



OPEN ACCESS

EDITED BY

Tingshao Zhu,
Chinese Academy of Sciences (CAS), China

REVIEWED BY

Xianlong Zeng,
Ohio University, United States
Anna Ceraso,
University of Brescia, Italy

*CORRESPONDENCE

Ang Li
✉ angli@bjfu.edu.cn

RECEIVED 15 March 2024

ACCEPTED 26 June 2024

PUBLISHED 08 July 2024

CITATION

Li A (2024) Predicting negative attitudes towards suicide in social media texts: prediction model development and validation study.

Front. Public Health 12:1401322.
doi: 10.3389/fpubh.2024.1401322

COPYRIGHT

© 2024 Li. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Predicting negative attitudes towards suicide in social media texts: prediction model development and validation study

Ang Li*

Department of Psychology, Beijing Forestry University, Beijing, China

Background: Implementing machine learning prediction of negative attitudes towards suicide may improve health outcomes. However, in previous studies, varied forms of negative attitudes were not adequately considered, and developed models lacked rigorous external validation. By analyzing a large-scale social media dataset (Sina Weibo), this paper aims to fully cover varied forms of negative attitudes and develop a classification model for predicting negative attitudes as a whole, and then to externally validate its performance on population and individual levels.

Methods: 938,866 Weibo posts with relevant keywords were downloaded, including 737,849 posts updated between 2009 and 2014 (*2009–2014 dataset*), and 201,017 posts updated between 2015 and 2020 (*2015–2020 dataset*). (1) For model development, based on 10,000 randomly selected posts from *2009 to 2014 dataset*, a human-based content analysis was performed to manually determine labels of each post (non-negative or negative attitudes). Then, a computer-based content analysis was conducted to automatically extract psycholinguistic features from each of the same 10,000 posts. Finally, a classification model for predicting negative attitudes was developed on selected features. (2) For model validation, on the population level, the developed model was implemented on remaining 727,849 posts from *2009 to 2014 dataset*, and was externally validated by comparing proportions of negative attitudes between predicted and human-coded results. Besides, on the individual level, similar analyses were performed on 300 randomly selected posts from *2015 to 2020 dataset*, and the developed model was externally validated by comparing labels of each post between predicted and actual results.

Results: For model development, the F1 and area under ROC curve (AUC) values reached 0.93 and 0.97. For model validation, on the population level, significant differences but very small effect sizes were observed for the whole sample ($\chi^2_1 = 32.35$, $p < 0.001$; Cramer's $V = 0.007$, $p < 0.001$), men ($\chi^2_1 = 9.48$, $p = 0.002$; Cramer's $V = 0.005$, $p = 0.002$), and women ($\chi^2_1 = 25.34$, $p < 0.001$; Cramer's $V = 0.009$, $p < 0.001$). Besides, on the individual level, the F1 and AUC values reached 0.76 and 0.74.

Conclusion: This study demonstrates the efficiency and necessity of machine learning prediction of negative attitudes as a whole, and confirms that external validation is essential before implementing prediction models into practice.

KEYWORDS

suicide, public attitudes, LIWC, machine learning, external validation

Background

Negative attitudes towards suicide have strong negative health effects on suicidal people, including reduced access to quality health care and poorer psychological well-being (1–3). The mass media contributes to the dissemination of misinformation about negative attitudes towards suicide (4–7). If the characteristics of such misinformation can be understood, it may be possible to design targeted interventions to counter its negative influence and promote public mental health awareness. Many studies employed human coders to manually analyze suicide-related misinformation in traditional mass media (e.g., newspapers). For example, Flynn and colleagues investigated the portrayal of 60 cases of homicide-suicide in newspapers from England and Wales, and found homicide-suicide was commonly described with negative stereotypes (8). Creed and colleagues collected and coded articles about Robin Williams's suicide from Canadian newspapers, and found a lack of necessary information that could help prevent suicide in 73% of articles (9). Sorensen and colleagues performed a thematic analysis of 78 suicide-related articles in Sri Lankan newspapers, and found the prevalence of biased news reporting (10). However, because of the sheer volume of mass media messages, it is difficult to efficiently screen for misinformation by human coders. Therefore, there is a dire need for automatic prediction of negative attitudes towards suicide, which may give us an efficient way to target misinformation among massive information or at least reduce the searching scope quickly for human coders.

As a representative of the new media, social media has introduced a new way to predict public attitudes towards mental health problems. Social media enables its millions of users to share their thoughts, feelings, and opinions with the public, and allows for viewing and downloading digital recordings of their publicly available online conversations. Therefore, by analyzing language use patterns of social media posts, many previous studies attempted to develop machine learning models for automatic prediction of misinformation regarding negative attitudes towards mental health problems, including depression, anxiety, and suicide (11–13). However, previous studies mainly focused on stigmatizing attitudes (negative stereotyping of suicidal people) and did not fully cover other forms of negative attitudes (without negative stereotyping of suicidal people), like dismissive, encouraging, cynical/indifferent, and disgusted attitudes associated with no stigmatizing expressions (6, 7, 14–16). Therefore, further studies are needed to incorporate varied forms of negative attitudes together and predict them as a whole.

In addition, before implementation into practice, external validation is needed to ensure that a prediction model is generalizable to new input data. External validation can be done by testing the model's output in data that is not the same as the data used to create the model. However, in many previous studies, only internal validation was performed by using cross-validation techniques (11, 17, 18). Therefore, further studies are also needed to carry out additional external validation after the development of a prediction model.

To address these concerns, by analyzing a large-scale social media dataset (Sina Weibo, a Chinese social media site that is similar to Twitter), this paper aims to consider varied forms of negative attitudes together and develop a classification model for predicting negative attitudes as a whole, and then to externally validate the developed

model on two different levels of granularity (population and individual levels). In this paper, negative attitudes refers to a set of emotions, beliefs, and behavioral intentions that may create negative outcomes for suicidal people.

Methods

This study was approved by the Institutional Review Board at the Institute of Psychology, Chinese Academy of Sciences (protocol number: H15009). Informed consent was not obtained as this study solely focused on publicly available data and involved no personally identifiable data collection or analysis.

This study consisted of four steps: (1) data collection, (2) data preprocessing, (3) model development, and (4) model validation (Figure 1).

Data collection

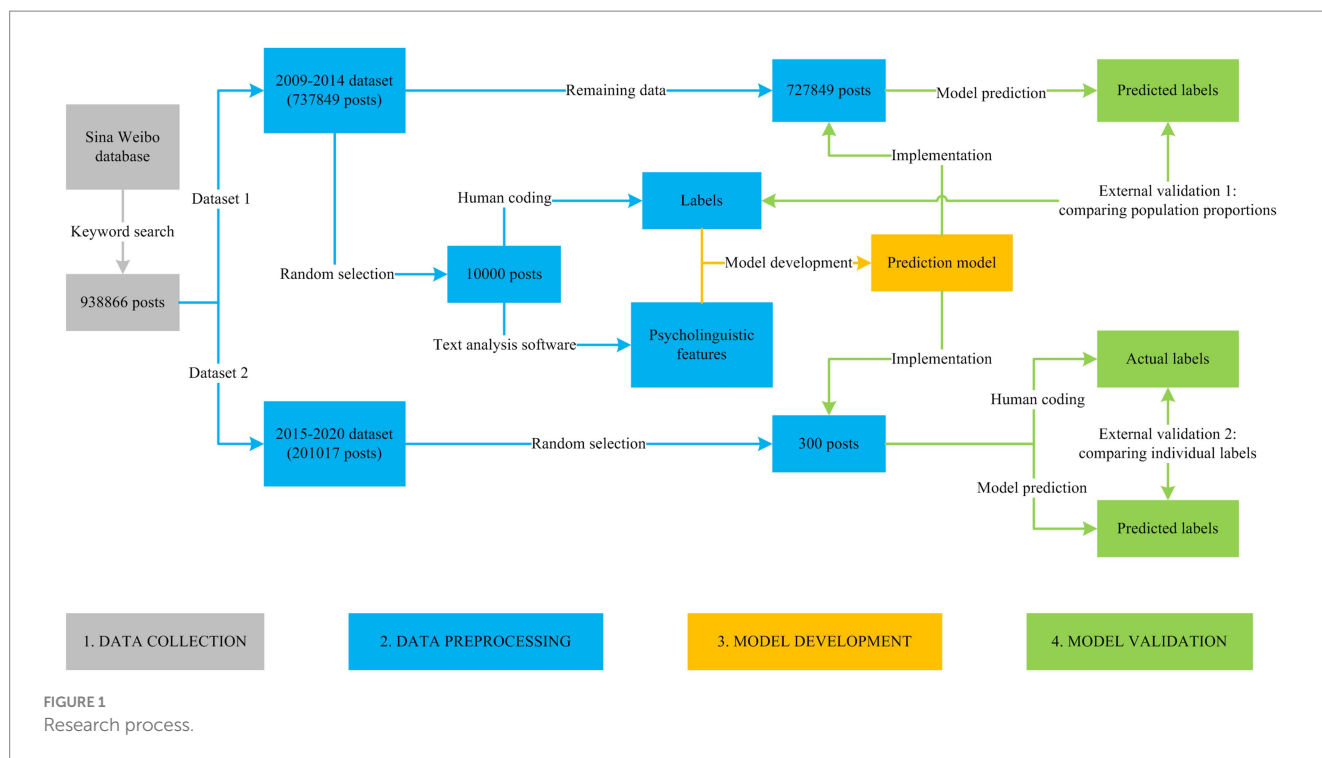
Suicide-related Sina Weibo posts were identified and downloaded. In a previous paper, a Sina Weibo database consisting of 1,953,485 active users and their Weibo posts since registration was created (19). In specific, an initial user (seed user) was randomly selected, and a breadth-first search algorithm was utilized to search for all other users incorporated in the seed user's social network. Then, the active users among them were identified based on their total number of posts and average number of posts per day, and all their Weibo posts since registration were downloaded via Application Programming Interface (API) to form the Sina Weibo database. After the creation of the database, according to a previous study (13), in order to identify as many suicide-related posts as possible, this study chose "suicide" (自杀) as the keyword to search the database and collected data in May 2020. Finally, a total of 938,866 posts with keywords were downloaded (men: 489138 posts; women: 449728 posts).

Data preprocessing

After data collection, a data preprocessing was performed on raw data to prepare it for further analysis.

First, the downloaded data was divided into two subsets: (1) *2009–2014 dataset* (737,849 posts updated between 29 August 2009 and 31 December 2014; men: 390756 posts; women: 347093 posts), and (2) *2015–2020 dataset* (201,017 posts updated between 1 January 2015 and 28 May 2020; men: 98382 posts; women: 102635 posts).

Second, in *2009–2014 dataset*, a total of 10,000 posts were randomly selected from 737,849 posts (men: 5341 posts; women: 4659 posts). In order to develop a machine learning model for predicting negative attitudes as a whole, content analyses were performed on selected 10,000 posts to obtain values of predictors and predicted outcomes, respectively. Specifically, for one thing, to obtain values of predictors, a computer-based content analysis was performed by a text analysis software (Simplified Chinese version of Linguistic Inquiry and Word Count, SCLIWIC) to extract psycholinguistic features from each post. SCLIWIC is a reliable and valid text analysis tool for automatic estimation of word frequency in



multiple psychologically meaningful categories (20). For another thing, to obtain values of predicted outcomes, a human-based content analysis was performed by two independent human coders to determine the label of each post (i.e., non-negative or negative attitudes).

Model development

After data preprocessing, by using 10,000 posts randomly selected from 2009 to 2014 dataset, a machine learning model was developed on selected psycholinguistic features to automatically predict the label of a given input data.

First, to improve prediction performance of the model, a subset of psycholinguistic features that were valid for differentiating between non-negative and negative attitudes were selected as key features for use in model construction.

In this paper, to avoid over-fitting of the model, the data used for feature selection remained independent of the data used for model training. Therefore, among 10,000 selected posts, a total of 2000 posts were randomly selected for feature selection and the remaining 8,000 posts were used for model training.

The process of feature selection was evaluated by 10-fold cross-validation, and key features were identified by three different feature selection methods. By using each feature selection method, all extracted psycholinguistic features were ranked by their importance for classification. Those ranked among the top-20 features by all three methods were considered as key features. In this paper, by using the Waikato Environment for Knowledge Analysis software (WEKA, version 3.9.6), three different feature selection methods were implemented to rank psycholinguistic features by their individual evaluations, including CorrelationAttributeEval (CAE), InfoGainAttributeEval (IGAE), and SVMAttributeEval (SVM AE).

Specifically, methods of CAE, IGAE, and SVM AE evaluate the worth of each feature by measuring the correlation between it and class, measuring the information gain with respect to the class, and using an support vector machine classifier, respectively.

Second, by using the automated machine learning package for WEKA (Auto-WEKA, <https://www.cs.ubc.ca/labs/algorithms/Projects/autoweika>), a classification model was developed based on selected key features. The Auto-WEKA is a tool for automating algorithm selection (including 30 machine learning algorithms) and hyperparameter optimization. After importing a dataset into WEKA, the Auto-WEKA can be run on it to automatically determine the best model and its parameters, and performs internal 10-fold cross-validation.

Model validation

After model development, predictive performance of the classification model was externally validated on a large-scale dataset, which was independent from the dataset used to create the model. In this paper, the external validation was carried out on two different levels of granularity (population and individual levels).

First, on the population level, in 2009–2014 dataset, the remaining 727,849 posts ($737,849 - 10,000 = 727,849$) were used for external validation 1 (men: 385,415 posts; women: 342,434 posts). Specifically, the developed classification model was implemented on the remaining 727,849 posts to automatically predict the label of each post. Because the 10,000 human-coded posts used for creating the model were randomly selected from the 2009–2014 dataset, it was assumed that there should be no significant differences in proportions of negative attitudes between predicted results (727,849 posts) and human-coded results (10,000 posts).

Second, on the individual level, in 2015–2020 dataset, a total of 300 posts were randomly selected and analyzed by two independent

human coders to determine the label of each post. Then, the developed classification model was implemented on the selected 300 posts to automatically predict the label of each post. The differences in labels of each post between predicted and actual results were examined for external validation 2.

Results

Human coding

The Cohen's *k* coefficients for attitudes reached the level of near perfect agreement (10,000 posts from 2009 to 2014 dataset: 0.92; 300 posts from 2015 to 2020 dataset: 0.95) (21). Results of human coding were shown in Table 1.

Feature selection

The top-20 ranked psycholinguistic features by three different feature selection methods were shown in Table 2, respectively. As a result, a total of six psycholinguistic features were selected as key features, including *adverbs*, *positive emotion*, *assent*, *health*, *swear*, and *2nd pers singular*.

Model development

A classification model using the C4.5 algorithm was developed on six key features. The classification performance of the C4.5 model was evaluated using precision, recall, F1, and area under ROC curve (AUC). The weighted-average F1 and AUC values reached 0.93 and 0.97, respectively (Table 3). Since model training outcomes did not indicate relationships between psycholinguistic features and predicted labels, a logistic regression analysis was performed to explore such relationships (Table 4).

Model validation

First, on the population level, significant differences but very small effect sizes were observed between predicted results (727,849 posts) and human-coded results (10,000 posts) for the whole sample

($\chi^2_1 = 32.35, p < 0.001$; Cramer's $V = 0.007, p < 0.001$), men ($\chi^2_1 = 9.48, p = 0.002$; Cramer's $V = 0.005, p = 0.002$) and women ($\chi^2_1 = 25.34, p < 0.001$; Cramer's $V = 0.009, p < 0.001$) (Table 5).

Second, on the individual level, between predicted and actual results (300 posts), the weighted-average F1 and AUC values reached 0.76 and 0.74, respectively (Table 6).

Discussion

Principal results

This paper incorporated varied forms of negative attitudes together and developed a classification model for predicting negative attitudes in social media texts, and then externally validated its predictive performance on two different levels of granularity (population and individual levels), using a large-scale dataset. Results of this paper demonstrate the efficiency and necessity of machine learning prediction of negative attitudes as a whole, and confirm that external validation is essential before implementing prediction models into practice.

First, it is necessary to consider varied forms of negative attitudes together to better understand the nature of public reactions to suicide. In this paper, human-coded results showed that, among suicide-related posts, the prevalence of negative attitudes towards suicide was approximately around 40%, which is obviously higher than the findings of previous studies that focused on limited forms of negative attitudes (23% ~ 35%) (15, 16). It means that, on Chinese social media, negative attitudes towards suicide may be actually much more prevalent than previously estimated. According to the Healthy China Action Plan (2019–2030), by 2030, health literacy of the entire Chinese population is expected to be greatly improved (22). Therefore, to promote public mental health awareness as planned, mental health professionals may be under greater pressure than ever before, and there is a dire need for innovative approaches to efficiently promote public reactions to suicide.

Second, the method of psycholinguistic analysis provides insights into the language use patterns of negative expressions towards suicide, and facilitates automatic prediction of negative attitudes in social media texts. Results of this paper showed that expressions of negative attitudes towards suicide were associated with increased use of words related to adverbs (e.g., *very*), positive emotion (e.g., *nice*), assent (e.g., *ok*), swear (e.g., *fuck*), and 2nd pers. singular (e.g., *you*); and decreased use of words related to health

TABLE 1 Results of human coding.

	2009–2014 dataset	%	2015–2020 dataset	%
All	10,000		300	
Non-negative attitudes	5,966	59.66%	176	58.67%
Negative attitudes	4,034	40.34%	124	41.33%
Men	5,341		143	
Non-negative attitudes	3,216	60.21%	93	65.03%
Negative attitudes	2,125	39.79%	50	34.97%
Women	4,659		157	
Non-negative attitudes	2,750	59.03%	83	52.87%
Negative attitudes	1,909	40.97%	74	47.13%

TABLE 2 The top 20 ranked psycholinguistic features.

	CorrelationAttributeEval	InfoGainAttributeEval	SVMAttributeEval
1	<i>Adverbs</i>	<i>Adverbs</i>	<i>Swear</i>
2	<i>Positive emotion</i>	Total function words	<i>Positive emotion</i>
3	Total pronouns	Personal pronouns	<i>Health</i>
4	Total function words	<i>Health</i>	Signal words for past tense
5	Personal pronouns	Total pronouns	<i>Adverbs</i>
6	<i>Assent</i>	Word count	Motion
7	<i>Health</i>	<i>Assent</i>	Sad
8	Semicolons	<i>Positive emotion</i>	<i>2nd pers singular</i>
9	1st pers. singular	1st pers. singular	Sexual
10	Social processes	Semicolons	Common verbs
11	<i>Swear</i>	Signal words for continuous tense	Anger
12	Word count	Exclamation marks	Time
13	Sad	Social processes	Dashes
14	Signal words for tenses	<i>Swear</i>	Apostrophes
15	Signal words for continuous tense	Signal words for tenses	<i>Assent</i>
16	Motion	<i>2nd pers singular</i>	3rd pers. plural
17	<i>2nd pers singular</i>	Biological processes	Negations
18	Affective processes	3rd pers. singular	Relativity
19	Biological processes	Affective processes	Question marks
20	Exclusive	Humans	Religion

Bold and italics indicate key features.

TABLE 3 Performance of the developed C4.5 model.

	Precision	Recall	F1	Area under ROC curve
Class 1: non-negative attitudes	0.92	0.96	0.94	0.97
Class 2: negative attitudes	0.94	0.88	0.91	0.97
Weighted-average	0.93	0.93	0.93	0.97

TABLE 4 Relationships between psycholinguistic features and predicted labels.

Predicted labels	Psycholinguistic features	β
Non-negative attitudes (1) vs. Negative attitudes (2)	Adverbs	5.75
	Positive emotion	12.48
	Assent	2.98
	Health	-12.56
	Swear	21.07
	2nd pers. singular	12.27

(e.g., clinic). According to previous studies, increased use of swear-related words reflects enhanced emotional-related impulsivity and frustration, and therefore are commonly considered as an indicator of negative affect (23–25). Besides, more frequent use of health-related words indicates people are more concerned about health problems (26). Therefore, in this paper, decreased use of health-related words may imply a view that suicide is not worth paying attention to, and was related to negative attitudes. Moreover, it is worthy to note here that increased use of words related to positive

emotion and assent are positive predictors for emotional support in previous studies (27, 28). However, in this paper, increased use of words indicating positive emotion and assent were related to negative attitudes towards suicide, which may be due to the prevalence of normalization or glorification of suicide (29, 30). With the help of psycholinguistic analysis techniques, compared with results of other similar studies (F1 values: 0.62 ~ 0.86) (11, 13, 17), the developed machine learning model performed well in classifying non-negative and negative expressions (F1 value: 0.76 ~ 0.93). It means that understanding the differences in language use patterns between non-negative and negative expressions would make the features used in machine-learning models more understandable for end user and make machine learning prediction more precise. Furthermore, the use of machine learning approaches for automatic prediction of negative attitudes provide us an efficient way to reduce the workload of human coders for searching misinformation. For example, in this paper, the developed classification model was implemented on 727,849 unlabeled posts, and then automatically labeled 273,425 of them as negative attitudes. It means that 62.43% of the workload of human coders can be reduced quickly (62.43% = (727849–273,425)/727,849).

TABLE 5 Comparison of predicted and human-coded results (2009–2014 dataset).

	Predicted	%	Human-coded	%
All	727,849		10,000	
Non-negative attitudes	454,424	62.43%	5,966	59.66%
Negative attitudes	273,425	37.57%	4,034	40.34%
Men	385,415		5,341	
Non-negative attitudes	239,996	62.27%	3,216	60.21%
Negative attitudes	145,419	37.73%	2,125	39.79%
Women	342,434		4,659	
Non-negative attitudes	214,428	62.62%	2,750	59.03%
Negative attitudes	128,006	37.38%	1,909	40.97%

TABLE 6 Comparison of predicted and actual results (2015–2020 dataset).

	Precision	Recall	F1	Area under ROC curve
Class 1: non-negative attitudes	0.79	0.80	0.80	0.74
Class 2: negative attitudes	0.71	0.70	0.71	0.74
Weighted-average	0.76	0.76	0.76	0.74

Third, external validation of machine learning prediction models is crucial before clinical application. Without the process of external validity, it is unclear about the extent to which the prediction performance of the developed model can be generalized to other situations. In this paper, by using an internal cross-validation technique, a machine learning model was developed, and the weighted-average F1 and AUC values reached a nearly perfect level of prediction performance (internal validity: 0.93 and 0.97). However, when the developed model was implemented to the new data that is not the same as the data used to develop the model, the weighted-average F1 and AUC values dropped to a satisfactory level of prediction performance (external validity: 0.76 and 0.74). It means that, although similar studies have already performed an internal validity of developed mental health machine learning models and have achieved good predictive performance, additional external validity tests are still needed to ensure that a prediction model is generalizable to new input data other than those in the derivative cohort. Results of the external validity may show us the actual predictive performance of developed models under realistic conditions, and provide implications for making large-scale automatic and non-intrusive forecasting of public attitudes towards suicide. Findings suggest that, under realistic conditions, population-based prediction (predicting proportions of negative attitudes) may be considerably more effective than prediction that targets individuals (predicting labels of each post).

Limitations

Limitations existed in this study. First, because the number of posts is extremely unbalanced and limited among varied forms of negative attitudes, machine learning models for predicting specific forms of negative attitudes cannot be developed in this paper. Second, this paper mainly focused on Sina Weibo users, who are not representative of Chinese populations or all Chinese social media users. Therefore, the

findings may not be applicable to other non-users. Third, this paper focused on a single social media platform (Sina Weibo) and only analyzed Chinese social media posts, so the findings of this paper may have limited generalizability. Fourth, since the Auto-WEKA package can only perform automatic algorithm selection on 30 classical machine learning algorithms, it is not known if using the latest machine learning algorithms (e.g., deep learning models and large language models) will improve the predictive performance of the classification model.

Conclusion

This paper demonstrates the efficiency and necessity of machine learning prediction of negative attitudes as a whole, and confirms that external validation is essential before implementing prediction models into practice. Findings suggest that, under realistic conditions, population-based prediction may be considerably more effective than prediction that targets individuals.

Data availability statement

The datasets presented in this article are not readily available because the raw data cannot be made public (if necessary, feature data can be provided). Requests to access the datasets should be directed to AL, angli@bjfu.edu.cn.

Ethics statement

The studies involving humans were approved by Institutional Review Board at the Institute of Psychology, Chinese Academy of Sciences. The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation was not required from the participants or the

participants' legal guardians/next of kin because this study solely focused on publicly available data and involved no personally identifiable data collection or analysis.

Author contributions

AL: Conceptualization, Investigation, Formal analysis, Methodology, Writing – original draft.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

References

- Saunders KEA, Hawton K, Fortune S, Farrell S. Attitudes and knowledge of clinical staff regarding people who self-harm: a systematic review. *J Affect Disord.* (2012) 139:205–16. doi: 10.1016/j.jad.2011.08.024
- Asare-Doku W, Osafo J, Akotia CS. The experiences of attempt survivor families and how they cope after a suicide attempt in Ghana: a qualitative study. *BMC Psychiatry.* (2021) 17:178. doi: 10.1186/s12888-017-1336-9
- Gselamu L, Ha K. Attitudes towards suicide and risk factors for suicide attempts among university students in South Korea. *J Affect Disord.* (2020) 272:166–9. doi: 10.1016/j.jad.2020.03.135
- Whitley R, Saucier A. Media coverage of Canadian veterans, with a focus on post traumatic stress disorder and suicide. *BMC Psychiatry.* (2022) 22:339. doi: 10.1186/s12888-022-03954-8
- Marzano L, Hawley M, Fraser L, Lainez Y, Marsh J, Hawton K. Media coverage and speculation about the impact of the COVID-19 pandemic on suicide: a content analysis of UK news. *BMJ Open.* (2023) 13:e065456. doi: 10.1136/bmjopen-2022-065456
- Westerlund M, Hadlaczky G, Wasserman D. Case study of posts before and after a suicide on a Swedish internet forum. *Br J Psychiatry.* (2015) 207:476–82. doi: 10.1192/bjp.bp.114.154484
- Ma J, Zhang W, Harris K, Chen Q, Xu X. Dying online: live broadcasts of Chinese emerging adult suicides and crisis response behaviors. *BMC Public Health.* (2016) 16:774. doi: 10.1186/s12889-016-3415-0
- Flynn S, Gask L, Shaw J. Newspaper reporting of homicide-suicide and mental illness. *BJPsy Bull.* (2015) 39:268–72. doi: 10.1192/pb.bp.114.049676
- Creed M, Whitley R. Assessing fidelity to suicide reporting guidelines in Canadian news media: the death of Robin Williams. *Can J Psychiatr.* (2017) 62:313–7. doi: 10.1177/0706743715621255
- Sorensen JB, Pearson M, Armstrong G, Andersen MW, Weerasinghe M, Hawton K, et al. A qualitative analysis of self-harm and suicide in Sri Lankan printed newspapers. *Crisis.* (2021) 42:56–63. doi: 10.1027/0227-5910/a000687
- Li A, Jiao D, Zhu T. Detecting depression stigma on social media: a linguistic analysis. *J Affect Disord.* (2018) 232:358–62. doi: 10.1016/j.jad.2018.02.087
- Zhu J, Li Z, Zhang X, Zhang Z, Hu B. Public attitudes toward anxiety disorder on Sina Weibo: content analysis. *J Med Internet Res.* (2023) 25:e45777. doi: 10.2196/45777
- Li A, Jiao D, Zhu T. Stigmatizing attitudes across cybersuicides and offline suicides: content analysis of Sina Weibo. *J Med Internet Res.* (2022) 24:e36489. doi: 10.2196/36489
- O'Dea B, Achilles MR, Larsen ME, Batterham PJ, Calear AL, Christensen H. The rate of reply and nature of responses to suicide-related posts on twitter. *Internet Interv.* (2018) 13:105–7. doi: 10.1016/j.invent.2018.07.004
- Fu KW, Cheng Q, Wong PWC, Yip PSF. Responses to a self-presented suicide attempt in social media: a social network analysis. *Crisis.* (2013) 34:406–12. doi: 10.1027/0227-5910/a000221

Conflict of interest

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

- Li A, Huang X, Hao B, O'Dea B, Christensen H, Zhu T. Attitudes towards suicide attempts broadcast on social media: an exploratory study of Chinese microblogs. *PeerJ.* (2015) 3:e1209. doi: 10.7717/peerj.1209
- O'Dea B, Wan S, Batterham PJ, Calear AL, Paris C, Christensen H. Detecting suicidality on twitter. *Internet Interv.* (2015) 2:183–8. doi: 10.1016/j.invent.2015.03.005
- Guan L, Hao B, Cheng Q, Yip PSF, Zhu T. Identifying Chinese microblog users with high suicide probability using internet-based profile and linguistic features: classification model. *JMIR Ment Health.* (2015) 2:e17. doi: 10.2196/mental.4227
- Li L, Li A, Hao B, Guan Z, Zhu T. Predicting active users' personality based on micro-blogging behaviors. *PLoS One.* (2014) 9:e84997. doi: 10.1371/journal.pone.0084997
- Zhao N, Jiao D, Bai S, Zhu T. Evaluating the validity of simplified Chinese version of LIWC in detecting psychological expressions in short texts on social network services. *PLoS One.* (2016) 11:e0157947. doi: 10.1371/journal.pone.0157947
- Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics.* (1977) 33:159–74. doi: 10.2307/2529310
- Zhang Z, Jiang W. Health education in the healthy China initiative 2019–2030. *China CDC Wkly.* (2021) 3:78–80. doi: 10.46234/ccdcw2021.018
- Vornfield R, Toma CL, Shah DV, Moon TJ, Gustafson DH. What do you say before you relapse? How language use in a peer-to-peer online discussion forum predicts risky drinking among those in recovery. *Health Commun.* (2018) 33:1184–93. doi: 10.1080/10410236.2017.1350906
- Stamatis CA, Meyerhoff J, Liu T, Sherman G, Wang H, Liu T, et al. Prospective associations of text-message-based sentiment with symptoms of depression, generalized anxiety, and social anxiety. *Depress Anxiety.* (2022) 39:794–804. doi: 10.1002/da.23286
- Park A, Conway M. Longitudinal changes in psychological states in online health community members: understanding the long-term effects of participating in an online depression community. *J Med Internet Res.* (2017) 19:e71. doi: 10.2196/jmir.6826
- Andy A, Andy U. Understanding communication in an online cancer forum: content analysis study. *JMIR Cancer.* (2021) 7:e29555. doi: 10.2196/29555
- Verberne S, Batenburg A, Sanders R, van Eenbergen M, Das E, Lambooi MS. Analyzing empowerment processes among cancer patients in an online community: a text mining approach. *JMIR Cancer.* (2019) 5:e9887. doi: 10.2196/cancer.9887
- Stevens KA, Ronan K, Davies G. Treating conduct disorder: an effectiveness and natural language analysis study of a new family-centred intervention program. *Psychiatry Res.* (2017) 251:287–93. doi: 10.1016/j.psychres.2016.11.035
- Batterham PJ, Han J, Calear AL, Anderson J, Christensen H. Suicide stigma and suicide literacy in a clinical sample. *Suicide Life Threat Behav.* (2019) 49:1136–47. doi: 10.1111/sltb.12496
- Ludwig J, Dreier M, Liebherz S, Härter M, von dem Knesebeck O. Suicide literacy and suicide stigma – results of a population survey from Germany. *J Ment Health.* (2022) 31:517–23. doi: 10.1080/09638237.2021.1875421