

# Characterizing the Differences in Descriptions of Violence on Reddit During the COVID-19 Pandemic

Journal of Interpersonal Violence  
1–25

© The Author(s) 2023




Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/08862605231163885

journals.sagepub.com/home/jiv



Lifang Li<sup>1</sup> , Lilly Neubauer<sup>2</sup>,  
Robert Stewart<sup>1,3</sup>, and Angus Roberts<sup>1</sup>

## Abstract

Concerns have been raised over the experiences of violence such as domestic violence (DV) and intimate partner violence (IPV) during the COVID-19 pandemic. Social media such as Reddit represent an alternative outlet for reporting experiences of violence where healthcare access has been limited. This study analyzed seven violence-related subreddits to investigate the trends of different violence patterns from January 2018 to February 2022 to enhance the health-service providers' existing service or provide some new perspective for existing violence research. Specifically, we collected violence-related texts from Reddit using keyword searching and identified six major types with supervised machine learning classifiers: DV, IPV, physical violence, sexual violence, emotional violence, and nonspecific violence or others. The increase rate (IR) of each violence type was calculated and temporally compared in five phases of the pandemic. The phases include one pre-pandemic phase (*Phase 0*, the date before February 26, 2020) and four pandemic phases (*Phases 1–4*) with separation dates of June 17, 2020, September 7, 2020, and June 4, 2021. We found that the number of

<sup>1</sup>King's College London, UK

<sup>2</sup>University College London, UK

<sup>3</sup>South London and Maudsley NHS Foundation Trust, London, UK

## Corresponding Author:

Lifang Li, Department of Biostatistics & Health Informatics, King's College London, London, Denmark Hill, London, SE5 8AF, UK.

Email: lifang.li@kcl.ac.uk

IPV-related posts increased most in the earliest phase; however, that for COVID-citing IPV was highest in the mid-pandemic phase. IRs for DV, IPV, and emotional violence also showed increases across all pandemic phases, with IRs of 26.9%, 58.8%, and 28.8%, respectively, from the pre-pandemic to the first pandemic phase. In the other three pandemic phases, all the IRs for these three types of violence were positive, though lower than the IRs in the first pandemic phase. The findings highlight the importance of identifying and providing help to those who suffer from such violent experiences and support the role of social media site monitoring as a means of informative surveillance for help-providing authorities and violence research groups.

### **Keywords**

violence types, increase rate, Reddit, data classification

### **Introduction**

Domestic violence (DV) is a great social issue (Mayshak et al., 2022). During the COVID-19 pandemic, DV and intimate partner violence (IPV) increased significantly according to the previous studies (Bettinger-Lopez & Bro, 2020; Sri et al., 2021; Wood et al., 2020). For example, in France, since the lockdown on March 17, 2020, DV cases rose by 30% (Sri et al., 2021). In North America, in the first months of the COVID-19 pandemic, higher number of DV calls to the police were observed (Matoori et al., 2020). In Cyprus and Singapore, calls regarding DV to helplines increased by 30% and 33% after the outbreak of COVID-19, respectively (Sri et al., 2021). In Canada, Spain, Germany, the United States (US), and the United Kingdom, DV cases and demand for emergency shelter also increased during the pandemic (Voth Schrag et al., 2022).

Frequently imposed quarantine increased the contact between perpetrators and victims, potentially leading to underlying increases in the occurrence rates of violence (Lyons & Brewer, 2021) and/or reports (Campbell, 2020). Increasing DV and IPV may also reflect worse population health, rising unemployment, increased anxiety and financial stress, scarcity of community resources, and other factors (Emezue, 2020). More specifically, from the perspective of abusers, heightened financial pressures and stress, increased consumption of alcohol or drugs, and the purchase or hoarding of guns in response to quarantine measures could all also be driving increases in violence (Bettinger-Lopez & Bro, 2020). Victims, on the other hand, may have found themselves to be isolated in violent homes, without access to resources, friendship, and/or family networks (Bettinger-Lopez & Bro, 2020).

To minimize the impact of DV and IPV on public health during pandemics, preventive measures and support programs are needed (Ertan et al., 2020). Several supporting strategies have been proposed, including facilitating online methods, developing alternative violence reporting mechanisms to identify victims (e.g., victim notification apps such as VINEmobile and DV hotlines) (Emezue, 2020; Xue et al., 2020), or maintaining vital sexual and reproductive health services for victims, as well as other emotional and/or financial support (Bettinger-Lopez & Bro, 2020). More specifically, researchers argued for a ban on alcohol sales and limits on access to weapons (Ertan et al., 2020). Digital interventions for DV and IPV could be one way to decrease the occurrences of such violent behavior (Emezue, 2020); however, these may be limited by the stress and burnout experiences of online facilitators and helpers (Sharifi et al., 2020). IPV and sexual assault helpers are reported to be experiencing increased stress and a lack of resources to provide help to survivors and to themselves since the outbreak of the COVID-19 pandemic (Wood et al., 2021). Moreover, using DV hotlines or other expert-based methods in dealing with increasing violence may themselves cause psychological stress, compassion fatigue, and burnout of the help providers (Emezue, 2020).

Social media provide individuals with additional or alternative platforms to share their experiences and can be utilized to raise awareness of the psychological impact of social isolation (Goh et al., 2020; Woodlock et al., 2020). Social media has been widely used to help mitigate loneliness (Hajek & König, 2021; Hamilton et al., 2020; Koh & Liew, 2022). Koh and Liew (2022) used Twitter data to characterize discussions of loneliness during the COVID-19 pandemic and suggested that social media could be an important route for applying potential interventions to deal with emerging COVID-19-related mental health problems (Koh & Liew, 2022). Further, characterizing and/or monitoring temporal changes in abusive behavior before and after the outbreak of COVID-19 becomes possible based on social media data (Babvey et al., 2021). Researchers found significant increase in abusive textual contents during quarantine in the US using Twitter data, and they found significant changes in mentions of DV after the outbreak of the COVID-19 pandemic using Reddit data (Babvey et al., 2021).

Previous studies shed light on the role of social media in allowing people to express abusive/violent experiences and provide social support or necessary information for protecting victims and survivors during the pandemic (Babvey et al., 2021; Koh & Liew, 2022; Xue et al., 2020). However, further investigation is still needed on the disclosure of various forms of violence represented in social media data. Moreover, the COVID-19 pandemic is an important triggering factor for the increase in DV, but it may at the same time

reduce the visibility of abusive/violent behaviors if the attention of health workers and resources are more focused on the pandemic itself (Bhavasar et al., 2021). Social media may have been a choice for the public, especially victims, to disclose their violent experiences and feelings in order that others might provide help to them during the COVID-19 pandemic (Boserup et al., 2020). However, studies till date have not yet reported social media data for different violent behavior identification and reporting, especially during different phases of the pandemic. Addressing these research gaps, this study aimed to investigate the trends of different types of violence using Reddit data, additionally focusing on those posts that directly referred to the COVID-19 pandemic.

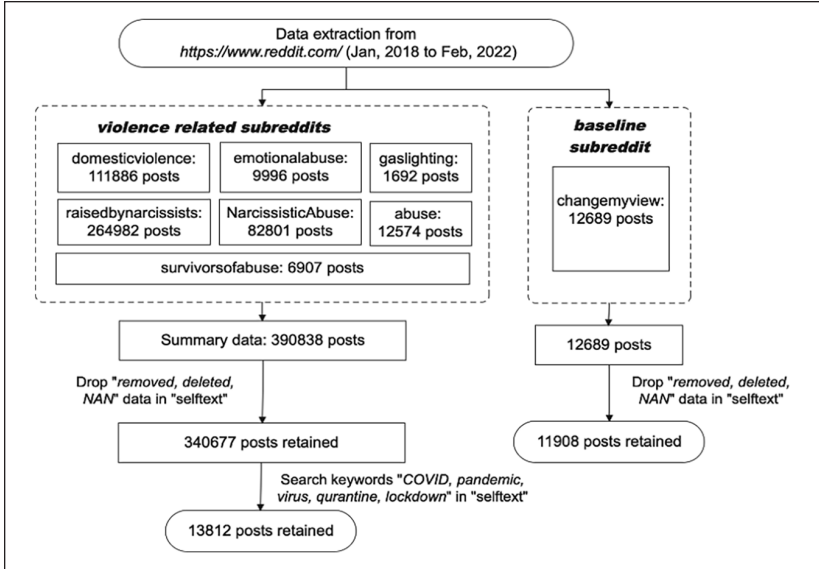
## Methods

### *Definitions of DV and IPV*

DV and IPV have not been consistently defined in previous studies. According to the US Department of Justice, DV refers to violence at the hands of a family member, roommate, intimate partner, or someone else in one's domestic setting (Yokotani & Takano, 2021). According to WHO (2017), IPV refers to violence at the hands of a current or former intimate partner (Barlett & Coyne, 2014). IPV in the US cuts across all genders, races, and social classes (Petrosky et al., 2017) and includes physical, financial, emotional, sexual, and psychological abuse (Larsen, 2016). As technology has evolved, other forms of violence have been characterized, such as coercive control and tech-facilitated violence or cyberviolence (PenzeyMoog & Slakoff, 2021). Based on these two definitions, researchers have argued that DV is a broader term that includes IPV, elder abuse, and child abuse in a family setting or cohabitation setting involving physical violence, sexual violence, and/or emotional violence (Boserup et al., 2020), whereas IPV represents physical violence, sexual violence, or emotional violence that occurs within the family, between intimate partners in cohabitation or non-cohabitation relationships. In this study, we followed the DV definition by US Department of Justice and the IPV definition by WHO.

### *Data Collection and Preprocessing*

To address our research goal, we first developed a classification tool to extract different types of violence from Reddit posts. We collected data from seven violence-related subreddits *r/domesticviolence*, *r/abuse*, *r/survivorsofabuse*, *r/emotionalabuse*, *r/NarcissisticAbuse*, *r/Gaslighting*, and *r/Raisedbynarcissists*



**Figure 1.** Data collection and preprocessing processes.

from January 1, 2018 to February 28, 2022. (For a description of each subreddit, please refer to Appendix A1.) Among these selected subreddits, four (*r/domesticviolence*, *r/abuse*, *r/survivorsofabuse*, and *r/emotionalabuse*) were chosen following previous work (Buchner et al., 2021). Three of the emotional violence-related subreddits (*r/NarcissisticAbuse*, *r/Gaslighting*, and *r/Raisedbynarcissists*) were chosen following our definitions of DV and IPV mainly because these subreddits are about the “gaslighting” and “narcissistic” behaviors which were potential keywords for emotional abuse. Data from subreddit *r/changemyview* were collected as a baseline, against which trends could be compared to the violence-related subreddits, again following previous work (Buchner et al., 2021). It was selected as a neutral comparison group because its members engage in discussions about personal opinions regarding political and social issues that are not directly related to violent behavior. Figure 1 shows the flowchart of data collection and data preprocessing procedures.

To be more specific, we applied a Python Reddit application programming interface wrapper to extract data for these subreddits and delete stop words (Balakrishnan & Lloyd-yemoh, 2014). Table 1 summarizes the number of posts for each subreddit and the COVID-19-related posts in each subreddit. To identify COVID-19-related posts, we used keyword searching.

**Table 1.** Summary of the Number of Posts for Each Subreddit.

| Subreddit Type             | Subreddit Name      | Original Posts | Retained Posts | COVID-19-Related Posts |
|----------------------------|---------------------|----------------|----------------|------------------------|
| Violence-related subreddit | domesticviolence    | 11,886         | 4,442          | 194                    |
|                            | emotionalabuse      | 9,996          | 9,443          | 356                    |
|                            | raisedbynarcissists | 264,982        | 252,914        | 10,434                 |
|                            | narcissisticabuse   | 82,801         | 55,651         | 2,290                  |
|                            | gaslighting         | 1,692          | 1,131          | 41                     |
|                            | survivorsofabuse    | 6,907          | 5,950          | 365                    |
|                            | abuse               | 12,574         | 11,146         | 132                    |
| Baseline subreddit         | changemyview        | 11,788         | 10,258         | —                      |

The keywords including *covid*, *pandemic*, *quarantine*, *virus*, *coronavirus*, *stay-at-home*, and *lockdown* to extract COVID-19-related posts in these violence subreddits.

## Data Analysis

**Data classification.** We used keyword searches to find posts containing violence-related keywords such as abuse, assault, punch, hurt, and so on to simplify the supervised learning procedure (the full keywords list is an integration of previous studies [28, 18] which we attach in Appendix Table A2). We found 203,097 posts with at least one of the abovementioned keywords (potentially violence-related posts), and 137,580 posts without any of the abovementioned keywords.

Next, we randomly selected 1,000 posts from the potentially violence-related posts and coded them into six subtypes including *DV*, *IPV*, *physical*, *sexual*, *emotional*, and *nonspecific violence and others* according to the coding book described in Appendix Table A3. Among all the randomly sampled posts, we conducted two rounds of coding. For the first round, we randomly sampled 200 posts for the pilot test, all of which were coded by two coders. In this round, the inter-annotator agreement (Cohen's kappa) of the two coders for DV, IPV, physical violence, sexual violence, emotional violence, and nonspecific violence and others were 0.67, 0.80, 0.84, 0.90, 0.72, and 0.80, respectively. Given the relatively low kappa scores for DV and emotional violence compared to other types of violence, we conducted a round of discussion to resolve differences, and improved these kappa values to 0.82 and 0.79, respectively, for the 200-sample data. To verify the coding results, the first 100 coded posts were coded by a third coder, giving a kappa value when

compared to the final decisions of the previous two coders for the six types of violence of 0.66, 0.62, 0.67, 0.79, 0.78, and 0.60, respectively. In the second round of coding, we coded the remaining 800 samples. The kappa values for the two coders are summarized in Table 2.

Then, we trained multiple machine learning classifiers, support vector machine,  $k$ -nearest neighbors (KNN), random forest (RF), decision tree, and naive Bayes for each type of violence. We chose these classifiers because they have previously been found to be effective in identifying IPV from social media data (Trinh Ha et al., 2022). For each type of violence, we applied 80% of the training data to train the classifiers and used the remaining 20% sampled data to calculate the accuracy of the classifiers. The best classifier for DV and nonspecific violence (or others) was KNN, the best classifier for PV and EV was the RF, and the best classifier for SV was the decision tree. The comparative performance summaries of classifier accuracies for each violence type are presented in Table 2.

**Definition of pandemic phases.** To define the phases of the pandemic for comparisons, we drew on time periods defined in previous articles. First, Leslie and Wilson (2020) compared the trends in DV service calls in 2019 and 2020 by separating data into two time periods with week 9 of 2020 as the separation date (Leslie & Wilson, 2020). To make the separation date more precise, we also referred to the data on confirmed cases of the COVID-19 pandemic in the US and separated the data into pre-pandemic and pandemic phases using the date February 26, 2020 (the third day of week 9 in 2020). The first nontravel US case had also been confirmed in California on February 26, 2020. Second, Savadori and Lauriola (2022) investigated the daily confirmed cases of coronavirus based on search trends of DV-related keywords (Savadori & Lauriola, 2022), and Baglivio et al. (2022) found that DV referrals were at their lowest from June 15, 2020 to June 21, 2020 (Baglivio et al., 2022). We therefore used the third day of this week, June 17, 2020, as a separation date point. Third, given that Google Trends of DV searches remained relatively stable after June 17, 2020, for the other two separation dates, we referred to the number of COVID-19-confirmed cases in the US and used two trough values of September 7, 2020 and June 4, 2021. The reason we used US timings is that around 50% of Reddit users are from the US (Davis & Graham, 2021). All five phases are illustrated and marked in Figure 2.

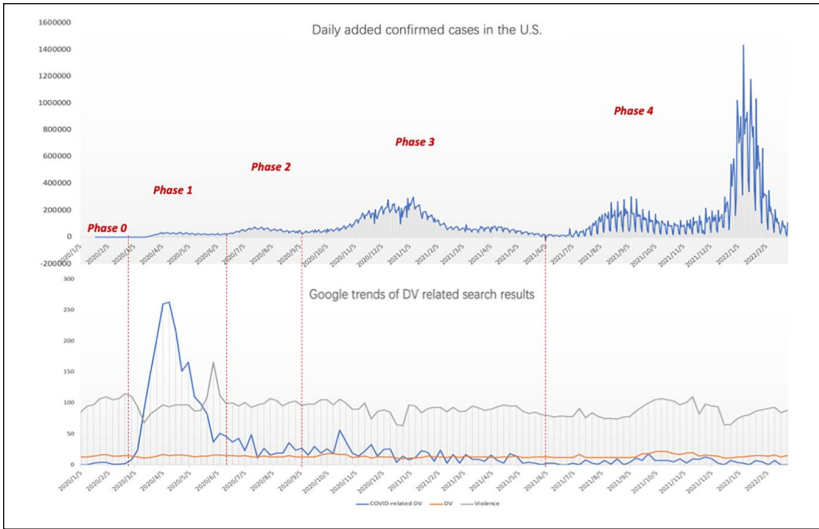
In addition, the keywords we applied for searching Google Trends and the way we presented the trends of COVID-related DV, DV, and violence are as follows. We applied the following keywords: *violence*, *abuse*, *domestic violence*, *domestic abuse*, *domestic violence COVID-19*, *domestic violence during coronavirus*, and *coronavirus domestic violence*. In Figure 2, the value of

**Table 2.** Summary of IAA, Manual Coding Results, Automatic Coding Results, and Best Performance Classifier and its Accuracy.

| Violence Type                      | IAA<br>(1,000<br>Samples) | Manual Labeling<br>Results (Label:<br>Counts) | Automatic Labeling<br>Results (Label:<br>Counts) | Best<br>Performance<br>Classifier | Accuracy |
|------------------------------------|---------------------------|---|--|-----------------------------------|----------|
| DV                                 | 0.80                      | 1: 452; 0: 548                                | 1: 87,803; 0: 115,294                            | SVM                               | 0.82     |
| IPV                                | 0.82                      | 1: 139; 0: 861                                | 1: 20,011; 0: 183,086                            | KNN                               | 0.94     |
| Physical violence                  | 0.84                      | 1: 344; 0: 656                                | 1: 8,616; 0: 194,481                             | RF                                | 0.89     |
| Sexual violence                    | 0.85                      | 1: 115; 0: 885                                | 1: 4,608; 0: 198,489                             | Decision Tree                     | 0.96     |
| Emotional violence                 | 0.78                      | 1: 307; 0: 693                                | 1: 51,347; 0: 151,750                            | RF                                | 0.86     |
| Nonspecific violence<br>and others | 0.80                      | 1: 94; 0: 906                                 | 1: 2,243; 0: 200,854                             | KNN                               | 0.96     |

Note. DV = domestic violence; IAA = inter-annotator agreement; IPV = intimate partner violence; KNN = k-nearest neighbors; RN = random forest; SVM = support vector machine.





**Figure 2.** Categories applied to the timing of the COVID-19 pandemic based on US trends in daily confirmed cases and DV-related searching keywords. DV=domestic violence.

*Violence* is the sum of searching results of the keywords *violence* and *abuse*. The value of *DV* is the sum of the search results of the keywords *domestic violence* and *domestic abuse*. The value of *COVID-related DV* is the sum of search results of the keywords *domestic violence COVID-19*, *domestic violence during coronavirus*, and *coronavirus domestic violence*.

*Increase rates of posts in subreddits.* To describe and compare changes in the number of posts from violence-related subreddits, the baseline subreddit, and across the whole of Reddit, we defined the increase rate (IR) of a subreddit (or certain type of subreddit) using equation (1):

$$IR_{pi} = \frac{AVP_{pi} - AVP_{pi-1}}{AVP_{pi-1}} \quad (1)$$

In equation (1),  $IR_{pi}$  indicates the IR of the posts of a subreddit at *Phase 1*.  $AVP_{pi}$  indicates the average number of posts per month of a subreddit at *Phase 1*. In this work, we considered IRs of two types of subreddits (violence-related subreddits and the baseline subreddit) and across the whole of the Reddit platform itself (Table 3). One point that needs noting is that to

**Table 3.** IRs of Posts Across Subreddits.

| Subreddit groups                         | IR <sub>p0</sub>   | IR <sub>p1</sub>  | IR <sub>p2</sub>  | IR <sub>p3</sub>  | IR <sub>p4</sub> |
|--|--------------------|-------------------|-------------------|-------------------|------------------|
|  | Years<br>2018–2019 | Phases<br>0–1     | Phases<br>1–2     | Phases<br>2–3     | Phases<br>3–4    |
| Violence-related subreddits              | 39.9%<br>(6146.7)  | 25.7%<br>(7728)   | 11.4%<br>(8608.7) | 6.8%<br>(9198.2)  | 5.9%<br>(9744.1) |
| Violence-related subreddits (normalized) | —<br>(0.328)       | 90.1%<br>(0.623)  | 26.4%<br>(0.788)  | 13.98%<br>(0.898) | 11.4%<br>(1)     |
| Baseline subreddit (changemyview)        | 4095.8%<br>(115.4) | 106.8%<br>(238.6) | −6.4%<br>(233.3)  | 28.3%<br>(286.5)  | 5.7%<br>(302.8)  |
| Baseline subreddit (normalized)          | —<br>(0.376)       | 109.3%<br>(0.786) | −2.2%<br>(0.768)  | 23.1%<br>(0.946)  | 5.7%<br>(1)      |
| All of Reddit                            | 18.9%              | 23.1%             | 34.4%             | 32.3%             | 30.3%            |

Note. For baseline and violence-related subreddits, the average number of monthly posts are 2.7 and 4392.25, respectively. For IR<sub>p0</sub>, the entire month of January 2020 was included in the data of year 2019 because of our data phases definitions.

calculate IR<sub>0</sub>, we divided *Phase 0* into two subperiods. The first period was the whole year of 2018, and the second period was from January 1, 2019 to January 31, 2020. Then, IR<sub>0</sub> is equal to the average monthly posts of the second subperiod minus the average monthly posts of the first subperiod divided by the average monthly posts of the first subperiod. For the IR of the whole of Reddit, detailed calculations of each phase are summarized in Appendix A3 and Table A2. We also calculated the IRs of various types of violence and of posts that identified specific types of violence as well as mentioned COVID-19-related keywords (details are in Tables 4 and 5).

## Results

### *Trends and IRs of Violence-Related Subreddits and the Baseline Subreddit*

We analyzed the trends of violence-related subreddits and the baseline subreddit (the changemyview) as shown in Figure 3. In this figure, the *x*-axis shows the date, and the *y*-axis shows the number of posts of the violence-related subreddits we applied and baseline subreddit (changemyview).

To investigate further the differences in trends of the seven violence-related subreddits, baseline subreddits, and Reddit itself, we calculated the IRs for the five different phases and the results are summarized in Table 3.

**Table 4.** IRs of Various Types of Violence-Related Posts.

| Violence types and baseline     | IR <sub>p0</sub>  | IR <sub>p1</sub>  | IR <sub>p2</sub>  | IR <sub>p3</sub> | IR <sub>p4</sub>  |
|---------------------------------|-------------------|-------------------|-------------------|------------------|-------------------|
|                                 | Years 2018–2019   | Phases 0–1        | Phases 1–2        | Phases 2–3       | Phases 3–4        |
| DV                              | 43.3%<br>(1580.1) | 26.9%<br>(2005.8) | 7.97%<br>(2165.7) | 1.3%<br>(2193)   | 0.94%<br>(2213.6) |
| IPV                             | 55.7%<br>(279.2)  | 58.8%<br>(443.4)  | 17.1%<br>(519)    | 16.3%<br>(603.5) | 3.5%<br>(624.8)   |
| Emotional violence              | 48.8%<br>(884.8)  | 28.8%<br>(1139.2) | 13.9%<br>(1297.7) | 2.7%<br>(1332.1) | 3.8%<br>(1382.7)  |
| Physical violence               | 64.3%<br>(149.6)  | 45.7%<br>(218)    | 4.3%<br>(227.3)   | 5.24%<br>(239.3) | -12.1%<br>(210.2) |
| Sexual violence                 | 49.8%<br>(90.8)   | 16.8%<br>(106)    | 4.7%<br>(111)     | -2.9%<br>(107.8) | 0.3%<br>(108.1)   |
| Nonspecific violence and others | 14.9%<br>(29.8)   | 63.3%<br>(48.6)   | 18.0%<br>(57.3)   | 12.3%<br>(64.4)  | 6.0%<br>(68.2)    |
| Reddit itself                   | 18.9%             | 23.1%             | 34.4%             | 32.3%            | 30.3%             |

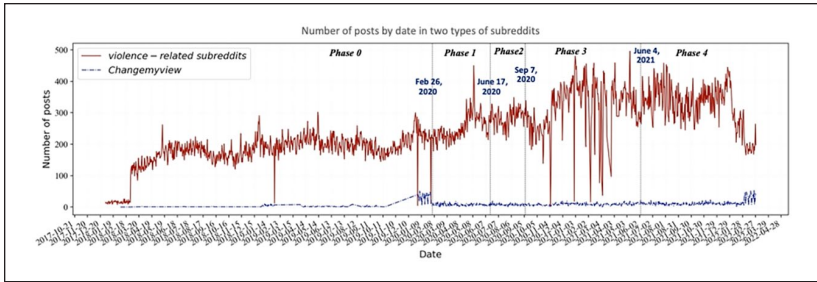
Note. DV = domestic violence; IPV = intimate partner violence. The calculation of Phase 0 data is the same as mentioned in the notes of Table 3. The average number of monthly posts of DV, IPV, EV, PV, SV, and nonspecific violence in the year 2018 are 1103, 179.25, 594.75, 91.08, 60.58, and 25.92, respectively.

**Table 5.** IRs of Various Types of Violence-Related Posts that Mention COVID-19.

| IRs                             | IR <sub>p2</sub> | IR <sub>p3</sub> | IR <sub>p4</sub> |
|---------------------------------|------------------|------------------|------------------|
|                                 | Phases 1–2       | Phases 2–3       | Phases 3–4       |
| DV                              | -6.4% (179)      | -0.6% (177.9)    | -35.5% (114.7)   |
| IPV                             | -8.6% (30.3)     | 36.8% (41.5)     | -39.8% (25)      |
| Emotional violence              | -4.3% (96.7)     | 10.3% (106.6)    | -37.3% (66.9)    |
| Physical violence               | -6.1% (16.3)     | 13.3% (18.5)     | -50.8% (9.1)     |
| Sexual violence                 | -20.1% (7.7)     | 22.3% (9.4)      | -15.9% (7.9)     |
| Nonspecific violence and others | -16.7% (3.7)     | 33.0% (4.9)      | -31.6% (3.3)     |

Note. DV = domestic violence; IPV = intimate partner violence.

For clarity, we also summarize the average number of monthly posts in parentheses after the IR value. To eliminate the data imbalance between seven violence-related subreddits and one baseline subreddit, we adopted the widely used min-max normalization method to normalize the average



**Figure 3.** Number of posts per day of the violence-related subreddits and the baseline subreddit.

number of monthly posts among all phases (Jain et al., 2005). Similarly, the normalized average monthly posts are summarized in parentheses after the normalized IR values.

Figure 3 and Table 3 show that in the pre-pandemic phase (*Phase 0*), the IR of the baseline subreddit was the largest among all three types of subreddits. This is because, in the year 2018, there were fewer posts on the baseline subreddit (the average number of monthly posts was only 2.7). For the violence-related subreddits, the average number of monthly posts in 2018 was 4392.25 and the average number of monthly posts in 2019 (including the first month of 2020) was 6146.7. Based on the normalized data, after the outbreak of the COVID-19 pandemic (*Phase 1*), the IR of violence-related subreddits was similar to the baseline subreddit but larger than all of Reddit. In *Phase 2*, the baseline subreddit's IR was negative, and the IR of the violence-related subreddits was positive but lower than the IR of Reddit itself. In *Phase 3*, the IR of violence-related subreddits was smaller than the baseline subreddit and Reddit itself.

### *Trends and IRs of Different Types of Violence*

To investigate further trends of different types of violence on Reddit, we visualized these in Figure 4 based on the classification results from section "Data Classification," and summarized the IRs in Table 4. In Figure 4, the *x*-axis shows the date, and the *y*-axis shows the number of posts that were classified as DV, IPV, physical violence, emotional violence, sexual violence, and nonspecific violence and others (we label it as "others" in the figure). In Table 4, we summarized the IRs (average number of posts) of all the six types of violence in each phase ( $IR_{p0}$  to  $IR_{p4}$ ). For comparison, we also summarized the IRs of Reddit itself in Table 4.

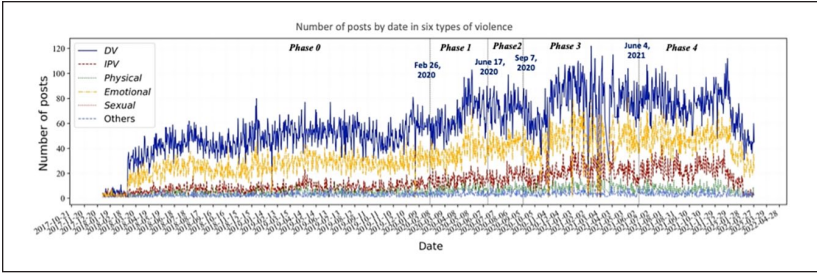


Figure 4. Trends of different types of violence.

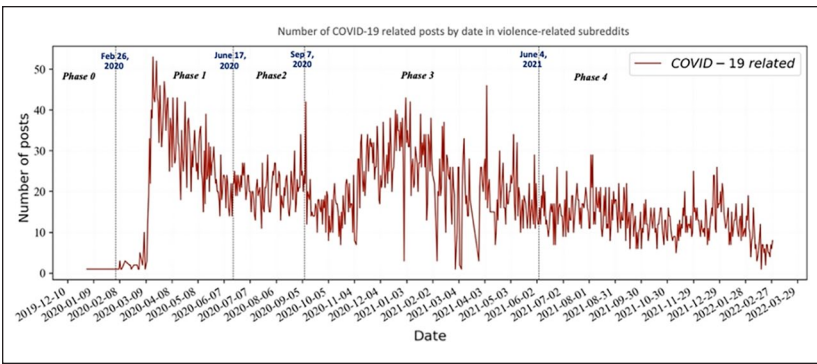


Figure 5. Trends of COVID-19-related posts.

In *Phase 0*, the IR of physical violence was the largest, followed by those for IPV, emotional violence, and DV, all of which were larger than the IR for Reddit overall. From *Phase 0* to *Phase 1*, the IR of IPV was the largest compared with other types of violence, and IRs for physical violence (45.7%), emotional violence (28.8%), and DV (26.9%) were also larger than the general Reddit IR (23.1%). In the later phases (*Phases 2, 3, and 4*), the IRs were all smaller than the general Reddit IR, although most showed increases, except for the *physical violence* from *Phase 3* to *Phase 4* and *sexual violence* from *Phase 2* to *Phase 3*.

### Trends and IRs of Posts Mentioning COVID-19

Figure 5 shows the trends in posts mentioning both violence and COVID-19. IRs for specific violence terms are displayed in Table 5 and the corresponding

number of posts is indicated in parentheses. The data were restricted to *Phases 1–4* owing to minimal eligible posts in *Phase 0*.

Table 5 shows that there was a decrease in the IRs for all types of violence associated with COVID-19 mentions from *Phase 1* to *Phase 2* followed by an increase from *Phase 2* to *Phase 3* for most types, followed by a decrease again in all types from *Phase 3* to *Phase 4*.

## Discussion

### *Role of Social Media in Investigating Trends in Different Types of Reported Violence*

We carried out this descriptive study on the assumption that social media such as Reddit may provide alternative outlets for people to report their experiences of violence, especially during the COVID-19 pandemic. Thus, there may be a potential for using social media data as an ancillary information source for public health surveillance, because quarantine periods influenced people's daily life substantially and more time on an average was spent on social media (Ellis et al., 2021). Other evidence has suggested that the COVID-19 pandemic and associated social responses potentially empowered abusive individuals to control their partners and placed the survivors into more dangerous situations because of the lack of formal and informal support during quarantine periods (Wood et al., 2021). People who have experienced abuse may sustain substantial psychological trauma, which may manifest in a negative internalizing or externalizing behavior (Greenwald, 2002). During the pandemic, concerns about increases in self-harming behavior and suicidal ideation were heightened due to the increased risk of abusive experiences (Job et al., 2020). Thus, using social media data to monitor reported experiences of violence (mainly of victims) during the COVID-19 pandemic could help researchers and authorities design and deliver preventive interventions to mitigate the negative pandemic impact on potential victims. Using Reddit data, this study successfully trained several machine learning classifiers to identify six types of violence and was, therefore, able to describe the trends in different types of reported violence across different phases of the COVID-19 pandemic.

Specifically, by calculating and describing IRs for six types of violence and comparing them with Reddit posts in total and one baseline subreddit (*r/changemyview*) as a control measure, we found that IPV showed particularly marked trends as a reported experience on Reddit compared to other specific types of violence. The increase in IPV can potentially be explained as an impact of lockdown policy of trapping victims with their abusers, especially

female victims (Wood et al., 2021). Such findings are consistent with a theory of exposure reduction which suggests that victims and their abusers had a high risk of conflict during quarantine (Hsu & Henke, 2021). Furthermore, during quarantine, more victims needed to live with their abusers, and the theory of exposure posits that people commit violence when they are free to do so (Dugan et al., 2003). The findings of this study support the development of online interventions, extracting these IPV-related posts and providing necessary emotional/informational support to victims using Reddit, especially during future circumstances that mimic *Phase 1* of the pandemic. Moreover, although the IRs of DV between different phases were not the highest, the average number of monthly reported posts of DV was the largest of all types of violence. Due to the large number of DV reports, the unmet mental health needs of Reddit users should be taken into consideration, particularly as an association has been reported between DV and increased risk of mental health problems within 4 weeks of the pandemic (Liu et al., 2021).

Around 50% of Reddit users are from the US (Davis & Graham, 2021); the results of this study, therefore, are more likely to reflect the violence experience of US citizens mainly in 2020 and 2021. Also, the results may reflect a younger, more tech-literate population since they are the regular Reddit users (Zhang et al., 2021). Based on the observed trends in different types of reported violence with mentions of COVID-19, we found that pandemic-related DV and emotional violence were at a high level during *Phase 1*. This could be attributed to the US Government's response to the pandemic in New York beginning with a full lockdown from March 2020 to April 2020, followed by a four-phase reopening plan by region from April 2020 to July 2020. During this phase, people experienced a full or partial lockdown policy which has been reported to be associated with increased violent behavior (Hsu & Henke, 2021). Later, from *Phase 3* to *Phase 4*, the IRs were negative, potentially as a result of vaccinations and less rigorous quarantine or other COVID-19-related policies compared to previous time periods. The intermediate positive IRs of violence posts including COVID-19 mentions from *Phase 2* to *Phase 3* for most violence types corresponds to the fact that US Government announced further social distance requirements for travelers to quarantine after an international travel. Also, the average daily number of confirmed coronavirus cases during this period was at a higher level compared with the previous phases.

The findings of this study suggest that government agencies and related nongovernmental organizations (NGOs) should pay attention to violence potentially resulting from public health policies. For instance, in the case of a crisis that needs quarantine, measuring the IR of a certain type of violence is necessary. If the IR value is large during a certain phase, the authorities

should apply proper interventions/policies to help those potential victims (i.e., IPV in *Phase 1* during the COVID-19 pandemic). Moreover, by calculating the mean value of monthly posts of DV and IRs of DV across social media data, authorities may be better able to monitor occurrence rates of specific types of violence (e.g., IPV, DV, emotional violence, sexual violence, or physical violence) and provide timely help to potential victims. In summary, social services and social welfare institutions can learn from the results of this and similar studies to develop responses to future epidemics or other major disasters which could increase the risk of violent behavior. A future surveillance system might therefore obtain social media data on different types of violence, capture reporting trends, and take appropriate steps to increase resources for preventive interventions.

### *Limitations and Future Study*

Despite the potential insights of this study, it has several limitations which need to be considered when drawing interpretations. First, we only considered five specific types of violence and categorized other types of violence such as financial violence and tech-facilitated violence into the type of *nonspecific and other violence*. Second, the separation of time phases of this study is inevitably approximate and arbitrary (although attempting to replicate previous practice by other researchers); therefore, our insights are limited to four phases during the pandemic. Future studies could separate time phases according to their research needs or intervention needs. Third, the keywords we applied to identify potential violence-related posts were limited to those deployed by three previous studies. In the future, a more comprehensive keywords list needs to be created. Fourth, the results can show the general trends of violent behavior discussions, but not rule out changes in Reddit user behavior due to the influence of the pandemic. For example, some people have increased their use of Reddit because of the pandemic, but some people, such as healthcare workers, may have reduced their use of Reddit. Last, although the discussions in this study are primarily focused on US Reddit users, as they were the predominant group within our data sample, it should be noted that the structure of Reddit's user base is subject to change over time, thus limiting the generalizability of our findings.

### **Conclusion**

Social media platforms such as Reddit provide useful sites for people to report violent experiences, especially during periods of relative social



isolation, such as those experienced during the COVID-19 pandemic when other tools and resources are limited. Uncovering trends in different types of violence between different phases of the pandemic may present opportunities for statutory authorities and/or NGOs—both as a potential method for surveillance and possibly as an impetus for the provision of specific online interventions or support. This study concludes that the number of posts related to IPV increased the most among other types of violence in *Phase 1* when the pandemic was in its early phase. Then in the later *Phase 3*, the increase in rate of IPV related to COVID-19 was the largest which suggests the high risk of IPV accompanying the pandemic. Meanwhile, the positive increases in reported DV, IPV, and emotional violence across all phases on Reddit indicate the necessity of identifying and providing help to those who suffer from such experiences, especially in *Phase 1*. Future studies are needed to further understand the reasons that different types of reported violence increased at different rates at different phases of the pandemic and how interventions might be feasibly designed and evaluated.

## Appendix A

### A1. Description of the Referred Subreddits

The following are subreddits used and their description on the Reddit platform:

- *r/domesticviolence*: “Information and support for victims, survivors, their friends, and family.”
- *r/emotionalabuse*: “A subreddit for people who are or were affected by emotional abuse. Share your stories and we can help each other overcome our pains, sometimes by the use of internet hugs.”
- *r/NarcissisticAbuse*: “This is a place for victims of narcissistic abuse to come together to support, encourage, learn from, share with, and validate one another. It is NOT a replacement for therapy or counseling. We support abusive relationships from romance, work, roommates, and friends. No family content except for you, your partner/ex and your children (no abusive kids). All others will be removed including mentions like family/friends or childhood content. No demands to explain our rules or ‘just curious’ requests.”
- *r/Gaslighting*: no specific introduction to the community.
- *r/Raisedbynarcissists*: “This is a support group for people raised by abusive parents (with toxic, self-absorbed, or abusive personality

traits, which may be exhibited by those who suffer from cluster B personality disorders). Please share your stories, your questions, your histories, your fears, and your triumphs. Significant others and friends are all welcome.”

- r/abuse: “the subreddit data are used by previous work (Babvey et al., 2021); currently, we can only download its data (from the year 2018 to the year 2021) but cannot reach it out anymore.”
- r/survivorsofabuse: a private community, “this is a subreddit designed to give a place for survivors of all abuse to come together to share their stories, vent, and assist one another in healing. We want to celebrate success, give acceptance, and support one another. The only rule is to respect one another and understand how very difficult it is to expose such a vulnerable part of one’s life. This will be a \*safe place.”

## A2. Keywords List

To find out the potentially violence-related content, we applied a keyword searching technique. The keywords are mainly from two previous studies. Botelle et al. (2022) applied 17 violence-related keywords to extract domestic, physical, and/or sexual violence-related text from clinical data [28]. Violence-related subreddits were among the topics with the highest growth after the COVID-19 outbreak. Babvey et al. (2021) used Twitter and Reddit data to access children’s exposure to violence during the COVID-19 pandemic, they listed the keywords related to intimate partner abuse, physical abuse, sexual abuse, and child abuse [18]. We applied these keywords in this study.

**Table A1.** Keywords Used in Finding Potentially Violence-Related Topics.

| Subjects                             | Keywords   |
|--------------------------------------|--|
| Violence related from the literature | abuse, abusive, assault, sex, sexual, abused, violence, violent, punch, violent, hurt, hit, beat, insult, choke, domestic, attack, stalk, manipulate, sexual, physical, mental, poorly treated, verbal, emotional, gaslighting, narcissism, victim, witness, injured, painful, verbal, pull, beat, throw, grab, neck, bleed, arm, remove, glass, smash, dish, bruise, weed, couch, rap, weird, porn, naked, nude, and mistreat |

(continued)

**Table A1. (continued)**

| Subjects                                | Keywords   |
|---|--|
| Nominated keywords                      | emotional violence, psychological, psychological violence, coercive, and control   |
| Relationship related (literature based) | Relationship, family, parent, dad, mom, husband, wife, boyfriend, girlfriend, co-parent, sister, brother, mother, father, cohabitate, my ex, lover, intimate, partner, stepmom, and kiss |

### A3. IRs Across the Whole of Reddit

To calculate the IRs across the whole of Reddit, we collected data from two statistical reports on Reddit (<https://www.oberlo.com/blog/reddit-statistics>; <https://blog.hubspot.com/marketing/reddit-stats>) and summarized the number of posts and number of active users per year. Based on these two values, we calculated the IR of the whole of Reddit. Due to the missing number of posts for 2017, we use the IR of “Number of active users” as a representation. For the phase(s) including the years 2020 and 2021, we use the average number of IR in 2020 and 2021.

**Table A2.** IR Calculations of Reddit Using Official-Reported Data.

| Year      | Increasing Rate (%) | Increasing NUMBER | Number of Posts (Million) | Number of Active Users (Million) |
|-----------|---------------------|-------------------|---------------------------|----------------------------------|
| 2017      | /                   | /                 | 124 (around)              | 250                              |
| 2018      | 18.9                | 29                | 153                       | 330 (32% increase)               |
| 2019      | 23.1                | 46                | 199                       | 430                              |
| 2020      | 34.4                | 104.4             | 303.4                     | /                                |
| 2021      | 30.3                | 131.6             | 435                       | /                                |
| 2020/2021 | 32.4                | /                 | /                         | /                                |

IRs = increase rates.

### A4. Coding Book

We need to classify the collected data based on the violence type and the role of the poster. First, the violence type may include *domestic, emotional (verbal), physical, sexual, and intimate partner violence (IPV), or nonspecific violence and others*. The specific coding rules and the definition of each type are shown in Table A3.

**Table A3.** Definitions of Each Type of Violence and the Coding Rules.

| Violence Types                  | Label                         | Definition   | Examples   |
|---------------------------------|-------------------------------|--|--|
| DV                              | 1: yes;<br>0: no/<br>not sure | Domestic violence or other abuse that occurs in a domestic setting, such as in a marriage or cohabitation. It centers on violence from intimate partners and other family members but may extend beyond this   | I was abused by my mother/father/wife/husband/ex-wife/ex-husband/brother/sister/aunt, etc.   |
| IPV                             | 1: yes;<br>0: no/<br>not sure | Similar to domestic violence, but the violence is limited to intimate partners. In this type of post, my ex, boyfriend, girlfriend, wife, husband, lover, and intimate relationship are often can be found   | I am a 29F. I was in a 2 year-long abusive relationship with a 28M, it included <i>emotional, financial, and physical</i> abuse, and one time it almost escalated to <i>sexual abuse</i>                                   |
| Physical                        | 1: yes;<br>0: no/<br>not sure | Posts that mainly contain the following words: punch, hit, hurt, beat, insult, panic, choke, manipulate, attacked, stalk, physically abused, physically harmed, or assault, etc.   | I've been hit, spit on, had a gun pulled on me, gaslighted, financially abused, sexually abused, emotionally abused. I not only went out with that level of abusive horrible guy, but the more subtle, emotionally abusive |
| Sexual                          | 1: yes;<br>0: no/<br>not sure | Posts that contain the words sexually abuse/harm/hurt/forced/, have sex without my permission. This type of posts often contains the following terms: my ex, boyfriend, girlfriend, wife, husband, lover, intimate assault, or sex (abuse, violence) | 1. A few days ago. I was sexually assaulted by a former friend.<br>2. I get now that I was repeatedly raped and sexually assaulted so many times. I never defined what went on as that until recently                      |
| Nonspecific violence and others | 1: yes;<br>0: no/<br>not sure | If they only mention that they are/were/had going through an abusive relationship but not clearly motioned which type of violence, it belongs to this type   | 1. I was mistreated and abused by my family and friends<br>2. A member of my community has just left their abusive household/my abusive brother and the effects it had on me.  |
| Emotional                       | 1: yes;<br>0: no/<br>not sure | Posts mention that some were going through an emotionally abusive relationship, or include verbal violence/abuse, emotional abuse/harm, etc., also gaslighting, narcissism, emotionally manipulate   | My first high school boyfriend, who was 3 years older than me (a large gap for a teenager) was emotionally abusive towards me, and I was too young/inexperienced to understand that's what it was                          |

Note. DV= domestic violence; IPV = intimate partner violence.

## Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

## Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: AR and RS are funded in part by the National Institute for Health Research (NIHR) Biomedical Research Centre at the South London and Maudsley NHS Foundation Trust and King's College London. AR and RS are supported by Health Data Research UK, an initiative funded by U.K. Research and Innovation, Department of Health and Social Care (England) and the devolved administrations and leading medical research charities. RS is additionally funded in part by the NIHR Applied Research Collaboration South London (NIHR ARC South London) at King's College Hospital NHS Foundation Trust, and by the DATAMIND HDR UK Mental Health Data Hub (MRC grant MR/W014386). The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health and Social Care.

In addition, this research was supported by the UK Prevention Research Partnership (Violence, Health and Society; MR-VO49879/1), which is funded by the British Heart Foundation, Chief Scientist Office of the Scottish Government Health and Social Care Directorates, Engineering and Physical Sciences Research Council, Economic and Social Research Council, Health and Social Care Research and Development Division (Welsh Government), Medical Research Council, National Institute for Health and Care Research, Natural Environment Research Council, Public Health Agency (Northern Ireland), The Health Foundation, and Wellcome. The views expressed in this article are those of the authors and not necessarily those of the UK Prevention Research Partnership or any other funder.

## ORCID iD

Lifang Li  <https://orcid.org/0000-0003-4917-5701>

## References

- Babvey, P., Capela, F., Cappa, C., Lipizzi, C., Petrowski, N., & Ramirez-Marquez, J. (2021). Using social media data for assessing children's exposure to violence during the COVID-19 pandemic. *Child Abuse and Neglect, 116*, 104747. <https://doi.org/10.1016/j.chiabu.2020.104747>
- Baglivio, M. T., Wolff, K. T., Reid, J. A., Jackson, S. L., & Piquero, A. R. (2022). Did Juvenile domestic violence offending change during COVID-19? *Youth Violence and Juvenile Justice, 20*(1), 63–79. <https://doi.org/10.1177/15412040211047266>
- Balakrishnan, V., & Lloyd-yemoh, E. (2014). Stemming and lemmatization: A comparison of retrieval performances. *Lecture Notes on Software Engineering, 2*(3), 174–179.

- Barlett, C., & Coyne, S. M. (2014). A meta-analysis of sex differences in cyber-bullying behavior: The moderating role of age. *Aggressive Behavior, 40*(5), 474–488. <https://doi.org/10.1002/ab.21555>
- Bettinger-Lopez, C., & Bro, A. (2020). *A double pandemic: Domestic violence in the age of COVID-19*. Council on Foreign Relations.
- Bhavsar, V., Kirkpatrick, K., Calcia, M., & Howard, L. M. (2021). Lockdown, domestic abuse perpetration, and mental health care: Gaps in training, research, and policy. *The Lancet Psychiatry, 8*(3), 172–174. [https://doi.org/10.1016/S2215-0366\(20\)30397-7](https://doi.org/10.1016/S2215-0366(20)30397-7)
- Boserup, B., McKenney, M., & Elkbuli, A. (2020). Alarming trends in US domestic violence during the COVID-19 pandemic. *American Journal of Emergency Medicine, 38*(12), 2753–2755. <https://doi.org/10.1016/j.ajem.2020.04.077>
- Botelle, R., Bhavsar, V., Kadra-Scalzo, G., Mascio, A., Williams, M. V., Roberts, A., Velupillai, S., & Stewart, R. (2022). Can natural language processing models extract and classify instances of interpersonal violence in mental healthcare electronic records: An applied evaluative study. *BMJ Open, 12*(2), 1–10.
- Buchner, V., Hamm, S., Medenica, B., & L Molendijk, M. (2021). Linguistic analysis of online domestic violence testimonies in the context of COVID-19. *Sage Open, 13*, 21582440221146135.
- Campbell, A. M. (2020). An increasing risk of family violence during the Covid-19 pandemic: Strengthening community collaborations to save lives. *Forensic Science International: Reports, 2*, 100089. <https://doi.org/10.1016/j.fsir.2020.100089>
- Davis, J. L., & Graham, T. (2021). Emotional consequences and attention rewards: The social effects of ratings on Reddit. *Information Communication and Society, 24*(5), 649–666. <https://doi.org/10.1080/1369118X.2021.1874476>
- Dugan, L., Rosenfeld, R., & Nagin, D. S. (2003). Exposure reduction or retaliation? The effects of domestic violence resources on intimate-partner homicide. *Law and Society Review, 37*(1), 169–198. <https://doi.org/10.1111/1540-5893.3701005>
- Ellis, A., Briggs, D., Lloyd, A., & Telford, L. (2021). A ticking time bomb of future harm: Lockdown, child abuse and future violence. *Abuse: An International Impact Journal, 2*(1), 37–48. <https://doi.org/10.37576/abuse.2021.017>
- Emezue, C. (2020). Digital or digitally delivered responses to domestic and intimate partner violence during COVID-19. *JMIR Public Health and Surveillance, 6*(3), e19831. <https://doi.org/10.2196/19831>
- Ertan, D., El-Hage, W., Thierrée, S., Javelot, H., & Hingray, C. (2020). COVID-19: Urgency for distancing from domestic violence. *European Journal of Psychotraumatology, 11*(1), 1800245. <https://doi.org/10.1080/20008198.2020.1800245>
- Goh, K. K., Lu, M. L., & Jou, S. (2020). Impact of COVID-19 pandemic: Social distancing and the vulnerability to domestic violence. *Psychiatry and Clinical Neurosciences, 74*(11), 612–613. <https://doi.org/10.1111/pcn.13130>
- Greenwald, R. (2002). The role of trauma in conduct disorder. *Trauma and Juvenile Delinquency: Theory, Research, and Interventions, 6771*(2002), 5–24. <https://doi.org/10.1300/J146v06n01>

- Hajek, A., & König, H. H. (2021). Social isolation and loneliness of older adults in times of the COVID-19 pandemic: Can use of online social media sites and video chats assist in mitigating social isolation and loneliness? *Gerontology*, 67(1), 121–123. <https://doi.org/10.1159/000512793>
- Hamilton, J. L., Nesi, J., & Choukas-Bradley, S. (2020). Staying socially connected while physically distant. *Journal of Adolescent Health*, 13(2), 1–35.
- Hedges, L. V., Pustejovsky, J. E., & Shadish, W. R. (2012). A standardized mean difference effect size for single case designs. *Research Synthesis Methods*, 3(3), 224–239. <https://doi.org/10.1002/jrsm.1052>
- Hsu, L. C., & Henke, A. (2021). COVID-19, staying at home, and domestic violence. *Review of Economics of the Household*, 19(1), 145–155. <https://doi.org/10.1007/s11150-020-09526-7>
- Job, E., Steptoe, A., & Fancourt, D. (2020). Abuse, self-harm and suicidal ideation in the UK during the COVID-19 pandemic. *British Journal of Psychiatry*, 217(4), 543–546. <https://doi.org/10.1192/bjp.2020.130>
- Jain, A., Nandakumar, K., & Ross, A. (2005). Score normalization in multimodal biometric systems. *Pattern Recognition*, 38(12), 2270–2285. <https://doi.org/10.1016/j.patcog.2005.01.012>
- Koh, J. X., & Liew, T. M. (2022). How loneliness is talked about in social media during COVID-19 pandemic: Text mining of 4,492 Twitter feeds. *Journal of Psychiatric Research*, 145, 317–324. <https://doi.org/10.1016/j.jpsychires.2020.11.015>
- Larsen, M. (2016). Health inequities related to intimate partner violence against women: The role of social policy in the United States, Germany, and Norway. In R. J. Angel & W. R. Avison (Eds.), *Health Inequities Related to Intimate Partner Violence Against Women* (pp. 1–244). Springer International Publishing.
- Leslie, E., & Wilson, R. (2020). Sheltering in place and domestic violence: Evidence from calls for service during COVID-19. *Journal of Public Economics*, 189, 104241. <https://doi.org/10.1016/j.jpubeco.2020.104241>
- Liu, M., Xue, J., Zhao, N., Wang, X., Jiao, D., & Zhu, T. (2021). Using social media to explore the consequences of domestic violence on mental health. *Journal of Interpersonal Violence*, 36(3–4), NP1965–NP1985. <https://doi.org/10.1177/0886260518757756>
- Lyons, M., & Brewer, G. (2021). Experiences of intimate partner violence during lockdown and the COVID-19 pandemic. *Journal of Family Violence*, 37, 969–977. <https://doi.org/10.1007/s10896-021-00260-x>
- Matoori, S., Khurana, B., Balcom, M. C., Koh, D. M., Froehlich, J. M., Janssen, S., Kolokythas, O., & Gutzeit, A. (2020). Intimate partner violence crisis in the COVID-19 pandemic: How can radiologists make a difference? *European Radiology*, 30(12), 6933–6936. <https://doi.org/10.1007/s00330-020-07043-w>
- Mayshak, R., Curtis, A., Coomber, K., Tonner, L., Walker, A., Hyder, S., Likhaitzky, P., & Miller, P. (2022). Alcohol-involved family and domestic violence reported to police in Australia. *Journal of Interpersonal Violence*, 37(3–4), NP1658–NP1685. <https://doi.org/10.1177/0886260520928633>
- PenzeyMoog, E., & Slakoff, D. C. (2021). As technology evolves, so does domestic violence: Modern-day tech abuse and possible solutions. In J. Bailey, A. Flynn, &

- N. Henry (Eds.), *The Emerald International Handbook of Technology-Facilitated Violence and Abuse* (pp. 643–662). Emerald Publishing Limited. <https://doi.org/10.1108/978-1-83982-848-520211047>
- Petrosky, E., Blair, J. M., Betz, C. J., Fowler, K. A., Jack, S. P. D., & Lyons, B. H. (2017). Racial and ethnic differences in homicides of adult women and the role of intimate. *Morbidity and Mortality Weekly Report*, *66*(28), 741–746. [https://www.cdc.gov/mmwr/cme/conted\\_info.html#weekly](https://www.cdc.gov/mmwr/cme/conted_info.html#weekly)
- Savadori, L., & Lauriola, M. (2022). Risk perceptions and COVID-19 protective behaviors: A two-wave longitudinal study of epidemic and post-epidemic periods. *Social Science & Medicine*, *301*, 114949. <https://doi.org/10.1016/j.socscimed.2022.114949>
- Sharifi, M., Asadi-Pooya, A. A., & Mousavi-Roknabadi, R. S. (2020). Burnout among healthcare providers of COVID-19; a systematic review of epidemiology and recommendations. *Archives of Academic Emergency Medicine*, *9*(1), e7. <https://doi.org/10.22037/aaem.v9i1.1004>
- Sri, A. S., Das, P., Gnanapragasam, S., & Persaud, A. (2021). COVID-19 and the violence against women and girls: 'The shadow pandemic.' *International Journal of Social Psychiatry*, *67*(8), 971–973. <https://doi.org/10.1177/0020764021995556>
- Trinh Ha, P., D'Silva, R., Chen, E., Koyutürk, M., & Karakurt, G. (2022). Identification of intimate partner violence from free text descriptions in social media. *Journal of Computational Social Science*, *5*, 1207–1233. <https://doi.org/10.1007/s42001-022-00166-8>
- Voth Schrag, R. J., Leat, S., Backes, B., Childress, S., & Wood, L. (2022). "So many extra safety layers:" Virtual service provision and implementing social distancing in interpersonal violence service agencies during COVID-19. *Journal of Family Violence*, *38*, 227–239. <https://doi.org/10.1007/s10896-021-00350-w>
- Wood, L., Baumler, E., Schrag, R. V., Guillot-Wright, S., Hairston, D., Temple, J., & Torres, E. (2021). "Don't know where to go for help": Safety and economic needs among violence survivors during the COVID-19 Pandemic. *Journal of Family Violence*, *37*, 959–967. <https://doi.org/10.1007/s10896-020-00240-7>
- Wood, L., Schrag, R. V., Baumler, E., Hairston, D., Guillot-Wright, S., Torres, E., & Temple, J. R. (2020). On the front lines of the COVID-19 pandemic: Occupational experiences of the intimate partner violence and sexual assault workforce. *Journal of Interpersonal Violence*, *37*, NP9345–NP9366. <https://doi.org/10.1177/0886260520983304>
- Woodlock, D., McKenzie, M., Western, D., & Harris, B. (2020). Technology as a weapon in domestic violence: Responding to digital coercive control. *Australian Social Work*, *73*(3), 368–380. <https://doi.org/10.1080/0312407X.2019.1607510>
- Xue, J., Chen, J., Chen, C., Hu, R., & Zhu, T. (2020). The hidden pandemic of family violence during COVID-19: Unsupervised learning of tweets. *Journal of Medical Internet Research*, *22*(11), 1–11. <https://doi.org/10.2196/24361>
- Yokotani, K., & Takano, M. (2021). Social contagion of cyberbullying via online perpetrator and victim networks. *Computers in Human Behavior*, *119*, 106719. <https://doi.org/10.1016/j.chb.2021.106719>



Zhang, S., Liu, M., Li, Y., & Chung, J. E. (2021). Teens' social media engagement during the covid-19 pandemic: A time series examination of posting and emotion on reddit. *International Journal of Environmental Research and Public Health*, 18(19), 10079. <https://doi.org/10.3390/ijerph181910079>

## Author Biographies

**Dr. Lifang Li** is a Research Associate in Health Text Analytics and Data Science in the Institute of Psychiatry, Psychology & Neuroscience, King's College London. She currently focuses on applying natural language processing techniques to extract various types of violence using Maudsley's Clinical Record Interactive Search (CRIS) dataset and to conduct new applications. She got her PhD from the City University of Hong Kong and Xi'an Jiaotong University. Her expertise is in data science, text mining, and network analytics, with a focus on public health, crisis management, and gender inequality.

**Lilly Neubauer** is currently a PhD candidate in the cybersecurity CDT at UCL, supervised by Dr Leonie Tanczer and Dr Enrico Mariconti. Her research is looking at the digital traces of coercive control, investigating how Natural Language Processing techniques might be used to identify abusive and controlling behavior. She completed her BSc in computer science at UCL in 2021.

**Prof. Robert Stewart** is a Professor of Psychiatric Epidemiology and Clinical Informatics and leads Clinical and Population Informatics in the Maudsley NIHR Biomedical Research Centre. He has been an Academic Lead for the Maudsley's Clinical Record Interactive Search (CRIS) platform since its development in 2007–2008. CRIS is an internationally unique data resource comprising de-identified full electronic mental health records from the South London and Maudsley NHS Foundation Trust (over 400,000 patients) with extensive enhancement through multiple data linkages and natural language processing (“text mining”) algorithms. Rob Stewart joined the Institute of Psychiatry as a Junior Researcher in 1996 and has worked on the epidemiology of dementia and other late-life mental disorders and on a range of International Mental Health initiatives, as well as his more recent activity in clinical informatics. He is a practicing consultant in Liaison Old Age Psychiatry and an NIHR Senior Investigator.

**Dr. Angus Roberts** joined the Biostatistics and Health Informatics Department at King's College London as a Senior Lecturer in Health Informatics in January 2018. Angus trained and worked as a Biomedical Scientist in the NHS before switching careers to computer science, and subsequently working as an NHS Software Developer. This led to an interest in the electronic health record (EHR), medical knowledge representation, and medical terminology. He completed his PhD in Natural Language Processing (NLP) at the University of Sheffield, where he went on to lead life science work for GATE, a widely used open-source NLP framework. He currently researches many aspects of NLP applied to the EHR, and the infrastructures needed to support this.