the virus. The underrepresentation of Chinese articles persisted even after controlling for various factors, including the research team’s pre-existing media sources and self-country citations, i.e., cites coming from the same country. Such media bias may come at the cost of global public health.

**Keywords** Citations · Media mentions · Altmetric · Media bias · Science of science

## Introduction

Social scientists have established the fact that the medialization of science, i.e., the coupling between science and media, bears profound social and political implications (Peters, 2013; Weingart, 1998). Media coverage of scientific discoveries shapes public understanding of and engagement with science, and is relevant for policymaking in related fields (Tan & Weaver, 2007). The importance of media in scientific communication is even more pronounced during a public health crisis, such as the COVID-19 pandemic. For instance, certain media outlets are politically motivated, and their reporting may adversely impact the public’s attitudes toward and behavioral responses to the disease (Loomba et al., 2021; Zhao et al., 2020). Distorted information in the media is also detrimental for international collaboration during the outbreak, as it hinders infection mitigations and undermines a coordinated effort to fight the virus (Meinhof, 2020; Zhang & Xu, 2020).

---

✉ Yinxian Zhang  
yinxian.zhang@qc.cuny.edu

<sup>1</sup> CUNY-Queens College, Queens, NY, USA

Yet, the problem of media bias not only stems from inaccurate information, but also occurs when important information is downplayed or neglected in the media, making it hard for people to access crucial messages. This is what the World Health Organization (WHO) calls an infodemic (Watts et al., 2021).

Examples of the media overlooking critical scientific findings started at the very beginning of the COVID-19 pandemic. As early as on January 24, 2020, a group of scientists reported findings about person-to-person transmission of SARS-CoV-2 in family settings (Chan et al., 2020). Soon in February 2020, another impactful article warned about asymptomatic carriers of the virus (Bai et al., 2020). These discoveries emphasized the importance of early quarantine, contact-tracing and mask wearing in curbing virus transmission.

The two articles were published even before the Chinese government acknowledged the outbreak of COVID-19. At the time, most people did not understand the nature of the disease or how to best protect themselves. The articles deviated from the Chinese official rhetoric and provided timely clarification and guidance on disease mitigation. They were published in world-leading medical journals, *the Lancet* and the *Journal of the American Medical Association (JAMA)*, and have drawn wide attention within the scientific community.

However, these articles attracted little attention from mainstream media<sup>1</sup> and the public at the time—the media mentions they received are about 300 times and 9 times less than that of other COVID-related articles in top medical and science journals with the same level of citations. Eventually, the United States (U.S.) missed the best time window for quarantine and contact tracing. Worse, mask wearing was not mandated until two to three months later. Such a delayed response was blamed for the ineffective COVID-19 mitigation in the U.S. (PBS, 2020).

Such examples cast doubts on the media's discretion in determining what scientific research is newsworthy. Ideally, the media should present the latest and most consequential scientific findings endorsed by the scientific community. Some studies indeed found a positive correlation between an article's media exposure and scientific impact (typically measured by its citation counts; see, e.g., Chapa et al., 2017; Costas et al., 2015). Nevertheless, the media is also influenced by market competition, the logic of profit-maximizing, and even personal interests and biases of editors and journalists (Baron, 2006; Gentzkow & Sharpio, 2006). As a result, media outlets may cover low-quality research or promote certain hot topics over other discoveries in exchange for eye-catching stories or promoting specific agendas (Althaus et al., 2021; Selvaraj et al., 2014).

In the context of the COVID-19 pandemic, scientific reporting is also influenced by overwhelming public sentiments and geopolitical tensions, as the disease has become a politically and ideologically charged issue that divides the world. The two articles in the *Lancet* and *JAMA* both happen to be authored by scientists from Chinese institutions. Note that these discoveries were independent from Chinese official rhetoric—in fact, they were even not in line with the political interests of the Chinese government who focused on damage control at the moment. Nonetheless, these scientific findings were largely overlooked in the media.

The aforementioned articles are not exceptions. As the pandemic unfolds, many researchers have restlessly investigated the origins and transmission dynamics of the virus. A *Lancet* article published in February 2020 finds that bats may be the original host of the virus (Lu et al., 2020), and a *Nature* article successfully isolated

<sup>1</sup> Unless specified, “the media” in this paper primarily refers to the Western media that dominates the global media market.

SARS-CoV-2-related coronavirus from pangolins, providing evidence to the claim that pangolins might be the intermediate hosts (Lam et al., 2020). Both articles were authored by Chinese scientists and, again, the media attention received by these articles are about 400 and 5 times less than other articles with the same level of scientific impact. In the meantime, lab leak theories have been heatedly discussed.

Even on less controversial topics that are of great interest to the public, impactful articles authored by Chinese scientists appear to be neglected. A *JAMA* article discusses the possibility of vertical transmission of the virus from an infected mother to her newborn (Zeng et al., 2020); an *Annals of Internal Medicine* (AIM) article warns about the risks of fecal aerosol transmission in high rise apartments (Kang et al., 2020). Yet they were not picked up by the media despite that these findings bear important implications for our day-to-day life in the pandemic era.

To sum, while the world decries the Chinese government for its mishandling of the COVID outbreak, frontline discoveries from China appear to be collateral damage. Even though they are highly impactful within the scientific community and are independent of the Chinese government's agenda, they do not draw much attention from the media and the public.

The present study asks whether COVID-19 discoveries from certain countries are systematically underrepresented, compared to these from other places of the world. Given the boiling public sentiments and the geopolitics between the U.S. and China, the media is predicted to downplay research findings authored by scientists from Chinese institutions, unintentionally or not.

More generally, this study investigates media coverage of scientific reporting during COVID-19 with a focus on the role of an article's country affiliation. Country affiliation is an understudied factor in the existing literature on media bias, but it became salient during the pandemic when geopolitical tensions arose. Knowing that there is intense international competition to study the virus and develop vaccines, the media might consciously or unconsciously promote more research from some countries than from others. The article aims to answer two questions: What is the relationship between an article's media attention and its scientific impact? How is this relationship moderated by the article's country affiliation?

I confined the scope of this study to the media coverage of only COVID-related articles published in 2020. Studying country affiliation in the media coverage of science is challenging partly because countries vary in their research focus. Some countries are advantaged by concentrating on the disciplines or topics that are inherently more newsworthy, such as climate change or astrophysics. By contrast, COVID-19 is a topic with global interest, and the best talents from around the world are focusing on the same topic simultaneously. Therefore, confining the scope to COVID-19 studies addresses the issue of the variations in topic preferences across countries.

## Results

### Scientific impact and media attention: a positive relationship, but less positive for some countries

To answer the research questions, this study investigated media reporting of 665 COVID-related research articles published in 13 top general science or medical journals

in 2020, such as *Science*, *Nature*, *JAMA*, and *The Lancet*. These articles were written by research teams from 33 countries/regions. The country/region of each article was determined by the corresponding authors' institution affiliations (see [Materials and Methods](#) for more details). I limited the sample to the 13 top journal publications, as those journals occupy the vast majority of media and scientific attention.

To study potential media bias, I examined how the relationship between articles' media attention and their scientific impact would be moderated by the articles' country affiliations. The main outcome variable was an article's media attention, measured by the number of mentions that an article received from media outlets globally (Mean = 127.69, Median = 49, SD = 233.16); the main independent variable was an article's scientific impact, proxied by its citation counts, which represent attention from within the scientific community (Mean = 272.03, Median = 91, SD = 694.33).

I reckoned that articles published earlier had more time to attract media mentions or citations, and therefore would score higher. To control for the time effect, I defined a 6-month window starting from the publication date, and counted only media mentions and citations that occur within this time frame for all the articles. Since the outcome variable was an over-dispersed count variable, with its variance being proportional to the square of its mean (variance = 3.33 mean-squared), I adopted negative binomial models throughout the analyses.

I first ran a regression and confirmed a positive relationship between media attention and scientific impact. Specifically, an increase in one citation count, on average, is associated with 0.12 additional media mentions ( $n = 665$ ,  $z = 16.88$ ,  $p < 0.001$ ).

I then tested the hypothesis that the relationship between an article's media attention and its scientific impact is moderated by its country affiliation. Among all countries/regions, only the U.S. ( $n = 308$ ), China ( $n = 102$ ), and the United Kingdom (UK,  $n = 74$ ) had more than 30 articles in the data—the minimally adequate sample size for a moderation analysis. Among these three countries, I discovered that compared with articles from the U.S. ( $z = 10.46$ ,  $p < 0.001$ ) and the UK ( $z = 4.83$ ,  $p < 0.001$ ), articles from China attracted less media attention given the same level of scientific impact. A typical COVID-related article, which received 92.5 citations in six months (median of the three-country sample), tended to gather 0.04 [95% CI 0.03, 0.05] additional media mentions for every additional citation if it was written by a research team from China. By contrast, the average marginal increase of media mention was 0.20 [95% CI 0.17, 0.24] for articles from the U.S., and 0.28 [95% CI 0.18, 0.38] for articles from the UK. I refer to this phenomenon as the “attention gap” between different countries. By contrast, articles from the U.S. and the UK did not differ in media attention for the same level of scientific impact ( $z = 0.11$ ,  $p = 0.91$ ).

## Preexisting media resources do not explain the attention gap

One possible confounder in the previous analysis was preexisting media resources. Research teams from certain countries may have limited relations with news outlets that dominate the global media scene. Hence, articles from China could fall outside of journalists' radar despite having been published in the same prestigious journals as other countries.

To account for research teams' unequal access to media resources, I computed two metrics to estimate the preexisting media resources available to them. The first metric gauges

the *corresponding authors'* maximum media resources. I collected all of the corresponding authors' publications between 2015 and 2019, the previous five years before their COVID-19 publications came out. I then located the most popular publication of the corresponding authors and used its media mention count to proxy the maximum media resources that the research team had access to.

The second metric gauges the *institutions'* maximum media resources. For each corresponding author's affiliated institutions, I located the most popular publication of each institution in the same 13 journals between 2015 and 2019. The media mention count of that article was used to proxy the maximum media resource that the institution could provide to the research team.

After the access to media resources were controlled for, the moderation effect of the country affiliation remained unchanged. Compared with articles from the U.S. ( $z = 10.83$ ,  $p < 0.001$ ) and the UK ( $z = 5.02$ ,  $p < 0.001$ ), articles from China still attracted less media attention given the same level of scientific impact. While a typical COVID-related article from China tended to gather 0.04 [95% CI 0.03, 0.05] additional media mentions for every additional citation, the average marginal increase of media mention was 0.19 [95% CI 0.16, 0.23] for articles from the U.S., and 0.28 [95% CI 0.18, 0.37] for articles from the UK. I continued to include the two metrics of media resources as covariates for the rest of the regression analyses in the article.

### Behind the attention gap: are countries underrepresented or over-cited?

Another factor that may explain the media attention gap comes from the other side of the relationship—some may argue that articles from certain countries are probably not *under-represented* in the media, but rather *over-cited* by other researchers from the same country. In other words, the relationship between citation counts and media mentions is attenuated for certain countries due to their research practice.

Along this line, Chinese research teams could disproportionately cite articles produced by other Chinese teams, leading to an inflation of citation counts. It is possible that the media does not ignore scientific discoveries from China. Rather, they reflect the true level of impact of these discoveries.

To control for the effect of “self-country citations”, I excluded citations from the same countries for each article. Using the *lean citation counts* as an alternative measure of scientific impact, the model indicates that the moderation effect of the country affiliation remained the same. Again, compared with articles from the U.S. ( $z = 10.75$ ,  $p < 0.001$ ) and the UK ( $z = 4.88$ ,  $p < 0.001$ ), articles from China attracted less media attention given the same level of *lean citations counts* without self-country cites. The average marginal increase of media mention was 0.05 [95% CI 0.03, 0.06] for a typical COVID-related article from China, and the average marginal increase was 0.25 [95% CI 0.21, 0.30] for articles from the U.S., and 0.30 [95% CI 0.20, 0.40] for articles from the UK. Preexisting media resources were controlled for.

Since self-country citations do not confound the relationship between media attention and scientific impact, I continued to use original citation counts for the rest of the main analysis. However, I included the alternative measure of lean citation counts (excluded self-country cites) in multiple alternative specifications in robustness checks (see Fig. 3).

## The attention gap pertains uniquely to articles from China during COVID

After considering the role of media resources, I proceeded to explore if this phenomenon pertains uniquely to articles from China. Do articles from other countries also receive media attention disproportionate to their scientific impact, compared with the U.S. or UK?

To this end, I combined articles published by other countries into Asia ( $n=32$ ) and Europe ( $n=112$ ) to form two additional comparison groups (see [Materials and Methods](#) for the specific countries). I then performed another moderation analysis contrasting the relationship between citation counts and media mentions among all five countries/regions. Access to media resources were again controlled for.

In addition to the U.S. and the UK, both Asian ( $z=3.50$ ,  $p<0.001$ ) and European articles ( $z=5.87$ ,  $p<0.001$ ) attracted more media mentions than those from China. Given the same level of scientific impact and after controlling for media resources, the average marginal increase of media mention was 0.04 [95% CI 0.03, 0.05] for a typical COVID-related article from China. By contrast, the average marginal increase of media mention was 0.20 [95% CI 0.10, 0.30] or 5 times that of China for articles from Asia, 0.14 [95% CI 0.10, 0.18] or 3.5 times that of China for articles from Europe, 0.19 [95% CI 0.16, 0.23] or 4.75 times that of China for articles from the U.S., and finally, 0.27 [95% CI 0.18, 0.37] or 6.75 times that of China for articles from the UK. Figure 1a plots the actual citation counts (x-axis) against media mentions (y-axis) by countries/regions, as well as predictions based on the regression model.

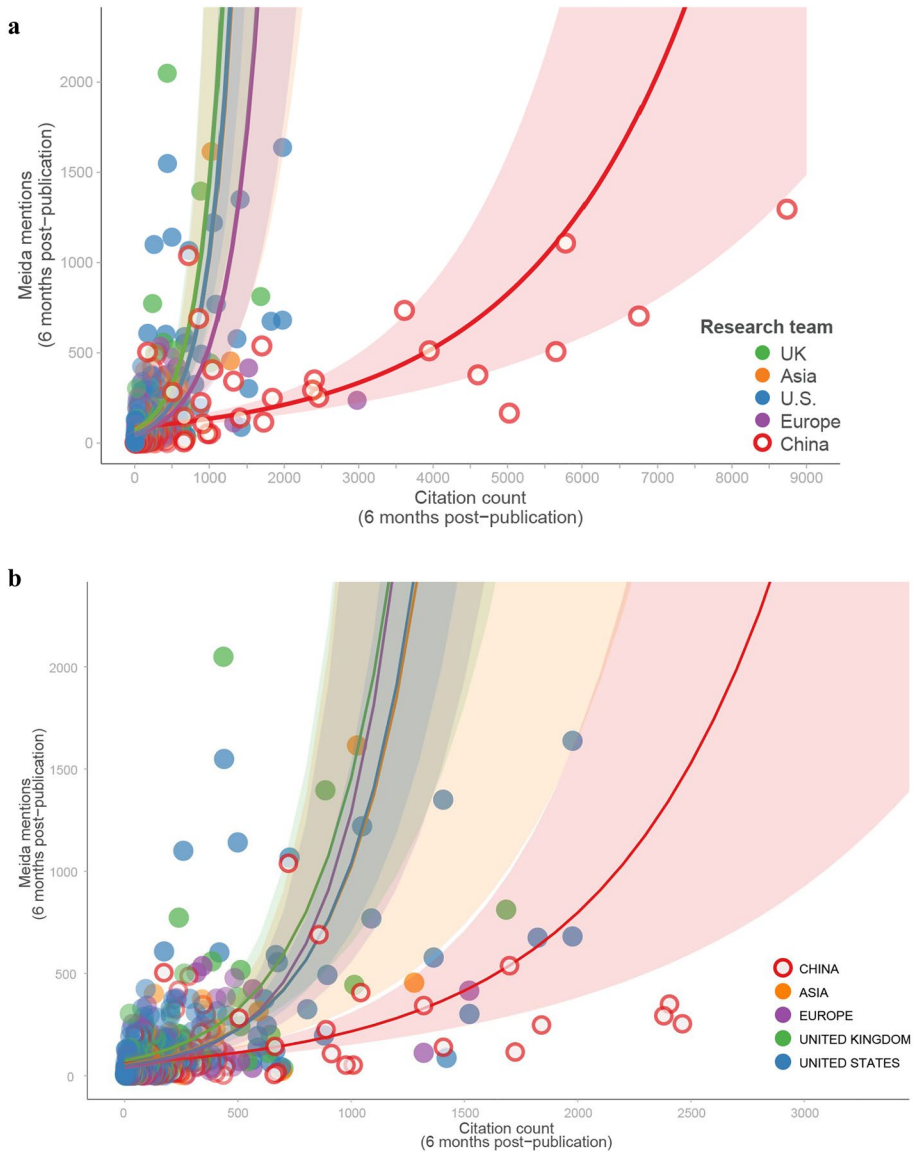
One may note that there appears to be some extreme outliers in Fig. 1a in terms of citation counts, which is also an over-dispersed count variable that follows a negative binomial distribution. Yet, a formal test showed no significant evidence for the presence of outliers in the model (see the Supplement Information (SI) for a detailed discussion). Moreover, removing the seemingly outlying data points ( $\geq 2500$  cites) does not alter the conclusions, even though the error margins of the predictions did increase. Figure 1b re-plots the relationship between citation counts and media mentions after removing the outliers. After confirming that outliers do not impact the analysis, I proceeded with the standard model.

Figure 2 displays the marginal increases in media mentions for every additional citation (i.e., marginal effects, y-axis) for articles at different levels of scientific impact (x-axis). Generally, the marginal effect was larger for articles with higher scientific impact. Yet, articles from China displayed a smaller marginal effect than articles from any other countries/regions.

Notably, even articles from seven Asian countries/regions—including Hong Kong, a special administrative region of China—did not differ from the U.S., the UK, or Europe in terms of the relationship between citation counts and media mentions (all  $ps > 0.40$ ).

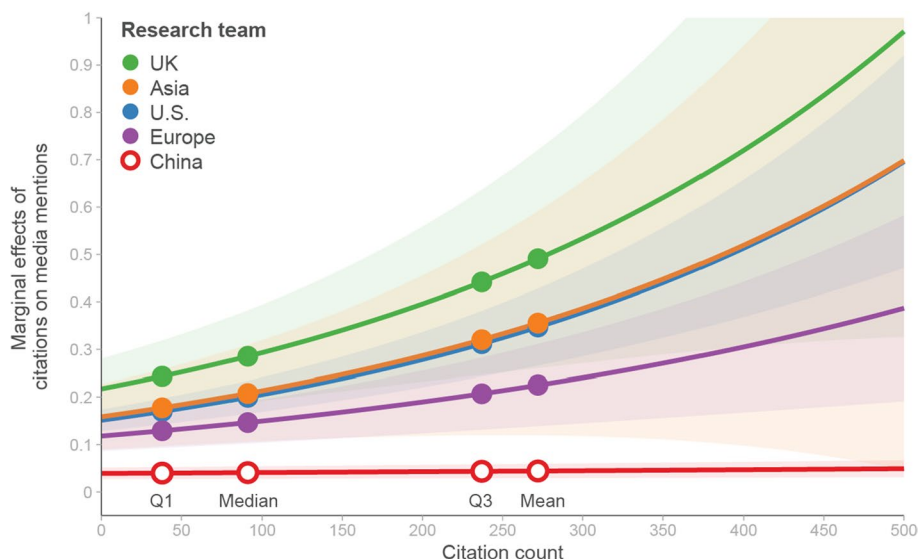
Apart from the five main countries and regions, the sample also includes 37 articles from other countries and regions that cannot be categorized into a coherent group—I labeled them as “Others”. Due to the significant disparities within this group, comparing China with “Others” may be meaningless. Regardless, including “Others” in the model did not change the conclusions. Figure S1 in the SI shows the citation counts against media mentions by countries/regions, including the “Others” group. A list of specific countries in this group can also be found in the SI.

Finally, I went on to do a binary comparison between articles from China and those from all other parts of the world combined, including the “Others” group. A typical COVID-related article, which received 91 citations in six months (median of the entire sample,  $n=665$ ), tended to gather 0.04 [95% CI 0.03, 0.05] additional media mentions



**Fig. 1** **a** Citation counts and media mentions for COVID-related research articles published in 2020, by the research team's country/region. An article's media attention was generally positively related to its scientific impact proxied by citation counts. However, compared to other countries/regions, articles from China attracted less media attention when they had the same level of scientific impact. Uneven access to media resources among the countries/regions was taken into account. **b** Citation counts and media mentions for COVID-related research articles published in 2020, by the research team's country/region, outliers with greater than 2500 cites removed. Compared to other countries/regions, articles from China still attracted less media attention given the same level of scientific impact and after controlling for uneven access to media resources





**Fig. 2** Increase in media mentions for every additional citation (y-axis) for COVID-related research articles. The effects are plotted for different levels of scientific impact (x-axis) and by the research team's country/region. Different levels of citation counts are highlighted in circles. For example, for a typical article from the UK with a median level of citation counts and media resources, an increase in one citation was associated with 0.27 [95% CI 0.18, 0.37] additional media mentions. For some countries/regions, the confidence intervals of the marginal effects grow rapidly beyond 500 citation counts. Hence, I restricted the x-axis to 500 (equivalent to the 90th percentile) for a clearer view

for every additional citation if it was written by a research team from China. By contrast, the average marginal increase of media mention was 0.21 [95% CI 0.18, 0.23] for articles from all other countries in the dataset, almost 5.25 times that of China. Preexisting media resources were controlled for.

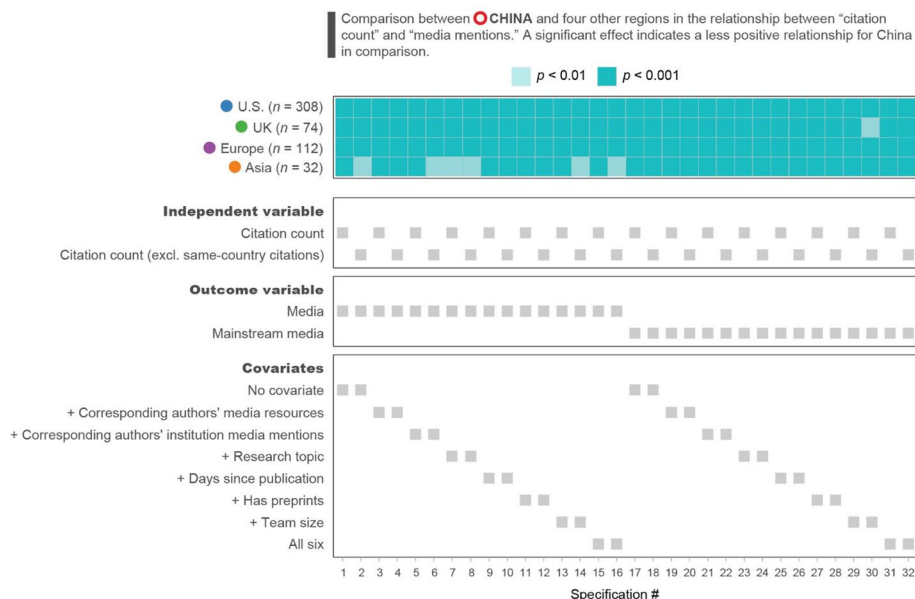
In conclusion, within the scope of the present study, China was the only country whose articles attracted fewer media mentions than those from other countries/regions, given the same level of scientific impact and after controlling for media resources.

### Testing the effect of author ethnicity

So far, an article's country affiliation was determined by the corresponding authors' main institution location. I tested an alternative approach in which affiliation was determined by the corresponding authors' ethnicity. This specification allowed us to consider scientists of mainland Chinese origin but working in non-Chinese institutions. In addition, I made a distinction between mainland and non-mainland Chinese (e.g., authors from Hong Kong and Taiwan).

To determine the ethnic composition of each research team, I manually coded the names of *corresponding authors*. Using names, especially surnames, to infer racial/ethnic groups has been a standard approach for race/ethnicity imputation when self-reported information is not available (Imai et al., 2022; Labgold et al., 2021). This approach is particularly popular in the fields of population studies, public health, and anthropology. A rich literature has verified the validity of this approach, demonstrating that surname analyses perform very





**Fig. 3** Multiple specifications of the model confirm that country affiliation significantly moderated the relationship between citation counts and media mentions. Articles from China received fewer media mentions than other countries/regions, given the same level of citation counts. The upper panel displays the comparisons between articles from China versus the U.S., the UK, Asia, or Europe respectively. The lower panel displays 32 specifications of the model with different combinations of independent variables, outcome variables and covariates. All comparisons were significant

well when compared against self-identified ethnicity (Elliott et al., 2009; Fiscella & Fremont, 2006; Mateos, 2007).

The present study focuses on separating other ethnic groups from Chinese and East Asian ethnicities, including non-mainland Chinese, Japanese, and Korean. Several studies have developed and validated surname lists to identify Chinese and other Asian ethnicities (Lauderdale & Kestenbaum, 2000; Quan et al., 2006; Shah et al., 2010). These surname lists have been proven to achieve high reliability and accuracy in predicting Chinese and Asian ethnic identities. Additionally, it is possible to differentiate between mainland and non-mainland Chinese names because the romanization system used in mainland China (*Hanyu Pinyin*) is different from the Chinese language romanization systems in other areas (such as *Standard Romanization* in Hong Kong and *Tongyong Pinyin* in Taiwan). Therefore, distinguishing between name spellings is effective in identifying mainland Chinese names, especially when both first and last names are used (Nicoll et al., 1986). Using the surname lists and the authors' local knowledge of Chinese and East Asian names, I manually coded the corresponding authors' names and categorized the articles into four groups:

(A) full mainland Chinese origin—all corresponding authors are of mainland Chinese origin ( $n = 101$ ); (B) partial mainland Chinese origin—multiple corresponding authors with varying origins that include mainland China ( $n = 23$ ); (C) full East Asian origins—all authors are of East Asian origins, i.e., non-mainland Chinese, Japanese, and Korean origins ( $n = 30$ ); and (D) other origins—all authors are of origins of other parts of the world ( $n = 511$ ).

The results showed that articles by authors of full mainland Chinese origin attracted less media attention given the same level of scientific impact, compared with articles by authors of full East Asian origins ( $z=4.20$ ,  $p<0.001$ ), partial mainland Chinese origin ( $z=2.64$ ,  $p=0.008$ ), or other origins ( $z=14.10$ ,  $p<0.001$ ). The average marginal increase of media mention was 0.04 [95% CI 0.03, 0.05] for a typical COVID-related article fully authored by mainland Chinese researchers, 0.08 [95% CI 0.04, 0.13] for articles partially authored by mainland Chinese researchers, 0.15 [95% CI 0.08, 0.22] for articles fully authored by other East Asian researchers, and 0.21 [95% CI 0.18, 0.24] for articles fully authored by researchers of other origins. Preexisting media resources were again taken into account. Figure S2 plots predictions based on the moderation analysis.

In addition, articles by authors of partial mainland Chinese origin also attracted significantly less media attention compared to articles by authors of other origins ( $z=5.32$ ,  $p<0.001$ ), but less so when compared to articles by authors of full East Asian origins ( $z=2.07$ ,  $p=0.04$ ). Lastly, articles by authors of full East Asian origins did not differ from articles by authors of other origins ( $z=1.46$ ,  $p=0.14$ ) in media attention given the same level of scientific impact. These results confirmed that the phenomenon pertains distinctly to articles written by mainland Chinese scientists. Scientists of other East Asian origins were largely unaffected.

## Robustness checks

Finally, I considered multiple alternative specifications of the main models to test the robustness of the results. For the independent variable, I excluded self-country citations for each article when calculating citation counts. For the outcome variable, I tested the model with mainstream media mentions instead of all media mentions. For the covariates, I included four variables: research topic, novelty effect, presence of preprints, and team size, in addition to the research team's media resources. The rationales for considering these factors are listed below. The results of 32 model specifications are summarized in Fig. 3. All alternative specifications of the main models supported the conclusion that, within the scope of this study, articles from China received fewer media mentions than those from all other countries and origins, given the same level of scientific impact.

Below, I describe the factors I explored that could have potentially altered the conclusion.

1. *Self-country citations*: As discussed above, for each article in the dataset, I excluded citations from the same country as the corresponding authors. I used this *lean citation count* as an alternative independent variable in 16 of the 32 specifications. All results remain the same and it indicated that self-country citations do not explain the attention gap between China and other regions.
2. *Mainstream media versus all media mentions*: The majority of media mentions collected by *Altmetric.com* were not from mainstream media outlets such as CNN or the BBC. I further investigated whether the attention gap was only present in less prominent outlets with limited global impact. Results showed that the attention gap pertaining to articles from China was also present among mainstream media outlets (see [Materials and Methods](#) for the classification of media outlets).
3. *Research topic*: various topics in COVID-related research attract media and scientific attention to different degrees. For example, clinical case reports may not be as news-

worthy to the media as they are valuable to peer scientists. China—as the initial virus hub—produced more clinical case reports than other countries in 2020, and thus is expected to receive less media attention. To control for the role of topic variance, I used the topic classification of COVID-19 articles provided by the *CoronaCentral* database (version April 2, 2023; Lever & Altman, 2021) and sorted the 665 articles into eight research topics: biology, epidemiology, clinical, treatment, policy, psychology, others, or multi-topic (refer to [Materials and Methods](#) for more details). I entered the topics as dummy variables in the regressions. The main moderation effect of country affiliations remained unchanged after controlling for topic variances across countries.

4. *Novelty effect (days since publication)*: Although I controlled for the time effect by defining a time frame for media mentions and citations, later COVID-related publications might have still received less media attention as the public grew accustomed to the pandemic. I added a “days-since-publication” variable, i.e., the number of days between publication date and the end date of my data collection, to control for the novelty effect. The moderating effect of country affiliation remained the same even though earlier publications did indeed receive more media mentions.
5. *Presence of preprints*: Many scientists openly post their article online (i.e., a preprint) before its eventual publication in a journal. The media has an opportunity to immediately report on these findings, and thus may not report again when the same research is later published in a journal. I thus included a dummy variable to control for the presence of preprints in the regression, and all patterns still held the same.
6. *Team size*: Previous research indicates that larger research teams have more media resources and higher impact among peer scientists (Larivière et al., 2015). I therefore added the number of authors on the article as a covariate in the regression. This, again, did not alter the conclusions.

## Discussion

This work investigates the relationship between media attention and scientific impact for COVID-related research published by scientists from different countries/regions. In addition to verifying the positive correlation between articles’ media attention and scientific impact (i.e., citation counts), this study identifies country affiliations as a source of media bias in scientific reporting.

Out of the 665 COVID-related articles examined in this study, only articles from China received significantly less media attention than other articles with the same level of scientific impact. Articles from other countries and regions did not differ in the level of media attention they received. This attention gap occurred in both mainstream and non-mainstream media outlets and applied to articles affiliated with Chinese institutions as well as articles authored by mainland Chinese scientists.

Specifically, articles from China have on average 5.25 times fewer media mentions than those from other places, including other parts of Asia. The gap cannot be explained away by research teams’ unequal access to media resources, self-country citations, or other factors such as topic choices, research team size, research novelty effect, and the presence of preprints.

In the wake of the COVID-19 pandemic, Sinophobia and orientalism have resurged around the globe. The scientific community has noted the disruptive impact of geopolitics

and the chilling effect of multiple punitive measures, including the China Initiative, in stirring up racial discrimination against Asians (Nature, 2021; Lewis, 2021). Multiple scholars have explicitly expressed concerns about the distrust of or bias against Chinese scientists in Western countries (Horton, 2020; Zhang & Xu, 2020). Media bias against research associated with Chinese scientists could be a result of the rising anti-China sentiments. However, I cannot conclude that the media intentionally discriminates against Chinese research due to this dynamic, as I have not explored all alternative explanations. The present study simply reveals that impactful COVID-19 discoveries from China are systematically less represented in the media.

Nonetheless, peer-approved scientific findings about the virus and mitigation strategies from China are generalizable knowledge that can advance public understanding and help other countries to respond to the challenge. Therefore, regardless of the underlying reasons, the underrepresentation of frontline findings from China may come at the cost of global public health.

## Limitations

Due to the limited scope of this study focusing only on COVID-related articles published in 13 top medical and science journals in 2020, I identified China as the only country that is subject to media bias. The conclusion could be different if we examine articles published at later stages of the pandemic. For instance, more recently, South African scientists have accused Western nations of ignoring their discoveries about the Omicron variant, and they claimed that the West does not trust the science because “it came from Africa.” (Harding, 2022). Future study can expand the sample size and/or extend the study to more recent years to track the moderation effect of country affiliations in COVID-19 reporting.

This study is limited to the context of COVID-19 where China was widely viewed as the epicenter and its government’s handling of the virus was heavily scrutinized. The period under study was marked by intense geopolitical tensions between China and the West, as well as a growing negative sentiment toward China in the international community. The underrepresentation of Chinese research in the media could be a side effect of this dynamic during this historic moment. Therefore, the findings may not be applicable in a different context where the power dynamics and geopolitical relationships between countries differ.

However, it is reasonable to assume that geopolitics, racism, and their interplay have a general impact on scientific reporting. The attention gap could occur to other countries or regions if we examine media reporting on scientific advances during other crucial historical moments, such as during the Cold War or after 9/11. Future research may expand the present study to other disciplines or topics in other historical periods to explore the general relationship among media attention, scientific impact, and country affiliation.

Furthermore, *Altmetric* collects news mention data from over 4000 outlets globally. However, due to the dominance of Western media in the global media market, the platform inevitably collects more data from Western sources. As a result, data from local outlets in non-Western regions, including China, is relatively limited.<sup>2</sup> The findings primarily highlight the underrepresentation of Chinese research in Western media outlets. Future research

---

<sup>2</sup> This is not to say that Chinese or other non-Western media outlets were excluded from this study. This study included data from well-known Chinese outlets such as *the China Daily*, *Global Times*, *Xinhua*, *Sina.com*, *163.com*, *CCTV+*, and more.

can consider the locality of media outlets and investigate, for instance, whether Chinese media outlets are biased against non-Chinese articles.

Finally, this study solely focused on scientific reporting in traditional media outlets. I also limited the analysis to the researchers and institutes' traditional media resources. Future work is necessary to determine whether country-based bias also exists on social media platforms and whether unequal access to social media resources, such as different usage of Twitter and Facebook by research teams, may impact the level of media attention they receive.

## Materials and methods

The data on COVID-19 research articles came from (1) *Dimensions COVID-19 publications* (*Dimensions COVID* hereafter; Resources, 2020), an online database containing all publications related to COVID-19; (2) *Dimensions API* (Digital Science, 2018), a large dataset of more than 150 million records of research information; and (3) *Altmetric database* (Altmetric, 2020), which provides alternative metrics data for academic publications.

*Dimensions COVID* was curated using a keyword search and was updated periodically. Information about each publication includes the DOI, title, publication date, journal, author names and affiliations, citation counts, and links to preprints (if any) of each article, among others. I collected additional data from each journal's website to complement article information, including the article type (e.g., empirical research, commentary, etc.), corresponding authors, and author affiliations.

I applied a few filters to obtain the final sample of 665 articles: (1) the publication date was in 2020; (2) the article type is an empirical research article as opposed to a review or opinion piece; and (3) the publication journal is one of the 13 top-tier general science or medical journals that occupy the vast majority of media and scientific attention. The list of top journals was curated based on the 2019 Journal Impact Factor (JIF) List provided by the *Journal Citation Reports of Clarivate Analytics (JCR)*.<sup>3</sup> I included all general medicine (e.g., *The Lancet*) and general science (e.g., *Nature*) journals in the Top 100 journals in the list. Then, I complemented the list with the Top 10 journals from the 2019 JIF List in the General Medicine & Public Health category of JCR. I excluded journals that focus on reviews (e.g., *Nature Reviews Drug Discovery*) since I hoped to investigate original empirical research. See SI for the full list of journals.

## Country affiliation

Relying on information from *Dimensions COVID* and journal websites, I determined an article's country affiliation from the location of the corresponding authors' institution(s). If an article had multiple corresponding authors, or an author is affiliated with institutions of multiple countries, I chose the country shared by most corresponding authors/institutions on the team. There were 12 articles where two or more countries appeared equally often in terms of author affiliations. In such cases, I randomly picked a country from the list. See SI for the full list of countries and regions included in this study.

<sup>3</sup> Refer to an introduction of JIF here: <https://clarivate.com/webofsciencegroup/essays/impact-factor/>

## Citation counts

Using the *Dimensions* API, I retrieved the exact publication date of each citation accumulated by the target articles in the sample. I then calculated the number of citations obtained by each article within six months of its first publication. I also collected information from the *Dimensions* API about the author affiliations of each paper citing the articles in the sample, and determined the country affiliations of these citations using the same method described above.

## Media mentions

Media mentions of the COVID-related academic articles were sourced from the *Altmetric database*. Media mentions were captured by *Altmetric* whenever a news article included an identifier of a research article. Each media mention includes the date, outlet, and title of the media article; the title of the academic publication it mentions; and a URL link to the media article. For each article in the sample, I computed the total number of media mentions that appeared within six months since the article's publication.

## Mainstream media mentions

*Altmetric* captures news mentions from more than 4000 media outlets around the world. I further identified 21 mainstream media outlets and calculated the number of news mentions in these outlets for each article. The metric was used as an alternative outcome variable in the robustness checks. The list of mainstream media outlets was curated based on the news sources asked about in the *American Trends Panel Wave 1 Survey* conducted by Pew Research Center (Mitchell et al., 2014). The survey investigated political polarization and media habits, and included widely known news sources that have large audiences. Pew Research Center included 36 media outlets in their list. I excluded four radio programs (*The Ed Schultz Show*, *The Glenn Beck Program*, *The Rush Limbaugh Show*, and *The Sean Hannity Show*); two television shows (*The Daily Show* and *The Colbert Report*); and six digital sources with niche audiences (*TheBlaze*, *Breitbart*, *Daily Kos*, *Politico*, *Slate*, and *ThinkProgress*). Three other outlets (*The Wall Street Journal*, the *Drudge Report*, and *MSNBC*) had no news mentions of the COVID-related publications in the *Altmetric* database, and were excluded. The final list features 21 media outlets (see SI for the full list).

## Authors' and institutions' preexisting media resources

I collected data on the preexisting media resources of (1) each corresponding author and (2) their affiliated institutions. For author-level media resources, I obtained the corresponding authors' academic publications in the previous five years (2015–2019) from *Dimensions API* and obtained each publication's media mention count from *Altmetric database*. For institution-level media resources, I located all academic publications in the 13 journals that were published in the previous five years and that are affiliated with the same institutions as the COVID-19 publications' corresponding authors are. Then, I obtained the

media mention count of each publication. I used the maximum media mention count as a proxy of the author- and institution-level maximum media resources.

## Author ethnicity

The corresponding author's ethnicity was manually coded based on their names. The articles were classified as follows: (a) full mainland Chinese origin; (b) partial mainland Chinese origin, including the mixture of mainland Chinese and East Asian corresponding authors, and the mixture of corresponding authors from mainland China and all other regions; (c) East Asian origins, including non-mainland Chinese, Japanese and Korean authors; and (d) all other origins. Articles with East Asian corresponding authors and corresponding authors from other origins were counted as "other origins".

## Research topics

Each article was assigned a research topic to control for topic variation across countries. I used the topic classification of COVID-19 publications provided by *CoronaCentral* (Lever & Altman, 2021). As of April 2, 2023, the resource utilized a BERT model to identify 39 topics from more than 442,000 COVID-19 publications. These topics were then sorted into seven categories, including biology, epidemiology, clinical, treatment, policy, psychology, and others. It should be noted that one article could be associated with multiple topics. Out of the 665 articles in the dataset, 620 were included in the *CoronaCentral* database, and they were classified into 33 topics. I used the seven-type categorization to label each article. For articles that belong to multiple categories, I labeled them as "multi-topic." It is worth noting that I did not use field-normalized citation scores because most articles ( $n=505$ ) in the dataset fall within the biomedicine field broadly defined (including topics "clinical", "treatment", "epidemiology", and "biology"), and I assumed that their citation patterns do not vary much across topics.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s11192-023-04779-2>.

## Declarations

**Conflict of interest** The author declares no competing interests.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

Althaus, S. L., Berenbaum, M. R., Jordan, J., & Shalmon, D. A. (2021). No buzz for bees: Media coverage of pollinator decline. *Proceedings of the National Academy of Sciences USA*, 118(2), e200255117.



- Altmetric. (2020). *Altmetric database in JSON format. Unpublished raw data* Retrieved from <https://www.altmetric.com/about-our-data>
- Bai, Y., Yao, L., Wei, T., Tian, F., Jin, D.-Y., Chen, L., & Wang, M. (2020). Presumed asymptomatic carrier transmission of COVID-19. *JAMA*, 323(14), 1406–1407.
- Baron, D. P. (2006). Persistent media bias. *Journal of Public Economics*, 90(1–2), 1–36.
- Chan, J. F. W., Yuan, S., Kok, K. H., To, K. K. W., Chu, H., Yang, J., ... & Yuen, K. Y. (2020). A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: a study of a family cluster. *The Lancet*, 395 (10223): 514–523
- Chapa, J., Haq, Z., & Cifu, A. S. (2017). Comparative analysis of the factors associated with citation and media coverage of clinical research. *Scientometrics*, 112(3), 1271–1283.
- Costas, R., Zahedi, Z., & Wouters, P. (2015). Do “altmetrics” correlate with citations? Extensive comparison of altmetric indicators with citations from a multidisciplinary perspective. *Journal of the Association for Information Science and Technology*, 66(10), 2003–2019.
- Digital Science. (2018). *Dimensions [Software]*. Retrieved from <https://app.dimensions.ai>
- Elliott, M. N., Morrison, P. A., Fremont, A., McCaffrey, D. F., Pantoja, P., & Lurie, N. (2009). Using the census Bureau’s surname list to improve estimates of race/ethnicity and associated disparities. *Health Services and Outcomes Research Methodology*. <https://doi.org/10.1007/s10742-009-0047-1>
- Fiscella, K., & Fremont, A. M. (2006). Use of geocoding and surname analysis to estimate race and ethnicity. *Health Services Research*, 41(4p1), 1482–1500.
- Fiscella, K., & Fremont, A. M. (2014). Political Polarization and Media Habits. *Appendix B. Pew Research Center*. Retrieved December 29, 2021, from <https://www.pewresearch.org/journalism/2014/10/21/appendix-b-the-news-sources/>
- Gentzkow, M., & Shapiro, J. M. (2006). Media bias and reputation. *Journal of Political Economy*, 114(2), 280–316.
- Harding, A. (2022). Was South Africa Ignored over Mild Omicron Evidence? *BBC News*. Retrieved January 2, 2022, from <https://www.bbc.com/news/world-africa-60039138>
- Horton, R. (2020). Offline: COVID-19 and the dangers of Sinophobia. *The Lancet*, 396(10245), 154.
- Imai, K., Olivella, S., & Rosenman, E. T. (2022). Addressing census data problems in race imputation via fully Bayesian Improved Surname Geocoding and name supplements. *Science Advances*, 8(49), ead9824.
- Kang, M., Wei, J., Yuan, J., Guo, J., Zhang, Y., Hang, J., ... & Zhong, N. (2020). Probable evidence of fecal aerosol transmission of SARS-CoV-2 in a high-rise building. *Annals of Internal Medicine*, 173(12): 974–980
- Labgold, K., Hamid, S., Shah, S., Gandhi, N. R., Chamberlain, A., Khan, F., Khan, S., Smith, S., Williams, S., Lash, T. L., & Collin, L. J. (2021). Estimating the unknown greater racial and ethnic disparities in COVID-19 burden after accounting for missing race/ethnicity data. *Epidemiology*, 32(2), 157.
- Lam, T. T. Y., Jia, N., Zhang, Y. W., Shum, M. H. H., Jiang, J. F., Zhu, H. C., ... & Cao, W. C. (2020). Identifying SARS-CoV-2-related coronaviruses in Malayan pangolins. *Nature*, 583(7815): 282–285
- Larivière, V., Gingras, Y., Sugimoto, C. R., & Tsou, A. (2015). Team size matters: Collaboration and scientific impact since 1900. *Journal of the Association for Information Science and Technology*, 66(7), 1323–1332.
- Lauderdale, D. S., & Kestenbaum, B. (2000). Asian American ethnic identification by surname. *Population Research and Policy Review*, 19, 283–300.
- Lever, J., & Altman, R. B. (2021). Analyzing the vast coronavirus literature with CoronaCentral. *Proceedings of the National Academy of Sciences USA*, 118(23), e2100766118.
- Lewis, M. K. (2021). Time to End the U.S. Justice Department’s China Initiative. *Foreign Policy*. Retrieved December 26, 2021, from <https://foreignpolicy.com/2021/07/22/china-initiative-espionage-mistrial-hu/>
- Loomba, S., Figueiredo, A., Poatek, S. J., Graaf, K., & Larson, H. J. (2021). Measuring the impact of COVID-19 vaccine misinformation on vaccination intent in the UK and USA. *Nature Human Behaviour*, 5, 337–348.
- Lu, R., Zhao, X., Li, J., Niu, P., Yang, B., Wu, H., ... & Tan, W. (2020). Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. *The Lancet*, 395(10224): 565–574
- Mateos, P. (2007). A review of name-based ethnicity classification methods and their potential in population studies. *Population, Space and Place*, 13(4), 243–263.
- Meinhof, M. (2020). Othering the Virus. *Discover Society*. Retrieved December 3, 2021, from <https://archive.discoverysociety.org/2020/03/21/othering-the-virus/>
- Nature editorial. (2021). Protect precious scientific collaboration from geopolitics. *Nature*. Retrieved December 26, 2021, from <https://www.nature.com/articles/d41586-021-01386-0>

- Nicoll, A., Bassett, K., & Ulijaszek, S. J. (1986). What's in a name? Accuracy of using surnames and forenames in ascribing Asian ethnic identity in English populations. *Journal of Epidemiology & Community Health*, 40(4), 364–368.
- PBS News Hour. (2020). Leading Republicans, with the exception of Trump, now support mask-wearing. Retrieved December 3, 2021, from <https://www.pbs.org/newshour/politics/republicans-with-exception-of-trump-now-push-mask-wearing>
- Peters, H. P. (2013). Gap between science and media revisited: Scientists as public communicators. *Proceedings of the National Academy of Sciences USA*, 110(3), 14102–14109.
- Quan, H., Wang, F., Schopflocher, D., Norris, C., Galbraith, P. D., Faris, P., Graham, M. M., Knudtson, M. L., & Ghali, W. A. (2006). Development and validation of a surname list to define Chinese ethnicity. *Medical Care*. <https://doi.org/10.1097/01.mlr.0000204010.81331.a9>
- Resources, Dimensions. (2020). “Dimensions COVID-19 publications, datasets and clinical trials. Dimensions.” Dataset. Retrieved from <https://doi.org/10.6084/m9.figshare.11961063.v38>
- Selvaraj, S., Borkar, D. S., & Prasad, V. (2014). Media coverage of medical journals: Do the best articles make the news. *PLoS ONE*, 9(1), e85355.
- Shah, B. R., Chiu, M., Amin, S., Ramani, M., Sadry, S., & Jack, VTu. (2010). Surname lists to identify South Asian and Chinese ethnicity from secondary data in Ontario, Canada: A validation study. *BMC Medical Research Methodology*, 10, 1–8.
- Tan, Y., & Weaver, D. H. (2007). Agenda-setting effects among the media, the public, and congress, 1946–2004. *Journalism & Mass Communication Quarterly*, 84(4), 729–744.
- Watts, D. J., Rothschild, D. M., & Mobius, M. (2021). Measuring the news and its impact on democracy. *Proceedings of the National Academy of Sciences USA*, 118(15), e1912443118.
- Weingart, P. (1998). Science and the media. *Research Policy*, 27(8), 869–879.
- Zeng, H., Xu, C., Fan, J., Tang, Y., Deng, Q., Zhang, W., & Long, X. (2020). Antibodies in infants born to mothers with COVID-19 pneumonia. *JAMA*, 323(18), 1848–1849.
- Zhang, Y., & Xu, F. (2020). Ignorance, orientalism and Sinophobia in knowledge production on COVID-19. *Tijdschrift Voor Economische En Sociale Geografie*, 111(3), 211–223.
- Zhao, E., Wu, Q., Crimmins, E. M., & Ailshire, J. A. (2020). Media trust and infection mitigating behaviours during the COVID-19 pandemic in the USA. *BMJ Global Health*, 5(10), e003323.