Social Science & Medicine 364 (2025) 117542

Contents lists available at ScienceDirect

Social Science & Medicine

journal homepage: www.elsevier.com/locate/socscimed

Media bias in portrayals of mortality risks: Comparison of newspaper coverage to death rates

Calvin Isch *

University of Pennsylvania, Wharton School & Annenberg School for Communication, Philadelphia, PA, 19104, United States

A R T I C L E I N F O

Handling editor: Cecilia Cheng

Keywords:
Mass media
Mortality risks
Linear mixed models
Autoregressive distributed lag models
Risk communication

A B S T R A C T

Chronic diseases are the leading cause of death in the United States, yet effective preventive measures receive minimal healthcare funding. This disparity may stem from public underestimation of these diseases' impact and controllability, with distorted media coverage overemphasizing sensational risks and underemphasizing chronic illnesses. This study compares media coverage of mortality risks to objective measures of death rates to investigate such distortions. Data were collected on 14 mortality risks, including monthly US deaths from CDC Wonder and 823,406 relevant articles from major US newspapers via LexisNexis. Regression analyses and qualitative evaluations using natural language processing tools were performed. From 1999 to 2020, a significant disconnect was found between the deadliest risks and their media coverage. Media coverage fluctuations correlated with death rate changes, yet only 1.7–2.8% of the coverage was explained by these rates. Chronic illnesses were described neutrally as individual challenges, while sensational risks were depicted negatively as collective problems. These results illustrate how the media depict a skewed view of the risks facing the public, with disproportionate coverage of sensational risks while comparatively ignoring chronic diseases. Consumers may consequently come to a distorted understanding of the most threatening risks they face and how to combat them.

1. Introduction

Humans have finite attention and resources to invest in avoiding the many risks that threaten our mortality. Unfortunately, we collectively are far from optimal in allocating those scarce resources. For example, in the United States, chronic diseases lead to 70% of deaths annually, but most of these could be prevented with proper behavior like eating well, exercising, avoiding tobacco and excessive alcohol, and getting regular health screenings (Centers for Disease Control and Prevention, National Center for Health Statistics, 2022). Despite spending $4.3 trillion on healthcare in 2021–18.3% of US GDP (Centers for Medicare & Medicaid Services, 2021)—only a small fraction goes to preventative care (Miller et al., 2008). This lack of funding reflects a considerable mismatch between the value of preventing these diseases and the amount that we currently invest in preventative measures (Hoffman, 2022). What can account for this mismatch? Structural issues such as poverty (Wolfe et al., 2020) and cultural barriers including stigma toward treatments (Byrow et al., 2020) explain some challenges individuals face. Additionally, the misalignment of incentives among key healthcare stakeholders (Feldman, 2020; Branning and Vater, 2016), combined with a health infrastructure that inadequately supports preventive care (Maani and Galea, 2020), contributes to the provision of more expensive and less effective treatments. This creates systemic inefficiencies, reinforcing the current reactive healthcare model. These issues may be compounded by media depictions of mortality risks that inadequately address the scale of preventable chronic illnesses. In this paper, we examine how biases in media representations of mortality risks present a distorted depiction of the most pressing public health threats, potentially influencing public attitudes, resource allocation, and policy decisions.

Many articles assess how media coverage connects to objective measures of mortality risks (Bomlitz and Brezis, 2008; Combs and Slovic, 1979; Frost et al., 1997; Pilar et al., 2020). Generally, these studies reveal a large disconnect, with media coverage over-representing some risks and under-representing others. This discrepancy may impact people's perceptions and behaviors; as evidenced by studies showing a gap between objective risk measures and public beliefs and actions (Brown et al., 2023b; Pilar et al., 2020). While providing important insights, these analyses are limited. They are largely crossectional and focus on counts rather than content. Our study builds on previous research by incorporating both longitudinal and

* 3620 Walnut St., Office 139, Philadelphia, PA, 19104, United States.
E-mail address: calvinis@upenn.edu.

https://doi.org/10.1016/j.socscimed.2024.117542
Received 5 September 2024; Received in revised form 14 October 2024; Accepted 21 November 2024
Available online 22 November 2024

qualitative analyses to explore how bias in media coverage of different risks emerges. We employed multi-level and autoregressive distributed lag models to examine whether fluctuations in death rates correspond to changes in media coverage over time. Additionally, we used natural language processing techniques to analyze the content of articles, coding for mentions of mitigation strategies and assessing the tone of coverage through sentiment analysis. In tandem, these techniques provide more comprehensive insights than previous approaches, assessing how bias emerges through multiple pathways.

News organizations have significant control over the topics they choose to cover, the facts they present from those topics, and how they frame those facts. For example, consider an article about the genetic determinants of cancer framed around an individual's tragic story in comparison to an article on exercise and diet tips for preventing heart disease. Both articles may present empirically accurate information but depict opposing perspectives about control over health. When media organizations collectively emphasize certain risks more than others, focus on specific facts, and frame those facts in particular ways, they present a distorted perspective that may influence people's perceptions and behaviors.

There are well-developed theories that help to make sense of how media may impact consumers' understanding of the world (Potter, 2011; Thomas, 2022). Agenda-setting theory, for example, suggests that media prioritizes certain issues, influencing the public's perception of what is important (Gilardi et al., 2022; Langer and Gruber, 2021; McCombs and Valenzuela, 2020). Under the current framework, agenda-setting works at the topic level, and its impacts may expand over time, a notion grounded in cultivation theory, which argues that prolonged engagement with media is important for shaping perceptions (Busselle & Van den Bulck, 2019; Gerbner et al., 1986; Hermann et al., 2023). Our longitudinal analysis examines how changes in mortality rates align with media coverage (agenda-setting distortions) and the stability of coverage over time (prolonged cultivation). Framing theory focuses on the qualitative aspects of media, suggesting that the way information is presented can influence audiences' perspectives (Gao et al., 2022; Guenther et al., 2021; Scheufele, 1999). For health risks, key frames include individual versus collective focus and threat or sentiment. Collective frames highlight community or policy solutions, while individual frames emphasize personal responsibility and behavior. Additionally, threatening or negative descriptions signal greater concern. Our qualitative analysis considers these frames.

In summary, this paper investigates distortions in media coverage of mortality risks, focusing on topic agendas and qualitative framing. We analyze a 20-year period to understand time-trends in this coverage.

2. Materials and methods

Risks: We selected a diverse set of risks including six chronic diseases that were responsible for most deaths within the United States (heart disease, cancer, strokes, chronic lower respiratory disease, Alzheimer's, and diabetes), four more sensational risks that we expected to be more common in the media (homicide, suicide, overdose, and terrorism), and four other major causes of death (traffic accidents, influenza, sexually transmitted disease, and pandemics/COVID-19).

Media Coverage: Our analysis focused on newspaper coverage of mortality rates across various risks. We first identified the ten largest U. S. newspapers by circulation in 2022 (Turvill, 2022), prioritizing these major outlets due to their extensive readership and broad influence. For the longitudinal analysis, we then selected four newspapers from this list that demonstrated stable and consistent coverage within the LexisNexis database from 1999 to 2020. The requirement for stable coverage over this period was essential for ensuring the reliability and continuity of our longitudinal assessment, thereby limiting our final sample to these four outlets. This final sample is limited in that it may not capture trends in local newspapers or represent other media forms, such as social media or television. Nonetheless, given the considerable reach of these outlets, any bias identified within this sample remains significant. To identify relevant articles, we generated comprehensive sets of keywords tailored to each risk (See Supplementary Table S1). We then queried LexisNexis using these keywords, resulting in a total of 823,406 articles. We combined the number of newspaper articles across journals to arrive at monthly counts of media coverage for each risk.

Mortality Rates: We identified mortality rates using data from the CDC's Underlying Cause of Death database (Centers for Disease Control and Prevention, National Center for Health Statistics, 2021). These data are based on death certificates from U.S. residents.

This analysis covers all demographics and locations, providing mortality rates by month, year, and risk. We supplemented these data for pandemics, using the CDC's COVID Data Tracker (Centers for Disease Control and Prevention, 2024), and for Terrorism using the Global Terrorism Database (LaFree and Dugan, 2007).

Longitudinal Analysis: After collecting the newspaper mentions for each outlet, we conducted a monthly count analysis to ensure comprehensive coverage. This quality check confirmed stable coverage across all outlets, but it also revealed two data challenges. First, the Tampa Bay Times showed unusually large surges in mentions during several months of 2019. Further investigation indicated that these surges were likely due to changes in data recording within the LexisNexis database rather than genuine increases in risk mentions, as they deviated several standard deviations from the norm, affected all risks, and were not observed in other outlets. To address this, we replaced the 2019 outliers in the Tampa Bay Times data with 2018 averages. Additionally, we found instances where deaths were attributed to both homicide and terrorism. To avoid double counting, we subtracted terrorism deaths from homicide counts. We then Winsorized the top and bottom 1% of monthly counts for each risk to handle outliers and log-transformed the modified counts. Notably, all adjustments were pre-planned and incorporated into our pre-registration prior to the primary analysis.

Different longitudinal analyses offer unique insights. For example, analyzing leveled data assesses if long-term mortality trends align with long-term media coverage trends, while differenced data reveals if short-term changes in coverage match short-term changes in deaths. Including lagged terms accounts for delays between changes in mortality and media coverage. Finally, monthly data may be too short of an interval or suffer from dynamic seasonality, so yearly trends are also analyzed. We ran and report four sets of regressions (1) multi-level regression with monthly mortality rates predicting leveled monthly media coverage, with random intercepts and slopes for each risk. (2) The same multi-level regression with the deseasonalized and differenced data. (3) Autoregressive distributed lag models using both leveled and differenced data, ran separately for each risk. (4) The same regressions as 1–3 but with yearly data rather than monthly.

Qualitative analysis: To gather articles for qualitative analysis, we employed a stratified random sampling method, where the population of articles was divided into strata based on the publication year, with 5 articles randomly selected from each year between 1999 and 2020. This approach ensured balanced representation of articles across time, reducing the potential for bias associated with temporal fluctuations in article content and frequency. We were interested in the presented facts about the causes of the risk-factor along with mitigation strategies, and how those facts were framed: whether the focus was on individuals or the collective, as well as the sentiment of the article's description. To assess these dimensions, we first conducted a manual review of a subset of articles to identify key themes and developed an initial coding framework that addressed the dimensions of interest. We then iteratively designed and refined a prompt for GPT-4, providing explicit instructions and examples to guide the model (see full prompt in Appendix S1.3). This development involved multiple rounds of testing and adjustment to ensure that GPT-4's outputs closely matched human judgments. The full text of each article was then processed individually through OpenAI's API using the finalized prompt with certain parameters (Temperature = 0.8, Max Tokens = 1,500). The model's responses were then recorded and aggregated to compare across risks.

While GPT-4 performs well in tasks like this coding exercise (Achiam et al., 2023), there are limitations to its use in qualitative analysis. Specifically, large language models may lack a nuanced understanding of social and cultural contexts, which can lead to misinterpretations (McIntosh et al., 2024). Additionally, biases in training data may affect coding outcomes, perpetuating demographic or cultural biases (Bano et al., 2024). We attempted to address these issues by focusing our prompt instructions on simpler, context-independent judgments, when possible, an approach that has been shown to increase GPT's accuracy in other tasks (Hou et al., 2024). We also supplemented this analysis with a more typical method of sentiment analysis, calculating the VADER score of every sentence within articles and reporting the mean across articles (See SI 1.5. Supplemental analysis with sentiment analysis using VADER; Hutto and Gilbert, 2014).

Our analysis plan was pre-registered: https://osf.io/mr2ay/?view_only=1b8c35b31a5c4d7c87cbebd08ef962f0. See minor divergences in Appendix S1.6. We report the results for all analyses conducted such that there are no "file-drawer" analyses.

3. Results

First, Table 1 shows the number of articles and deaths by risk, highlighting a significant disparity between mortality rates and news coverage, with one article per 323 heart disease deaths compared to 36 articles per single terrorism death. In general, risks that are deadlier have lower ratios of media coverage to deaths; we find a strong negative relationship between these measures. Due to the concentration of deaths from pandemics/COVID-19 and terrorism in short periods, we analyze these risks separately for the remaining analyses. We also limit our data to 1999–2019 because news coverage changed drastically during the COVID-19 pandemic.

To understand how changes in mortality rates correspond to media coverage, we plot both time series for each risk in Fig. 1 (See SI Figure S1 for pandemics/Covid19 and Terrorism plot, and Figs. S2–S3 for alternative visualizations). For some risks (e.g., homicide), media mentions are far higher than the death rate, whereas for others (e.g., heart disease), deaths outpace media mentions. This illustrates unequal media coverage for these risks. Further, these plots suggest that there may be associations between these measures (see e.g., Alzheimer's which has trends that move together) but that the associations are not consistent or strong (see e.g, Cancer which has largely unrelated trends). Finally, we looked at the ratios from Table 1 over time and found that they were stable from 1999 to 2020, with sensational risks consistently over-represented in the news and chronic illnesses consistently under-represented, an important result for cultivation theory which suggests that prolonged engagement can increase media effects. Next, we report our modeling results to statistically assess these associations.

Table 1
Counts of articles and deaths for each risk from 1999 to 2020. The last column displays the ratio of articles to deaths for each risk and the table is ordered by this column.

Risk	Articles	Deaths	Articles to Deaths
Heart Disease	48,024	15,499,612	1:323
Respiratory Diseases	15,282	3,064,049	1:200
Cancer	132,016	12,644,869	1:96
Stroke	41,433	3,184,602	1:77
Alzheimer's	27,026	1,872,576	1:69
Diabetes	25,395	1,674,724	1:66
Influenza	20,264	1,257,088	1:62
Overall	**823,406**	**42,879,844**	**1:52**
Traffic Accidents	28,534	956,960	1:34
Overdose	41,412	892,857	1:22
Suicide	60,913	838,850	1:14
STDs	40,320	210,774	1:5
Pandemics	92,295	385,666	1:4
Homicide	127,383	393,756	1:3
Terrorism	123,109	3,461	36:1

We begin with a multi-level regression on the leveled data fitted by REML, using Satterthwaite's method for t-tests with the LMER package in R (Bates et al., 2005):

$$y_{ij} = (\beta_0 + u_j) + (\beta_1 + v_j)x_{ij} + f_{month} + \epsilon_{ij} \tag{1}$$

In this model, monthly deaths from a risk (x_{ij}, i indicating month and j indicating risk) predict monthly media mentions (y_{ij}) along with random intercepts (u_j) and slopes (v_j) for each risk and a monthly dummy (f_{month}) to account for seasonality. The fixed effect of deaths was positive and significant (b $= 0.38$, p $= 0.004$) and the random intercepts and slopes allowed the full model to significantly outperform the model with only fixed effects ($p < 0.001$, See also Supplemental Tables S2–S3). These results indicate that higher death rates from a risk lead to increased media coverage, though the relationship varies by risk. Fig. 2 visually represents this relationship, showing that for all risks except cancer and respiratory disease, there is a positive correlation between deaths and media mentions, despite significant variations in the number of deaths and media coverage for each risk.

We ran a similar regression using differenced and deseasonalized mortality ($\hat{x}_{ij}$) and media mentions ($\hat{y}_{ij}$):

$$\hat{y}_{ij} = (\beta_0 + u_j) + (\beta_1 + v_j)\hat{x}_{ij} + \epsilon_{ij} \tag{1a}$$

Once more, we found significant results for the random slope model compared to a model with only fixed effects ($p < 0.001$). The fixed effect was significant and positive ($b = 0.42, p = 0.014$, See also Supplemental Tables S4–S5). This suggests that when there were significantly more deaths than the previous month for a given risk, newspapers increased their coverage of that risk, and the random slopes suggest that this relationship varies across risks.

The above analyses were contemporaneous, but mortality information may take time to reach the media. Therefore, we also ran autoregressive distributive lag models with lagged media mentions and mortality rates for each risk. We ran separate regressions for each risk with both the leveled and differenced data, comparing full models with lagged mortality information to models without these terms. For leveled data models (Supplementary Table S6), including lagged mortality improved model fit for three of the 12 risks at typical significance thresholds (Supplementary Table S7). Collectively, these models were jointly significant ($p = 0.003$, Fisher's method). For differenced and deseasonalized data, models were significant for four of the 12 risks (Supplementary Tables S8–S9), producing a jointly significant result ($p < 0.001$).

While we found significant relationships between deaths and media mentions across risks, death information only explains about 2% of the media mentions a risk receives (leveled mean change in $R^2 = 0.017$, differenced mean change in $R^2 = 0.028$, both in comparison to models without death information). This demonstrates a weak connection between these measures. Finally, following our pre-registration, we ran similar models on the annual data for robustness and found very similar results (See Supplemental Analysis S1.1). While these results demonstrate how risks' mortality and media coverage are largely disconnected, it is important to note that the risks included in the primary analysis are perennial problems that have long plagued humanity. If one considers more novel risk factors (e.g., pandemics/COVID-19 and terrorism), the media is initially very responsive to these threats. Still, these too quickly become self-perpetuating stories that are largely unexplained by changes in the death rate (see Appendix S1.2).

Next, we present our qualitative analysis, focusing on how articles discuss risk causes and mitigation strategies and how they frame these facts in terms of effectiveness, individual vs. collective emphasis, and sentiment. Using GPT-4, we found substantial variation across risks in these dimensions (See Appendix S1.4 for more details).

When discussing risk causes, articles mainly cited environmental and

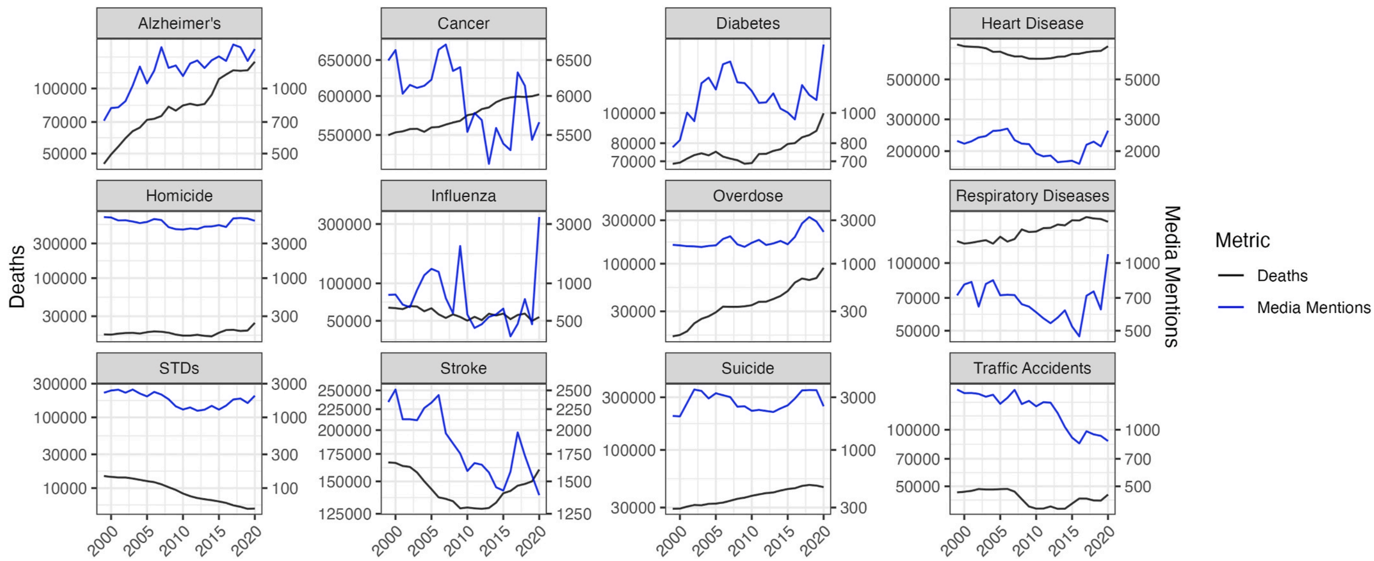

Fig. 1. Annual deaths (black) and media mentions (blue) of twelve risks from 1999 to 2020. While some risks, like Alzheimer's (top left), show similar trends, others are largely disconnected, like cancer (one row right), revealing how there is not a strong, consistent relationship between these measures across time. Note that the y-axes are scale-free to highlight trends within each series; however, in each plot, the media mentions axis is 10× the death-rate axis. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

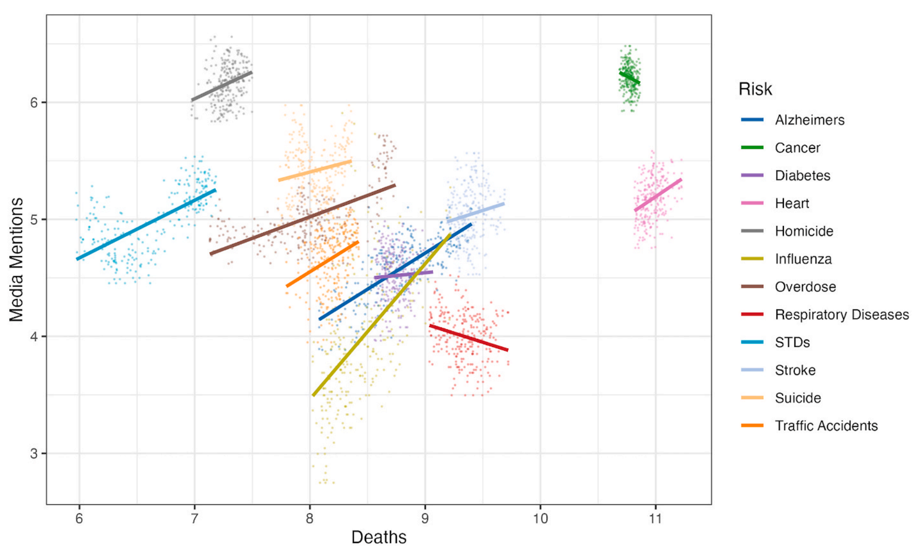

Fig. 2. Log monthly deaths on log monthly media mentions for each risk. Most of the slopes are positive, suggesting that in months when a risk results in more deaths, the media is more likely to cover that risk.

lifestyle factors, though this varied by risk. For example, chronic respiratory disease was often linked to environmental factors, diabetes to lifestyle choices, and homicide to a mix of both. Mitigation strategies fell into three categories: policy, behavioral, and technological solutions. Articles on chronic diseases emphasized behavioral and technological solutions, while those on sensational risks focused more on policy solutions (Fig. 3, Panel A). Articles also varied substantially in framing (Fig. 3). Sensational risks were described with more negative emotions than chronic illnesses, which were generally presented neutrally (Fig. 3, panel B). Sentiment ratings from VADER show these trends are stable over the 20-year period analyzed. Fig. 3, Panel C, illustrates the mean annual sentiment for all risks. See Appendices S1.4 and S1.5 for more detailed results.

To determine if our qualitative analysis revealed new or similar distortions compared to our quantitative results, we compared the mean sentiment and narrative focus by risk ($N = 14$) to the ratios from Table 1 and the random slopes from our mixed regression. We found that risks with more articles to deaths were also covered in a more negative sentiment ($\rho = -0.60, p = 0.025$) and were more likely to present collective solutions ($\rho = 0.73, p = 0.003$). In contrast, our qualitative results were unrelated (both $p > 0.29$) to the slope adjustments from our mixed regressions. These patterns, along with their stability over time, suggest that there are different forms of bias at play.

4. Discussion

Like prior studies (Bomlitz and Brezis, 2008; Combs and Slovic, 1979; Frost et al., 1997; Pilar et al., 2020), we found a substantial gap between the mortality risks in the US and those covered in the news. Chronic illnesses are relatively less covered compared to sensational risks, indicating how the news sets an agenda that prioritizes certain risks (Gilardi et al., 2022; Langer and Gruber, 2021; McCombs and

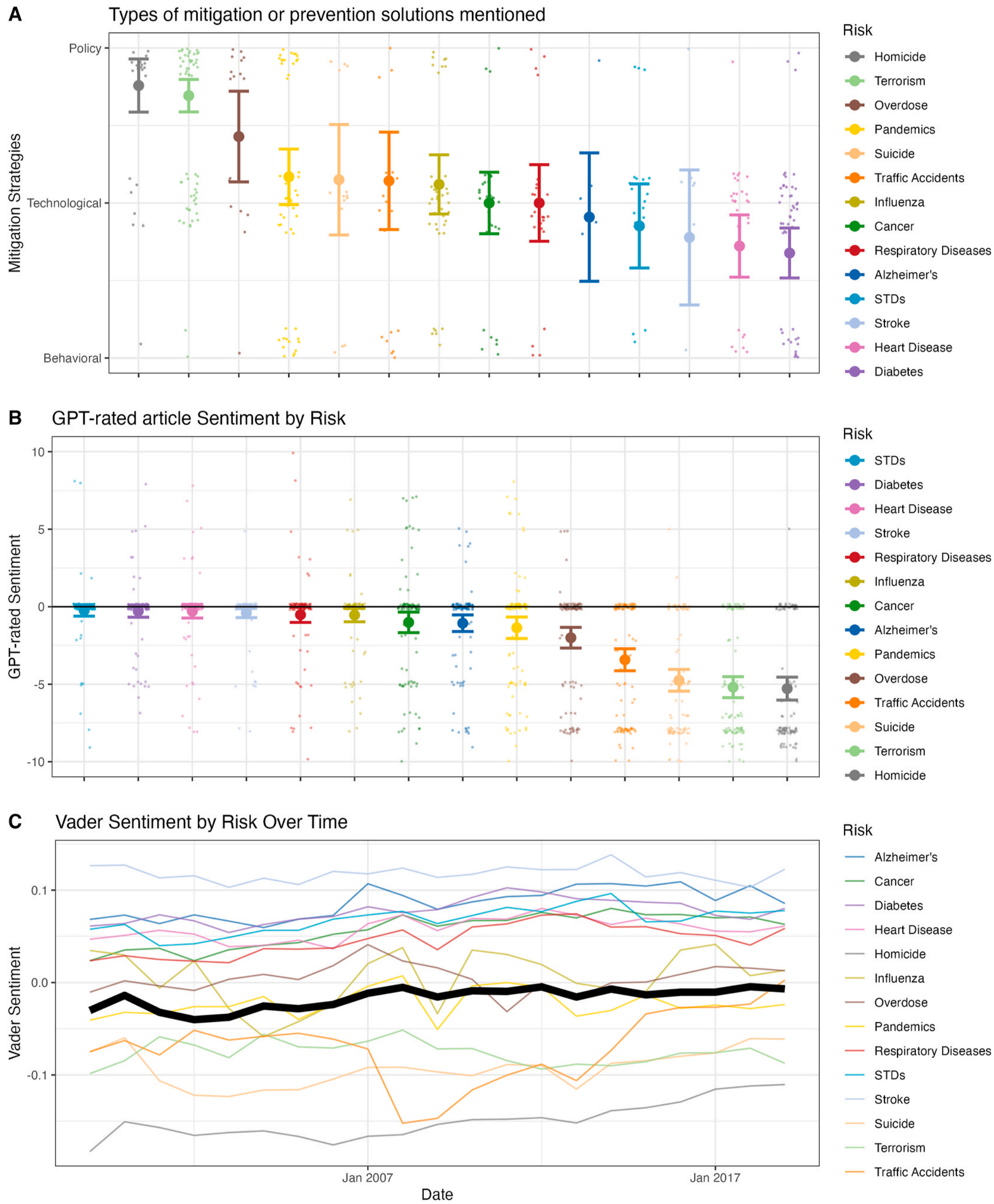

Fig. 3. Panel A depicts solution strategies mentioned in articles, including behavioral, technological, and policy solutions, revealing how articles about sensational risks tend to present collective, policy solutions whereas articles on chronic illnesses share technological or individual behavioral mitigation strategies. Panel B shows GPT-4 sentiment (−10 negative to 10 positive), revealing that sensational risks are described much more negatively than chronic illnesses. For panels A and B, the large dots and error bars show means and 95% CIs and the small dots represent individual articles. Panel C shows mean yearly sentiment, as scored by VADER, for the entire study period for each risk, with the dark black line showing the overall mean sentiment. This panel illustrates how the differences in sentiment remain largely stable over time.

Valenzuela, 2020). Qualitative analysis revealed similar biases: chronic diseases were portrayed neutrally with individual-focused solutions, whereas sensational risks were depicted negatively and with collective-focused strategies. These framings suggest that sensational risks are more concerning and require community or policy solutions and that such solutions are less necessary for chronic illnesses. Additionally, our sentiment analysis of newspaper articles aligned with recent studies on stigma in social media posts about health risks (Brown et al., 2023c), indicating that differences in how various risks are described may exhibit similar patterns across different media platforms. Finally, longitudinal analysis revealed that distortions in framing and topic remained stable from 1999 to 2020, an important result under cultivation theory (Busselle & Van den Bulck, 2019; Hermann et al., 2023).

Despite these distortions, we identified a significant though weak relationship between changes in mortality rates and media coverage. While other factors explain the majority of coverage variance, the media somewhat reflects changes in risk patterns. For more novel risks like terrorism and pandemics/COVID-19, the media was much more initially responsive. Perhaps newspapers are focusing on, well, news and so are less motivated to discuss perennial mortality risks. Still, we found similar distortions for longstanding sensational risks (e.g., homicide, suicide, drug overdose), suggesting novelty only partially explains this disconnect.

Media attention to different topics may largely be a product of consumer demand. In this case, news organizations may cater to consumer preferences, focusing on sensational risks that engage consumers because of evolutionary or cultural factors. Indeed, major themes from evolutionary psychology (e.g., norm violations, tribalism) often result in sensational news stories appearing more salient and engaging to consumers (Davis and McLeod, 2003). These evolutionary demands may interact with other cultural norms that drive consumer demand for sensational risks, and news organizations may amplify sensational content out of competition for audience members (Arrese et al., 2019). As a result, even if certain news outlets were to adjust their coverage to better reflect the risks people face, audience preferences for sensational content could push them toward alternative sources that offer more engaging, albeit less accurate, information. The importance of consumer demand highlights a key consideration in interpreting our analysis. While we examine the disconnect between mortality rates and media coverage, it is neither practical nor ideal to expect proportional representation of risks in the media based solely on their incidence. Instead, we encourage journalists to be mindful of how their selection of topics and framing may inadvertently lead to inaccurate risk perceptions, potentially resulting in suboptimal behaviors and policy decisions. Journalists should aim to provide a more balanced view by noting the relative risks when covering different mortality risks, publishing articles that contextualize the scale of various risks, and critically assessing how much attention their outlet devotes to certain risks compared to others.

We found statistical support for some initially counterintuitive relationships. First, there is a strong negative relationship between media coverage per death and the total number of deaths across risks. This likely results from the death rate being part of the ratio, as media systems have limited bandwidth, the number of stories about a risk may asymptote rather than increase linearly with the death rate. Alternatively, more frequent causes of death may be less appealing to discuss than rare, sensational risks due to consumer preferences. The second unexpected result is that the differences in random slopes across risks showed varying relationships. While most risks had a positive relationship between death rate and media mentions, some had negative slopes. These variations may stem from differences in public awareness across risks. For instance, widespread anti-cancer campaigns during the study period likely drove media coverage without reflecting short-term death rate fluctuations. In contrast, risks like influenza showed a strong positive relationship between media coverage and death rates, likely due to clear seasonality and public anticipation of yearly strains.

Although our mortality data is derived from aggregate statistics provided by the CDC, it is important to acknowledge that media organizations typically do not base their coverage on these statistics. Instead, news outlets tend to focus on specific events, with unique homicides, suicides, terrorist attacks, or deaths of prominent individuals from chronic illnesses prompting immediate coverage. While some articles may reference aggregate data, this is not the primary lens through which risks are reported. Therefore, any delay in the publication of aggregate mortality statistics is unlikely to materially drive our results. Additionally, major publications often cover international events, which may influence the relationship between deaths and media coverage observed in our study. Such international coverage could be a source of the bias underlying our results that could still shape consumers' perceptions of the risks they face. Future research should explore other sources of bias that may lead media organizations to emphasize risks beyond those directly relevant to their audiences.

Our study has limitations regarding our sample and analytical techniques. Despite our large dataset covering four major US newspapers, news coverage makes up only a small portion of people's media-diets (Allen et al., 2020). Individuals get most of their media from other sources, so future work could benefit by examining descriptions of mortality risks from other information sources. Of particular interest is digital media—including social media and other internet platforms—because it has gained widespread adoption (Twenge et al., 2019). Consequently, risk communication on these platforms may be especially important in shaping individual beliefs and behaviors. Additionally, our observational study cannot speak to causal patterns between these measures or their impact on consumers, as we did not explore how our findings correspond to perceptions, behaviors, or policy changes. Future research can address this concern with experimental designs or surveys combined with observational data. Two approaches are particularly promising for establishing causality. The first, outlined by Hornik et al. (2022), uses observational data on media coverage of different risks from numerous information sources, pairing coverage fluctuations with related health and consumption behaviors measured across time with a rolling national-representative survey. The second approach employs experimental techniques to measure the impacts of various articles on a given risk behavior (Allen et al., 2024), then leverages machine learning and the wisdom of crowds to extrapolate these effects to a larger sample of articles, combining the results with consumption data to determine the net impact of media coverage of risks on behaviors. Research that includes such designs offer a crucial next step in understanding the impacts of distorted media coverage of mortality risks on population health.

Although we did not include belief or behavioral outcomes in our analysis, we briefly speculate on the potential consequences of our findings. First, individuals often use simple heuristics when making risk-decisions (Gigerenzer and Todd, 1999; Siegrist and Arvai, 2020), which can be skewed by media distortion. For example, relying on the availability heuristic, people may overestimate the prevalence of heavily covered risks (Hertwig et al., 2005), viewing them as more problematic and worthy of attention. Consequently, they may neglect less-covered but more prevalent and preventable chronic risks. Similarly, the framing of articles may impact peoples' beliefs about these risks (Gao et al., 2022; Guenther et al., 2021), leading consumers to underestimate the importance of collective solutions.

Beliefs about mortality risk may impact other seemingly unrelated behaviors. For instance, if individuals feel at risk from uncontrollable factors, they may value deferred rewards less (Pepper and Nettle, 2017; Brown and Pepper, 2024b), leading to more present-oriented behavior and ultimately poorer health outcomes. This behavioral hypothesis has empirical support (Pepper and Nettle, 2017; Brown and Pepper, 2024a), and research shows that beliefs about mortality risk are weakly tied to actual threats (Brown et al., 2023a; Isch et al., 2023). The media distortions identified here may contribute to that disconnect.

Addressing the negative impacts of media bias will require initiatives from several stakeholders. Journalists can provide information about the relative rates of different mortality risks to ensure viewers are aware of the true risks they face. Public health agencies can partner with media outlets to promote more accurate and frequent coverage of chronic diseases, shaping narratives that better reflect the prevalence and preventability of such diseases. Health practitioners should also be proactive in addressing misconceptions, ensuring that patients receive accurate information to counter misguided beliefs shaped by skewed media coverage. Finally, media literacy programs can empower the public to critically evaluate the information they consume. These initiatives can equip individuals with the skills to recognize how media reporting may distort the perception of risk, ultimately leading to more informed health decisions and behaviors. Together, these strategies can help align public perceptions with actual health risks, potentially improving public health outcomes.

In conclusion, we identified significant distortions in media coverage of mortality risks with newspapers significantly over-representing sensational risks and under-representing chronic risks. Looking within risks over time, we found that there is a significant relationship between mortality rates and media coverage, but that this relationship is extremely weak. Additionally, sensational risks were framed more negatively and were often presented with more collective solution strategies than chronic illnesses. Ultimately, understanding these distortions is crucial for improving public awareness and informing media literacy initiatives, ensuring a more accurate and balanced perception of mortality risks.

Ethics approval

This study involved the analysis of publicly available and commercial, observational data and did not include any interaction with human participants, identifiable private information, or interventions. Therefore, ethics approval was not required for this research.

Declaration of competing interest

The authors report there are no competing interests to declare.

Acknowledgements

CI would like to thank Prof. Michul Shin for his extensive support in writing this paper. He would also like to thank Prof. Duncan Watts and members of the CSS lab who attended an early talk on this topic and provided feedback. Finally, CI is grateful to the Institute for Humane Studies for their support (grant no. IHS018154).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.socscimed.2024.117542.

Data availability

Data will be made available on request.

References

Achiam, J., Adler, S., Agarwal, S., Ahmad, L., Akkaya, I., Aleman, F.L., Almeida, D., Altenschmidt, J., Altman, S., Anadkat, S., et al., 2023. Gpt-4 technical report. arXiv preprint arXiv:2303.08774. https://arxiv.org/abs/2303.08774.

Allen, J., Howland, B., Mobius, M., Rothschild, D., Watts, D.J., 2020. Evaluating the fake news problem at the scale of the information ecosystem. Sci. Adv. 6 (14), eaay3539. https://doi.org/10.1126/sciadv.aay3539.

Allen, J., Watts, D.J., Rand, D.G., 2024. Quantifying the impact of misinformation and vaccine-skeptical content on Facebook. Science 384 (6699), eadk3451. https://doi.org/10.1126/science.adk3451.

Arrese, A., Medina, M., Sánchez-Tabernero, A., 2019. Consumer demand and audience behavior. In: A Research Agenda for Media Economics. Edward Elgar Publishing, pp. 59–76. https://doi.org/10.4337/9781788119061.00010.

Bates, D., et al., 2005. Fitting linear mixed models in r. R. News 5 (1), 27–30. https://archive.linux.duke.edu/cran/doc/Rnews/Rnews2005-1.pdf#page=27.

Bano, M., Hoda, R., Zowghi, D., Treude, C., 2024. Large language models for qualitative research in software engineering: exploring opportunities and challenges. Autom. Software Eng. 31 (1), 8. https://doi.org/10.1007/s10515-023-00407-8.

Bomlitz, L.J., Brezis, M., 2008. Misrepresentation of health risks by mass media. J. Publ. Health 30 (2), 202–204. https://doi.org/10.1093/pubmed/fdn009.

Branning, G., Vater, M., 2016. Healthcare spending: plenty of blame to go around. Am. Health Drug Benefits 9 (8), 445. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5394555/.

Brown, R., Sillence, E., Coventry, L., Branley-Bell, D., Murphy-Morgan, C., Durrant, A.C., 2023a. Health stigma on Twitter: investigating the prevalence and type of stigma communication in tweets about different conditions and disorders. Front. Commun. 8, 1264373. https://doi.org/10.3389/fcomm.2023.1264373.

Brown, R., Sillence, E., Pepper, G., 2023b. Individual characteristics associated with perceptions of control over mortality risk and determinants of health effort. Risk Anal. https://doi.org/10.1111/risa.14243.

Brown, R., Sillence, E., Pepper, G., 2023c. Perceptions of control over different causes of death and the accuracy of risk estimations. J. Publ. Health 1–14. https://doi.org/10.1007/s10389-023-01910-8.

Brown, R., Pepper, G., 2024a. The relationship between perceived uncontrollable mortality risk and health effort: replication, secondary analysis, and mini meta-analysis. Ann. Behav. Med. 58 (3), 192–204. https://doi.org/10.1093/abm/kaad072.

Brown, R.D., Pepper, G.V., 2024b. The uncontrollable mortality risk hypothesis: theoretical foundations and implications for public health. Evol., Med., Publ. Health 12 (1), 86–96. https://doi.org/10.1093/emph/eoae009.

Busselle, R., Van den Bulck, J., 2019. Cultivation Theory, Media, Stories, Processes, and Reality. Media effects, pp. 69–82. In: https://www.taylorfrancis.com/chapters/edit/10.4324/9780429491146-5/cultivation-theory-media-stories-processes-reality-rick-busselle-jan-van-den-bulck.

Byrow, Y., Pajak, R., Specker, P., Nickerson, A., 2020. Perceptions of mental health and perceived barriers to mental health help-seeking amongst refugees: a systematic review. Clin. Psychol. Rev. 75, 101812. https://doi.org/10.1016/j.cpr.2019.101812.

Centers for Disease Control and Prevention, 2024. COVID data tracker. https://covid.cdc.gov/covid-data-tracker.

Centers for Disease Control and Prevention, National Center for Health Statistics, 2021. National vital statistics system, mortality 1999-2020 on CDC WONDER online database. http://wonder.cdc.gov/ucd-icd10.html.

Centers for Disease Control and Prevention, National Center for Health Statistics, 2022. National center for chronic disease prevention and health promotion: about the center. https://www.cdc.gov/nccdphp/about/?CDC_AAref Val=https://www.cdc.gov/chronicdisease/center/index.htm.

Centers for Medicare & Medicaid Services, 2021. National health expenditures highlights. https://www.cms.gov/data-research/statistics-trends-and-reports/national-health-expenditure-data/historical.

Combs, B., Slovic, P., 1979. Newspaper coverage of causes of death. Journal. Q. 56 (4), 837–849. https://doi.org/10.1177/107769907905600420.

Davis, H., McLeod, S.L., 2003. Why humans value sensational news: an evolutionary perspective. Evol. Hum. Behav. 24 (3), 208–216. https://doi.org/10.1016/S1090-5138(03)00012-6.

Feldman, R., 2020. Perverse incentives: why everyone prefers high drug prices-except for those who pay the bills. Harv. J. Legislation 57, 303. https://heinonline.org/HOL/LandingPage?handle=hein.journals/hjl57&div=13&id=&page=.

Frost, K., Frank, E., Maibach, E., 1997. Relative risk in the news media: a quantification of misrepresentation. Am. J. Publ. Health 87 (5), 842–845. https://doi.org/10.2105/AJPH.87.5.842.

Gao, R., Guo, H., Li, F., Liu, Y., Shen, M., Xu, L., Yu, T., Li, F., 2022. The effects of health behaviours and beliefs based on message framing among patients with chronic diseases: a systematic review. BMJ Open 12 (1), e055329. https://doi.org/10.1136/bmjopen-2021-055329.

Gerbner, G., Gross, L., Morgan, M., Signorielli, N., 1986. Living with television: the dynamics of the cultivation process. Perspect. Media Effects 17–40, 1986. http://wiki.commres.org/pds/CultivationTheory/LivingWithTelevisionTheDynamicsoftheCultivationProcess.pdf.

Gigerenzer, G., Todd, P.M., 1999. Simple Heuristics that Make Us Smart. Oxford University Press, USA dangoldstein.com/papers/Czerlinski How Good Are_Simple Heuristics Chapter_1999.pdf.

Gilardi, F., Gessler, T., Kubli, M., Müller, S., 2022. Social media and political agenda setting. Polit. Commun. 39 (1), 39–60. https://doi.org/10.1080/10584609.2021.1910390.

Guenther, L., Gaertner, M., Zeitz, J., 2021. Framing as a concept for health communication: a systematic review. Health Commun. 36 (7), 891–899. https://doi.org/10.1080/10410236.2020.1723048.

Hermann, E., Morgan, M., Shanahan, J., 2023. Cultivation and social media: a metaanalysis. New Media Soc. 25 (9), 2492–2511. https://doi.org/10.1177/14614448231180257.

Hertwig, R., Pachur, T., Kurzenhäuser, S., 2005. Judgments of risk frequencies: tests of possible cognitive mechanisms. J. Exp. Psychol. Learn. Mem. Cognit. 31 (4), 621. https://doi.org/10.1037/0278-7393.31.4.621.

Hoffman, D., 2022. Commentary on chronic disease prevention in the us in 2022. Ann. Bioeth. Clin. App. 5 (2), 10–23880. https://doi.org/10.23880/abca-16000231.

Hornik, R., Binns, S., Emery, S., Epstein, V.M., Jeong, M., Kim, K., et al., 2022. The effects of tobacco coverage in the public communication environment on young people's decisions to smoke combustible cigarettes. J. Commun. 72 (2), 187–213. https://doi.org/10.1093/joc/jqab052.

Hou, C., Zhu, G., Zheng, J., Zhang, L., Huang, X., Zhong, T., et al., 2024. Prompt-based and fine-tuned GPT models for context-dependent and-independent deductive coding in social annotation. In: Proceedings of the 14th Learning Analytics and Knowledge Conference, pp. 518–528. https://doi.org/10.1145/3636555.3636910.

Hutto, C., Gilbert, E., 2014. Vader: a parsimonious rule-based model for sentiment analysis of social media text. Proc. Int. AAAI Conf. Web Soc. Media 8 (1), 216–225. https://doi.org/10.1609/icwsm.v8i1.14550.

Isch, C., Brown, R., Todd, P.M., Aktipis, A., Pepper, G., 2023. Objective risk exposure, perceived uncontrollable mortality risk, and health behaviors. J. Publ. Health 1–11. https://doi.org/10.1007/s10389-023-01994-2.

LaFree, G., Dugan, L., 2007. Introducing the global terrorism database. Terrorism Polit. Violence 19 (2), 181–204. https://doi.org/10.1080/09546550701246817.

Langer, A.I., Gruber, J.B., 2021. Political agenda setting in the hybrid media system: why legacy media still matter a great deal. Int. J. Press/Polit. 26 (2), 313–340. https://doi.org/10.1177/1940161220925023.

Maani, N., Galea, S., 2020. COVID-19 and underinvestment in the public health infrastructure of the United States. Milbank Q. 98 (2), 250, 10.1111%2F1468-0009.12463.

McCombs, M., Valenzuela, S., 2020. Setting the Agenda: Mass Media and Public Opinion. John Wiley & Sons. https://lccn.loc.gov/2020017656.

McIntosh, T.R., Liu, T., Susnjak, T., Watters, P., Halgamuge, M.N., 2024. A reasoning and value alignment test to assess advanced GPT reasoning. ACM Trans. Interact. Intell. Syst. https://doi.org/10.1145/3670691.

Miller, G., Roehrig, C., Hughes-Cromwick, P., Lake, C., 2008. Quantifying national spending on wellness and prevention. In: Beyond Health Insurance: Public Policy to Improve Health, vol. 19. Emerald Group Publishing Limited, pp. 1–24. https://www.emerald.com/insight/content/doi/10.1016/S0731-2199%2808%2919001-X/full/html.

Pepper, G.V., Nettle, D., 2017. The behavioural constellation of deprivation: causes and consequences. Behav. Brain Sci. 40, e314. https://doi.org/10.1017/S0140525X1600234X.

Pilar, M.R., Eyler, A.A., Moreland-Russell, S., Brownson, R.C., 2020. Actual causes of death in relation to media, policy, and funding attention: examining public health priorities. Front. Public Health 8, 545333. https://doi.org/10.3389/fpubh.2020.00279.

Potter, W.J., 2011. Conceptualizing mass media effect. J. Commun. 61 (5), 896–915. https://doi.org/10.1111/j.1460-2466.2011.01586.x.

Scheufele, D.A., 1999. Framing as a theory of media effects. J. Commun. 49 (1), 103–122. https://doi.org/10.1111/j.1460-2466.1999.tb02784.x.

Siegrist, M., Arvai, J., 2020. Risk perception: reflections on 40 years of research. Risk Anal. 40 (S1), 2191–2206. https://doi.org/10.1111/risa.13599.

Thomas, F., 2022. A methodological framework for analyzing the appearance and duration of media effects. J. Commun. 72 (3), 401–428. https://doi.org/10.1093/joc/jqac013.

Turvill, W., 2022. Top 25 US Newspaper Circulations: Print Sales Fall Another 12% in 2022. Press Gazette. Retrieved. https://pressgazette.co.uk/news/us-newspaper-circulations-2022/. (Accessed 1 March 2024).

Twenge, J.M., Martin, G.N., Spitzberg, B.H., 2019. Trends in US Adolescents' media use, 1976–2016: the rise of digital media, the decline of TV, and the (near) demise of print. Psychol. Popular Media Cult. 8 (4), 329. https://doi.org/10.1037/ppm0000203.

Wolfe, M.K., McDonald, N.C., Holmes, G.M., 2020. Transportation barriers to health care in the United States: findings from the national health interview survey, 1997–2017. Am. J. Publ. Health 110 (6), 815–822. https://doi.org/10.2105/AJPH.2020.305579.