

---

## Research and Applications

# An NLP approach to identify SDoH-related circumstance and suicide crisis from death investigation narratives

Song Wang  <sup>1</sup>, Yifang Dang  <sup>2</sup>, Zhaoyi Sun <sup>3</sup>, Ying Ding <sup>4</sup>, Jyotishman Pathak <sup>3</sup>,  
Cui Tao <sup>2</sup>, Yunyu Xiao <sup>3</sup>, and Yifan Peng  <sup>3</sup>

<sup>1</sup>Cockrell School of Engineering, The University of Texas at Austin, Austin, Texas, USA, <sup>2</sup>School of Biomedical Informatics, The University of Texas Health Science Center at Houston, Houston, Texas, USA, <sup>3</sup>Population Health Sciences, Weill Cornell Medicine, New York, New York, USA, and <sup>4</sup>School of Information, The University of Texas at Austin, Austin, Texas, USA

Corresponding Author: Yifan Peng, PhD, Population Health Sciences, Weill Cornell Medicine, 425 E 61ST ST DIV 305, New York, NY 10065, USA; yip4002@med.cornell.edu

Yunyu Xiao and Yifan Peng are co-last authors.

Received 5 December 2022; Revised 10 March 2023; Editorial Decision 28 March 2023; Accepted 29 March 2023

## ABSTRACT

**Objectives:** Suicide presents a major public health challenge worldwide, affecting people across the lifespan. While previous studies revealed strong associations between Social Determinants of Health (SDoH) and suicide deaths, existing evidence is limited by the reliance on structured data. To resolve this, we aim to adapt a suicide-specific SDoH ontology (Suicide-SDoHO) and use natural language processing (NLP) to effectively identify individual-level SDoH-related social risks from death investigation narratives.

**Materials and Methods:** We used the latest National Violent Death Report System (NVDRS), which contains 267 804 victim suicide data from 2003 to 2019. After adapting the Suicide-SDoHO, we developed a transformer-based model to identify SDoH-related circumstances and crises in death investigation narratives. We applied our model retrospectively to annotate narratives whose crisis variables were not coded in NVDRS. The crisis rates were calculated as the percentage of the group's total suicide population with the crisis present.

**Results:** The Suicide-SDoHO contains 57 fine-grained circumstances in a hierarchical structure. Our classifier achieves AUCs of 0.966 and 0.942 for classifying circumstances and crises, respectively. Through the crisis trend analysis, we observed that not everyone is equally affected by SDoH-related social risks. For the economic stability crisis, our result showed a significant increase in crisis rate in 2007–2009, parallel with the Great Recession.

**Conclusions:** This is the first study curating a Suicide-SDoHO using death investigation narratives. We showcased that our model can effectively classify SDoH-related social risks through NLP approaches. We hope our study will facilitate the understanding of suicide crises and inform effective prevention strategies.

**Key words:** suicide, social determinants of health, natural language processing, social risks

---

## INTRODUCTION

Suicide, as a fatal component of suicidal behavior, results from the complex interactions of individual, interpersonal, social, and environmental influences.<sup>1–4</sup> In the United States, the number of suicide deaths has increased by more than 30% in the past 2 decades,

with 45 797 Americans dying by suicide and another 1.2 million attempting suicide in 2020.<sup>5,6</sup> Though suicide rates vary by race and ethnicity, certain populations are disproportionately impacted due to systematic racism, structural disadvantages, and cultural contexts.<sup>7–11</sup>

Social Determinants of Health (SDoH) contain conditions in which people live that affect a wide range of health, quality-of-life outcomes, and risks,<sup>12</sup> such as social and community context and economic stability. While SDoH strongly influence health equity, little is known about how they impact disparities in suicides.<sup>2,10</sup> Understanding these interactions is necessary to identify the underlying mechanisms that can inform the development of effective suicide prevention strategies.<sup>13,14</sup>

The National Violent Death Reporting System (NVDRS) is a state-based violent death reporting system in the United States that provides information and contexts on when, where, and how violent deaths occur and who is affected.<sup>15</sup> It gathers and links detailed investigative information from several sources, including death investigation reports, toxicology, and death certificates. The detailed investigative information gathered in NVDRS can provide an overall picture of the circumstances contributing to violent deaths. Specifically, each incident in NVDRS is accompanied by 2 death investigation narratives by a coroner or medical examiner (CME) and law enforcement (LE) reporter describing the circumstances (eg, social contexts, interpersonal relationships, life events, mental illness, etc.). Such circumstances are judged as the potential causes of suicide deaths. Although NVDRS has a standard coding manual that defines these circumstances, manual data labeling is labor- and cost-intensive. Furthermore, circumstance variables have been added and coded to NVDRS at various times. For example, more than 20% of the circumstance variables were added in 2013, 10 years after NVDRS was created. Therefore, only data from the year the variable was officially added to NVDRS and onward could be included for analysis. These limitations raise substantial barriers to fully utilizing the rich data in NVDRS to improve suicide prevention.

In addition to SDoH-related circumstances, the crisis variables in NVDRS available after 2013 are important as they identify suicide that appears to involve an element of impulsivity.<sup>16</sup> Formally, a “Crisis” is an acute event (within 2 weeks before the suicide) that is indicated in one of the CME or LE reports to have contributed to the suicide. Figure 1 shows an example of SDoH-related circumstances and suicide crises. The LE report mentions “V’s medical history

of mental illness, alcoholism, and pain medication abuse,” which respectively points to the presence of mental health, alcohol, and substance abuse problems. Also, “V and his wife were going through a separation, and a divorce” reveals the presence of an intimate partner problem. As for suicide crises, an intimate partner crisis is present because V “shot the boyfriend” of his wife when the incident occurred. An alcohol crisis is identified because “V’s toxicology was positive for alcohol.” Figure 1 also shows that the actual timing of the crisis may not be mentioned in the text, so we need a deeper understanding of natural language as a clue to differentiate suicide crises and noncrisis circumstances.

Understanding the impact of SDoH on suicide risks is important to design suicide interventions. However, this is limited because SDoH information in NVDRS is embedded in the unstructured narratives reported by CME and LE. Therefore, to directly extract suicide-specific SDoH from free-text death investigation narratives, natural language processing (NLP) solutions are well-suited for their proven capability of representing and analyzing free-text human language.<sup>17</sup>

Most prior NLP-informed efforts of extracting standard SDoH information are conducted using clinical texts<sup>18</sup> and mainly rely on 3 methods: (1) rule-based methods such as keyword matching,<sup>19-21</sup> regular expressions,<sup>22</sup> or similarity matching<sup>23-25</sup> to identify SDoH from clinical documents; (2) supervised methods such as traditional feature-based machine learning algorithms (eg, support vector machine and random forests)<sup>26</sup> and transformer-based deep learning models<sup>27</sup>; and (3) unsupervised approaches including topic modeling<sup>28,29</sup> and latent class analysis.<sup>30</sup> Despite their contributions to SDoH extractions using NLP, they were based on data sources that are primarily clinical notes and thus cannot generalize well to death investigation narratives. There is also a lack of understanding regarding the temporal sequences of causes of deaths, especially regarding the crises signaling the impulsive warning signs of suicide.

To bridge these gaps, we propose to adopt an NLP approach to automatically detect SDoH-related circumstances and crises before suicide deaths from NVDRS death investigation narratives. Specifically, enabled by transformer-based models, we developed a deep

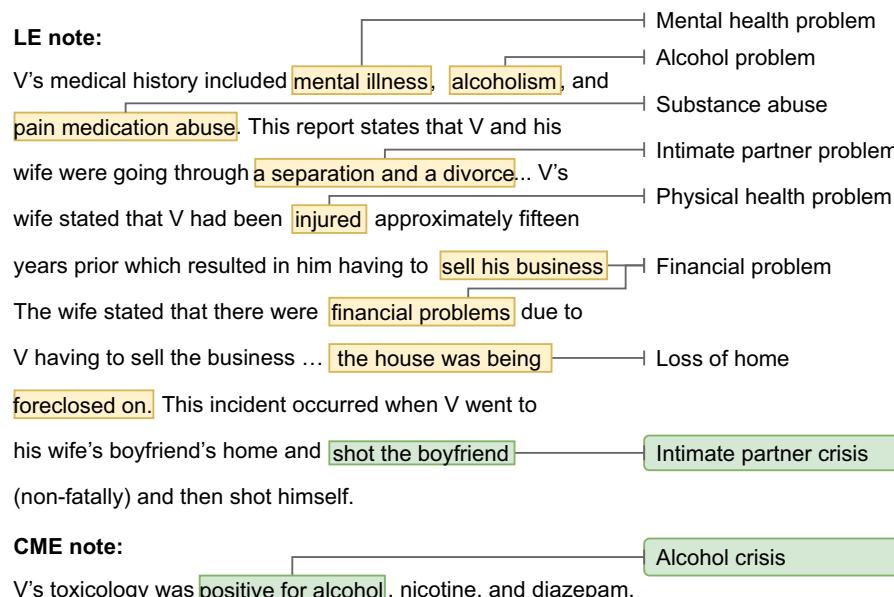


Figure 1. Example of the SDoH-related circumstances and suicide crises.

learning model that takes death investigation narratives as input and learns to classify SDoH-related circumstances and crisis. Through experiments, we showed the advantages of using data-driven approaches to effectively extract SDoH from free-text death investigation narratives to improve designing suicide prevention programs. We further conducted an in-depth crisis trend analysis by sex and age groups.

## MATERIALS AND METHODS

### Task definition

This study aims to identify 2 types of individual-level, SDoH-related social risks related to suicide. First is the “Circumstances,” defined as the precipitating events contributing to the infliction of each fatal injury incident (eg, “mental illness” as a mental health problem, “alcoholism” as an alcohol problem in Figure 1).<sup>15</sup> The second one is “Crisis,” defined as the circumstances that occurred within 2 weeks before the suicide death (eg, “toxicology was positive for alcohol” as alcohol crisis in Figure 1).

We split the task of suicide-related SDoH classification into 2 tasks: SDoH-related circumstance classification (ie, classifying if the victim had one or more SDoH-related circumstances present) and crisis classification (ie, classifying if the victim had one or more suicide crises present). Specifically, we approach the tasks in a text classification manner instead of through named entity recognition. This is because our focus is to classify whether the victim-level

SDoH-related circumstances or suicide crises are present (text classification) instead of identifying the exact text fragments where they are mentioned (named entity recognition). Moreover, extracting SDoH information often relies on comprehending relatively long textual contexts, which makes it difficult to extract accurate SDoH information (especially suicide crises) by recognizing the exact named entity mentioned.

### Data description

The latest NVDRS dataset contains 267 804 suicide death incidents in 50 US states, Puerto Rico, and the District of Columbia from 2003 to 2019.<sup>15</sup> This study is approved by the NVDRS Restricted Access Dataset (RAD) proposal.

Each incident is accompanied by 2 suicide death narratives: one that summarizes the sequence of events from the perspective of the CME record and one that summarizes the sequence of events of the incident from the perspective of the LE report. NVDRS coded over 600 unique data elements for each incident that provide valuable SDoH information about violent deaths, such as SDoH-related circumstances and suicide crises. SDoH-related circumstances and suicide crises are reported based on the content of CME and LE reports. In this study, we focus on 16 individual social risks extracted from the NVDRS codebook in 2019, including SDoH-related circumstances (Table 1) and their corresponding suicide crises.

**Table 1.** Circumstances and descriptions

First-level superclass	Second-level superclass	Circumstance	Description
Social community context	Mental health	Mental health problem	Current mental health problem.
		Family relationship problem	Victim had relationship problems with a family member (other than an intimate partner) that appear to have contributed to the death.
		Other relationship problem	Problems with a friend or associate (other than an intimate partner or family member) appear to have contributed to the death.
	Safety concern	Intimate partner problem	Problems with a current or former intimate partner appear to have contributed to the suicide or undetermined death.
		Criminal legal problem	Criminal legal problem(s) appear to have contributed to the death.
		Civil legal problem	Civil legal (noncriminal) problem(s) appear to have contributed to the death.
		Recent suicide of friend or family	Suicide of a family member or friend appears to have contributed to the death.
		Disaster exposure	Exposure to a disaster was perceived as a contributing factor in the incident.
		Physical health problem	Victim's physical health problem(s) appear to have contributed to the death.
		Job problem	Job problem(s) appear to have contributed to the death.
Behavior and lifestyle	Substance abuse	School problem	Problems at or related to school appear to have contributed to the death.
		Alcohol problem	Person has alcohol dependence or alcohol problems.
		Substance abuse	Person has a nonalcohol-related substance abuse problem.
Economic stability	Financial distressing	Other addiction	Person has an addiction other than alcohol or other substance abuse, such as gambling, sexual, etc., that appears to have contributed to the death.
		Financial problem	Financial problems appear to have contributed to the death.
		Eviction or loss of home	A recent eviction or other loss of the victim's housing, or the threat of it, appears to have contributed to the death.

*Note:* The hierarchical class structure comes from the suicide-specific SDoHO. The circumstance variables and descriptions are from the NVDRS codebook.<sup>16</sup> SDoHO: Social Determinants of Health Ontology.

Incidents with no present SDoH-related circumstances were first left out, resulting in a total of 236 606 incident entries. Given that 10 out of 16 SDoH-related circumstances were not added to NVDRS until August 2013, we selected incidents after August 2013 (133 254 instances) for training and testing, while incidents before August 2013 (103 352 instances) were used for inferencing and crisis trend analysis only. Training and testing data were split with a ratio of 8:2 at the victim level to avoid data leakage, resulting in 106 603 instances in the training set and 26 651 instances in the test set. The distributions of SDoH-related circumstances and corresponding suicide crises can be found in [Figure 2](#), the detailed statistics can be found in [Supplementary Table SA1](#). The average number of SDoH-related circumstances per victim is 2.04, and the average number of suicide crises per victim is 0.66.

### Suicide-specific SDoH ontology

SDoH ontology (SDoHO) is a well-defined ontology with a hierarchical class structure and properties that comprehensively represent SDoH collected from various sources.<sup>31</sup> There are 9 main categories covered by SDoHO, including aspects of behavior and lifestyle, demographics, education, social and community context, health care, economic stability, neighborhood, food, and measurements. YX first verified a set of circumstance variables and suicide crisis variables that are related to SDoH under 5 NVDRS categories ("Crime and Criminal Activity," "Manner Specific Circumstances for Homicide," "Manner Specific Circumstances for Suicide," "Mental Health, Substance Abuse, and Other Addictions," and "Relationships, Abuse, and Life Stressors") based on their definitions in NVDRS. We (YD and CT) further manually mapped each variable to its best-matched SDoHO category based on its definition in NVDRS. A subset of 57 of the SDoH-related circumstance variables and suicide crisis variables can find a matching SDoHO category, which we call Suicide-SDoHO. [Figure 3](#) shows the hierarchical suicide-specific SDoHO (Suicide-SDoHO), where the 16 focus individual social risk variables of this work (as listed in [Table 1](#)) are highlighted in green.

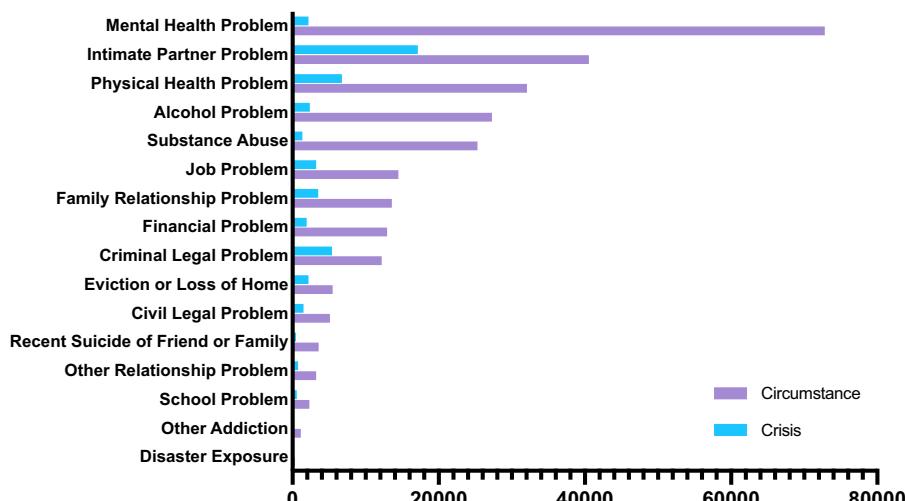
### Transformer-based method

In this work, we focus on classifying 16 SDoH-related circumstances and their corresponding suicide crises. The adapted Suicide-SDoHO

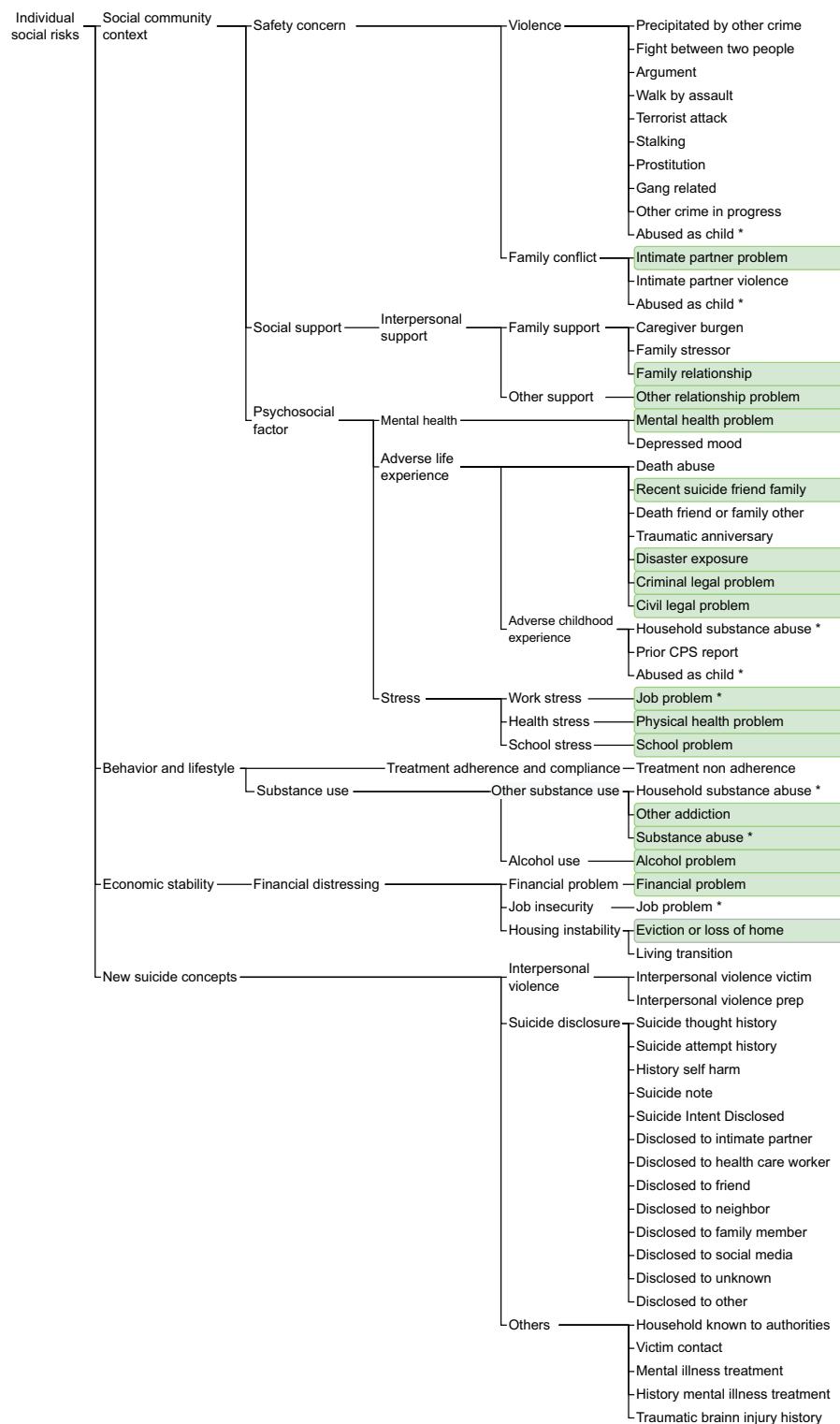
categorizes the 16 circumstance variables into 7 second-level superclasses (ie, safety concern, interpersonal support, mental health, adverse life experience, stress, substance use, financial distressing), and further into 3 first-level superclasses (ie, social community context, behavior and lifestyle, economic stability), which naturally partitions SDoH-related circumstances into 16 classes, 7 classes, and 3 classes in different levels. We treated the SDoH-related circumstance (ie, the events contributing to the infliction of each fatal injury incident) classification and suicide crisis (ie, the circumstances that occurred within 2 weeks before the suicide death) classification as 2 separate multilabel text classification problems. Our model uses Bidirectional Encoder Representations from Transformers (BERT)<sup>24</sup> as a backbone, leverages the state-of-the-art transformer-based language model (LM), and further fine-tunes the pretrained LM in the classification downstream tasks. After tokenizing the input texts, BERT takes a sequence of tokens with a maximum length of 512 and produces a 768-dimensional sequence representation vector. A fully connected layer is appended on top of the BERT model to map the representation vector to the target label space for classification. We trained 2 different BERT models on different sets of annotations, respectively the circumstance annotations and the crisis annotations. We expect the BERT models to learn to differentiate crisis versus circumstance by learning to identify the underlying temporal relations.

### Crisis trend analysis by sex and age groups

To make a fair comparison between population groups of various sizes, we computed the group-wise crisis rates for crisis trend analysis. Specifically, the crisis rate is defined as the percentage of the group's total suicide population that has the crisis present. To compute the average crisis rates, we used the ground truth labels for years after 2013 and used the model's predictions for years prior to 2013 whose ground truth labels were not available. In this study, we calculated the 3-year rolling average suicide crisis rates of 3 coarse-grained SDoH-related suicide crises: social community context crisis, behavior and lifestyle crisis, and economic stability crisis. We further conducted the crisis trend analysis after regrouping victims by sex and age.



**Figure 2.** Distributions of SDoH-related circumstances and suicide crises.



**Figure 3.** Suicide-specific SDoHO. The focus circumstance/crisis variables in this study are highlighted in highlight boxes. \*Circumstances belonging to multiple superclasses. SDoHO: Social Determinants of Health Ontology.

## Experiment settings

We selected the BioBERT<sup>32</sup> as our backbone model, which takes the concatenation of each victim's CME and LE reports (whichever is shorter in length will be put first) as input. For each task, 3 multilabel classifiers have been developed for the classification tasks whose

label space has 16 classes, 7 classes, and 3 classes, respectively. We used the Adam optimizer<sup>33</sup> and binary Cross-Entropy loss for parameter optimization. We used a learning rate of 10E-6, batch size of 12, and 20 epochs of training with Early-Stopping (patience = 5) to prevent overfitting. Intel Core i9-9960X 16 cores

processor, NVIDIA Quadro RTX 5000 GPU, and a memory size of 128G were used in this work. We evaluated the model performance using 4 commonly used classification metrics, including Precision, Recall, F-1, and ROC-AUC scores. We have made our code publicly available at [https://github.com/bionlplab/suicide\\_sdoh](https://github.com/bionlplab/suicide_sdoh).

## RESULTS

### SDoH-related circumstance classification

**Table 2** shows the performances of the 16-class, 7-class, and 3-class multilabel SDoH-related circumstance classifiers, respectively. For the 16-class multilabel classification, we achieved a weighted AUC of 0.966 and a weighted F-1 score of 0.837. Among all circumstances, the classification of mental health problem has the best F-1 score of 0.920. For the 7-class multilabel classification, we achieved a weighted AUC of 0.962 and a weighted F-1 of 0.863. The classification of mental health problem has the highest F-1 score of 0.922. For the 3-class multilabel classification, we achieved a weighted AUC of 0.948 and a weighted F-1 score of 0.944. The classification of social community context has the highest F-1 score of 0.979.

### Suicide crisis classification

**Table 2** also shows the 16-class, 7-class, and 3-class multilabel suicide crisis classification results. For the 16-class multilabel

classification, the weighted AUC is 0.942 and the weighted F-1 score is 0.563, with the classification of the intimate partner crisis having the highest F-1 score of 0.687. For the 7-class multilabel classification, the weighted AUC score is 0.938 and the weighted F-1 score is 0.578, with the classification of the safety concern crisis having the highest F-1 score of 0.696. For the 3-class multilabel classification, the weighted AUC score is 0.885 and the weighted F-1 score is 0.634, with the social community context crisis having the highest F-1 score of 0.671.

### Crisis trend analysis by sex and age groups

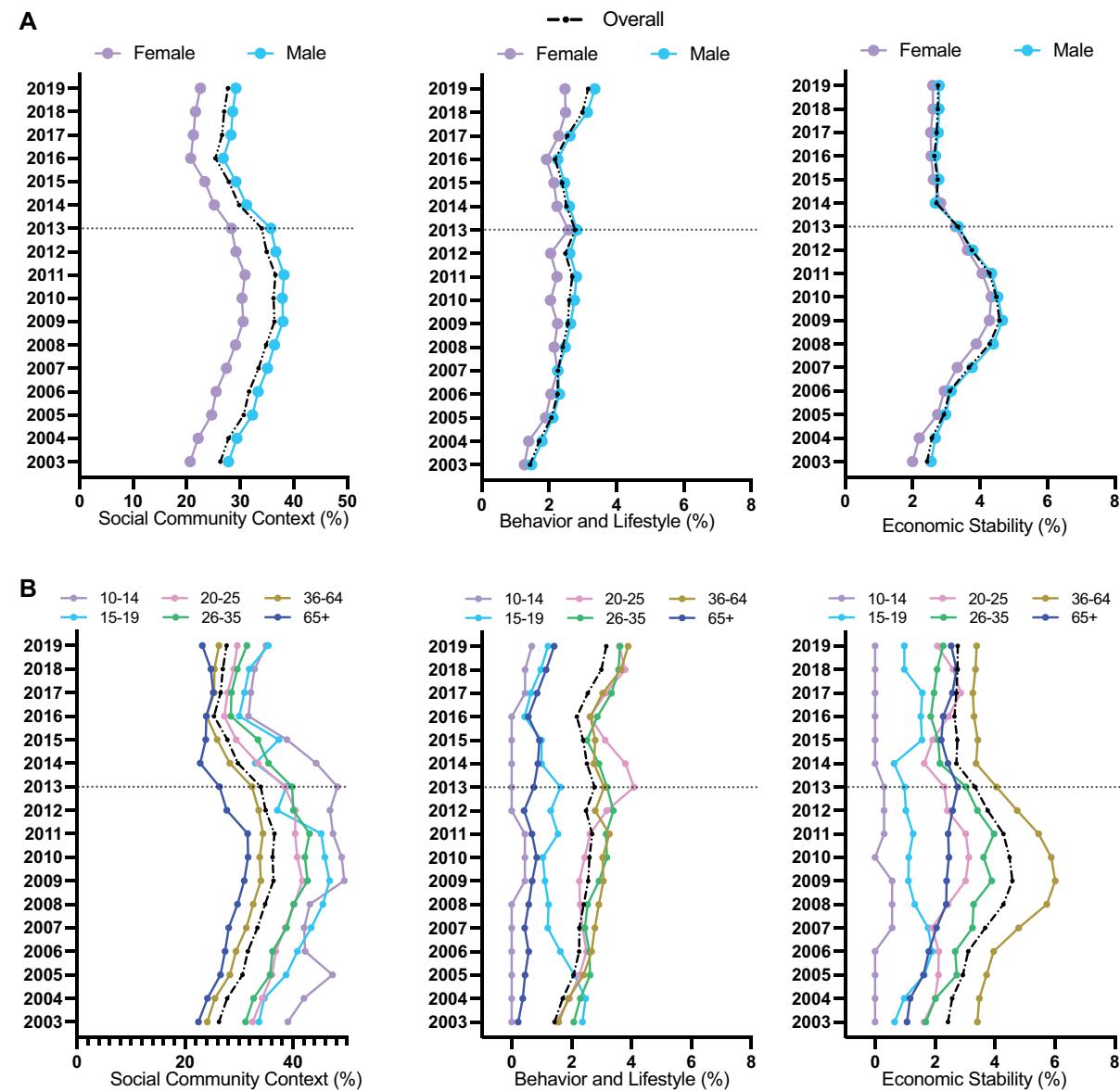
**Figure 4** shows how the 3-year rolling average crisis rates change among sex and age groups. **Figure 4A** compares the crisis rates by sex group. For the social community context crisis, the crisis rates among males are higher than those of females, and the overall population's crisis rates remained stable between 2003 and 2019. For the behavior and lifestyle crisis, the crisis rate difference between males and females is less significant, but from 2003 to 2019, the overall population's crisis rates slightly increased. From 2005 to 2009, the rates of the economic stability crisis among males and females significantly increased by a relative 50%.

**Figure 4B** compares the crisis rates by age group. We observed that the rates of the social community context crisis among adolescents aged 10–14 are higher than other age groups, while suicide

**Table 2.** Results of the SDoH-related circumstance and suicide crisis classification

	Circumstance				Crisis			
	P	R	F	AUC	P	R	F	AUC
<b>16 class</b>								
Intimate partner problem	0.839	0.950	0.891	0.975	0.575	0.853	0.687	0.953
Family relationship problem	0.628	0.746	0.682	0.929	0.440	0.656	0.527	0.949
Other relationship problem	0.338	0.441	0.383	0.879	0.204	0.282	0.237	0.912
Mental health problem	0.877	0.967	0.920	0.969	0.259	0.329	0.290	0.842
Recent suicide of friend or family	0.648	0.839	0.731	0.963	0.325	0.424	0.368	0.924
Disaster exposure	0.310	0.391	0.346	0.949	0.500	0.167	0.250	0.914
Criminal legal problem	0.658	0.888	0.756	0.967	0.523	0.764	0.621	0.966
Civil legal problem	0.429	0.637	0.513	0.932	0.298	0.459	0.361	0.926
Job problem	0.737	0.899	0.810	0.970	0.483	0.638	0.550	0.944
Physical health problem	0.774	0.907	0.835	0.964	0.472	0.649	0.547	0.933
School problem	0.700	0.863	0.773	0.981	0.364	0.708	0.481	0.968
Alcohol problem	0.832	0.932	0.879	0.975	0.380	0.557	0.451	0.939
Substance abuse	0.748	0.912	0.822	0.966	0.271	0.412	0.327	0.908
Other addiction	0.448	0.496	0.471	0.910	0.421	0.400	0.410	0.949
Financial problem	0.749	0.877	0.808	0.972	0.336	0.510	0.405	0.937
Eviction or loss of home	0.591	0.806	0.682	0.964	0.449	0.713	0.551	0.959
Weighted avg	0.780	0.906	0.837	0.966	0.476	0.692	0.563	0.942
<b>7 class</b>								
Safety concern	0.857	0.941	0.897	0.974	0.593	0.843	0.696	0.953
Interpersonal support	0.631	0.712	0.669	0.912	0.422	0.630	0.505	0.942
Mental health	0.884	0.963	0.922	0.971	0.199	0.431	0.272	0.870
Adverse life experience	0.693	0.833	0.757	0.947	0.508	0.728	0.598	0.947
Stress	0.801	0.910	0.852	0.955	0.455	0.703	0.552	0.922
Substance use	0.853	0.920	0.886	0.968	0.303	0.569	0.395	0.925
Financial distressing	0.758	0.857	0.804	0.958	0.412	0.631	0.498	0.944
Weighted avg	0.821	0.910	0.863	0.962	0.483	0.723	0.578	0.938
<b>3 class</b>								
Social community context	0.967	0.992	0.979	0.943	0.567	0.823	0.671	0.877
Behavior and lifestyle	0.879	0.902	0.890	0.964	0.350	0.522	0.419	0.908
Economic stability	0.811	0.817	0.814	0.946	0.400	0.595	0.478	0.939
Weighted avg	0.932	0.955	0.944	0.948	0.535	0.779	0.634	0.885

F: F-1 score; P: precision score; R: recall score.



**Figure 4.** Crisis trends of the social community context crisis, behavior and lifestyle crisis, and economic stability crisis with reference to sex (A) and age (B). Overall population's crisis rates are shown as black dotted curves.

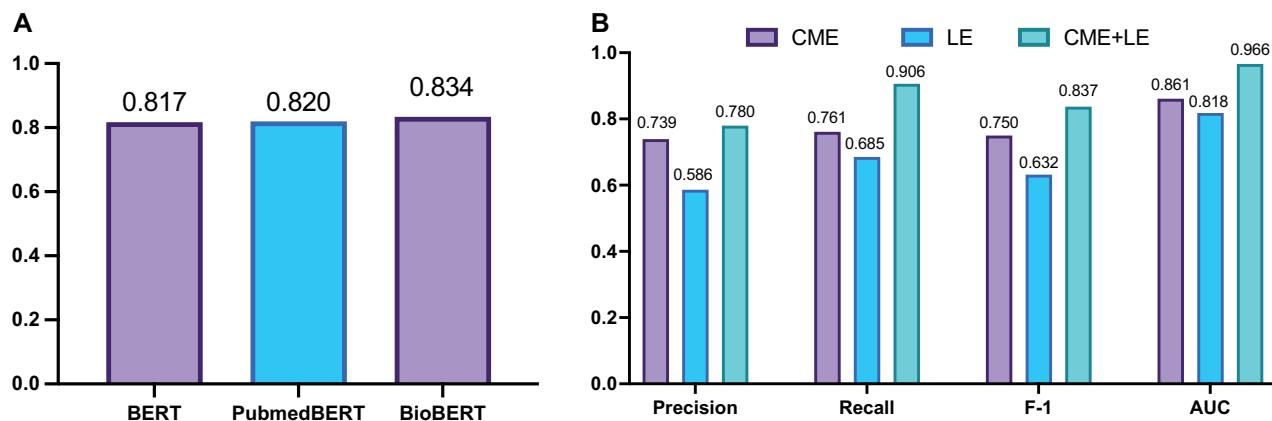
victims aged over 65 have the lowest rates. We also observed that the rates of the behavior and lifestyle crisis (eg, substance abuse) among adolescents (aged 10–19) and late adulthood (aged over 65) are lower than those among early to late middle adulthood (aged 20–64). Meanwhile, suicide victims aged 36–64 have the highest rates related to the economic stability crisis. Compared with the overall population's crisis rates, our results suggest that victims aged 36–64 may suffer more from economic instability during the Great Recession period between 2007 and 2009.

Note that although the ground truth circumstance labels from 2003 to 2019 were available, the ground truth crisis labels were only available for years after 2013. To make a fair comparison of the circumstance classification task and the crisis classification task, we trained both the circumstance classifiers and the crisis classifiers only using NVDRS data after 2013. To address the concerns about the model's performance gap on data after 2013 and on data prior to 2013,

Supplementary Figure SA1 shows how the model's ROC-AUC score changes from 2003 to 2019 in the circumstance classification task. An ROC-AUC score boost was observed in 2018, while the ROC-AUC scores remained consistent among other years, which may suggest that the distribution shift in the data after 2013 and the data prior to 2013 is not significant in terms of its impact on classification performance.

#### State-wise comparisons

To study the impacts of different documentation practices across US states, a state-level analysis was further conducted on the test set by computing the weighted average F-1 scores of the circumstance classification task (Supplementary Table SA2) and the crisis classification task (Supplementary Table SA3) per state. In the circumstance classification task, the states that have the highest, second highest and lowest weighted average F-1 scores are Virginia (0.906), North Carolina (0.898), and Hawaii (0.719), respectively. After a deeper



**Figure 5.** F-1 score comparisons on BERT variations (A) and CME or LE narratives (B). BERT: Bidirectional Encoder Representations from Transformers; CME: coroner or medical examiner; LE: law enforcement.

**Table 3.** Error cases

Death investigation record	Ground truth	Prediction
... According to this report, V had been separated from his wife for a few months and they were in the process of a divorce. V commented to a friend within the last week of his life that he was lonely; however, V's estranged wife advised that the dissolution of their marriage was V's idea and V had seemed much happier since they decided to end their relationship. ...	Intimate partner crisis	
... Earlier in the day the victim's mother received an email stating the victim was failing one of her classes. Her mother picked her up and lectured her all the home. The two argued and the victim stormed into the house. ...	Family relationship crisis School problem crisis	Family relationship crisis

investigation into Hawaii's class-wise F-1 scores, we found that Hawaii's F-1 scores of minority circumstances (eg, other addiction, eviction or loss of home) were lower than other states. Compared to the circumstance classification task, the state-wise performance variance is more significant in the crisis classification task, whose highest weighted average F-1 score is 0.700 (Utah), and lowest weighted average F-1 score is 0.287 (Rhode Island). It is worth noting that compared to other states, Rhode Island's F-1 scores of majority crisis (eg, intimate partner problem, physical health problem, financial problem) were significantly lower. This echoes that the crisis classification presents a more challenging task than the circumstance classification task and reveals that different documentation practices across states can introduce challenges to training one universal model.

### Backbone model comparisons

Our method employs the BERT model as the backbone; hence the model performance can vary when using different pretrained BERT models. To explore the impact of backbone models, we selected and compared 3 pretrained BERT models, including the basic BERT,<sup>34</sup> PubMedBERT,<sup>35</sup> and BioBERT.<sup>32</sup> We trained 3 16-class multilabel SDoH-related circumstance classification models using different backbones and compared their F-1 scores (Figure 5A). The BioBERT model yielded the best F-1 score of 0.834, which is 1.3% better than PubMedBERT and 1.7% better than base BERT.

### Inference using only CME or LE reports

During training and evaluation, the concatenations of each victim's CME and LE reports are fed to the model as input. To study whether the CME or LE report contains more information necessary to make a correct decision, we fed them separately to the trained 16-

class multilabel SDoH-related circumstance classifier and compared the classification results (Figure 5B). Though it is not surprising to observe a noticeable performance drop compared to using the concatenation of both reports, using the CME report alone shows a better performance than solely using the LE report.

### Classification error analysis

In the SDoH-related circumstance classification task, compared to the 16-class and 7-class multilabel classifiers, the 3-class multilabel classifier shows the best classification performance (ie, F-1 score). Similarly, in the suicide crisis classification task, the 3-class multilabel classifier performs better than the other 2 classifiers. There is a classification performance gap between the SDoH-related circumstance classification task and the suicide crisis classification task.

By analyzing the error cases where the suicide crisis classifier failed to make a correct classification, we show that suicide crisis classification is challenging. Table 3 shows 2 error case examples. In the first example, our suicide crisis classifier failed to classify the intimate partner problem to be a suicide crisis, although both the CME and LE records described the ongoing marriage separation between the victim and his wife. In the second example, the suicide crisis classifier correctly identified that the family relationship problem was a suicide crisis (eg, "Earlier in the day the victim and her mother were reported to have been arguing"), but the classifier was not convinced that the school problem was also a suicide crisis.

## DISCUSSION

In Table 2, we observe that not every circumstance variable can be classified equally well by our classifiers, where our classifiers

demonstrated a better classification performance on circumstances like mental health problems and intimate partner problems, while further improvements would be needed for circumstances like disaster exposure, other relationship problems, and other addiction problems. This performance gap can be closely related to the imbalanced training data distribution, which makes classifying the minority circumstance classes more challenging. The comparisons between the circumstance classification and the crisis classification in **Table 2** suggest that classifying crisis tends to be a more challenging task than classifying circumstances. Thus, a deeper understanding of the temporal contexts is necessary to differentiate between crisis and circumstances.

Through the crisis trend analysis, we observed that *not everyone is equally affected by SDoH-related social risks*. For example, the social community context crisis rates among adolescents are significantly higher than among other groups (**Figure 4B**). This finding may elucidate the necessity of social support that can protect adolescents during this vulnerable life stage.

For the economic stability crisis, our result shows a significant increase in crisis rate between 2007 and 2009, in parallel with the Great Recession period. While the crisis rates increased for both sex groups, we found that these increases are particularly significant among working-age victims (aged 26–64). The observed concentration of increase among working-age suicide victims is consistent with previous research,<sup>36</sup> which suggests a potential protective intervention—helping the newly unemployed return to work.

Finally, it is noteworthy that the crisis variables were not available for data before 2013 in the NVDRS dataset. In this study, we applied our NLP model retrospectively to automatically code the crisis variables using death investigation narratives for data between 2003 and 2013. The observations in **Figure 4** are consistent with previous studies that there has been a substantial increase in “economic suicides” during the Great Recession afflicting the United States.<sup>37</sup> Such observations indicate that our NLP approach can enrich the data in NVDRS and has the potential to facilitate the understanding of suicide crises and inform the development of effective suicide prevention strategies.

## CONCLUSIONS

In this work, we adapted a Suicide-SDoHO to comprehensively represent SDoH concepts specific to suicidal descriptions. We further developed a transformer-based model that classifies SDoH-related circumstances and suicide crises in free-text death investigation narratives. Through experiments, we demonstrated our model’s capability of accurately classifying SDoH-related circumstances and suicide crises in various granularity settings. We further applied our model retrospectively to annotate narratives whose crisis variables were not coded in NVDRS. This is helpful to maximize the sample size and assist with analyzing the yearly crisis trends among different population groups. Prompt-based learning methods transform the classification task into a masked language modeling task. Prompt-based methods take as input the textual string with a one-token answer slot [MASK], where the language model maps it to a sequence of token embeddings and learns to select one answer for the [MASK] token that can be mapped to the label space. Prompting methods have shown success in various natural language tasks,<sup>38,39</sup> and hence we propose to explore prompt-based learning methods in the future to enhance the efficiency of SDoH extraction, which will benefit more preventative efforts across violent deaths. Integrating structural SDoH data from place-based features can also improve

our prediction of suicide risks beyond individual social risks documented in the NVDRS.

## FUNDING

This work was supported by the National Library of Medicine (NLM) of the National Institutes of Health (NIH) under grant number 4R00LM013001, National Institute on Aging (NIA) of NIH under grant number RF1AG072799, National Institute of Allergy and Infectious Diseases (NIAID) of NIH under grant number 1R01AI130460, National Science Foundation under grant number 2145640, and the Center for Health Economics of Treatment Interventions for Substance Use Disorder, HCV, and HIV (NIDA P30DA040500) and the National Institute for Health Care Management Research and Educational Foundation. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH and NSF.

## AUTHOR CONTRIBUTIONS

SW: Conceptualization, Methodology, Software, Validation, Investigation, Formal analysis, Writing—original draft, Writing—review & editing. YD: Conceptualization, Methodology, Software, Writing—review & editing. ZS: Software, Validation, Writing—review & editing. YD: Writing—review & editing, Funding acquisition. JP: Investigation, Supervision, Writing—review & editing. CT: Conceptualization, Investigation, Supervision, Funding acquisition, Writing—review & editing. YX: Conceptualization, Methodology, Resources, Investigation, Formal analysis, Supervision, Funding acquisition, Writing—review & editing. YP: Conceptualization, Methodology, Resources, Investigation, Formal analysis, Supervision, Funding acquisition, Writing—review & editing.

## SUPPLEMENTARY MATERIAL

Supplementary material is available at *Journal of the American Medical Informatics Association* online.

## ACKNOWLEDGMENTS

Research reported in this publication was also supported by the NSF AI Institute for Foundations of Machine Learning (IFML).

## CONFLICT OF INTEREST STATEMENT

None declared.

## DATA AVAILABILITY

The data that support the findings of this study are available from the NVDRS Restricted Access Database (RAD) at <https://www.cdc.gov/violenceprevention/datasources/nvdrs/dataaccess.html>. Restrictions apply to the availability of these data. The data are available by request for users meeting certain eligibility criteria.

## REFERENCES

1. Stone DM, Crosby AE. Suicide prevention. *Am J Lifestyle Med* 2014; 8 (6): 404–20.
2. Xiao Y, Cerel J, Mann JJ. Temporal trends in suicidal ideation and attempts among us adolescents by sex and race/ethnicity, 1991–2019. *JAMA Netw Open* 2021; 4 (6): e2113513.

3. Xiao Y, Brown TT. The effect of social network strain on suicidal ideation among middle-aged adults with adverse childhood experiences in the US: a twelve-year nationwide study. *SSM Popul Health* 2022; 18: 101120.
4. Xiao Y, Romanelli M, Lindsey MA. A latent class analysis of health lifestyles and suicidal behaviors among us adolescents. *J Affect Disord* 2019; 255: 116–26.
5. Centers for Disease Control and Prevention National Center for Injury Prevention and Control. Web-Based Injury Statistics Query and Reporting System (WISQARS). Fatal Injury Reports. [www.cdc.gov/injury/wisqars](http://www.cdc.gov/injury/wisqars). Accessed February 9, 2023.
6. Gebbia R, Moutier C. The American Foundation for Suicide Prevention (AFSP). In: Wasserman D, ed. *Oxford Textbook of Suicidology and Suicide Prevention*. 2nd ed. Oxford Textbooks in Psychiatry. Oxford: Oxford University Press; 2021: 781–4. doi: [10.1093/med/9780198834441.003.0093](https://doi.org/10.1093/med/9780198834441.003.0093).
7. Daniel H, Bornstein SS, Kane GC, et al. Addressing social determinants to improve patient care and promote health equity: an American College of Physicians Position Paper. *Ann Intern Med* 2018; 168 (8): 577–8.
8. Himmelstein DU, Woolhandler S. Determined action needed on social determinants. *Ann Intern Med* 2018; 168 (8): 596–7.
9. Xiao Y, Hinrichs R, Johnson N, et al. Suicide prevention among college students before and during the COVID-19 pandemic: protocol for a systematic review and meta-analysis. *JMIR Res Protoc* 2021; 10 (5): e26948.
10. Xiao Y, Lu W. Temporal trends and disparities in suicidal behaviors by sex and sexual identity among Asian American adolescents. *JAMA Netw Open* 2021; 4 (4): e214498.
11. Singh GK, Daus GP, Allender M, et al. Social determinants of health in the United States: addressing major health inequality trends for the nation, 1935–2016. *Int J MCH AIDS* 2017; 6 (2): 139–64.
12. CDC—Social Determinants of Health. <https://www.cdc.gov/socialdeterminants/index.html>. Accessed December 1, 2022.
13. Shiu KY, Naumann RB, Proescholdbell S, et al. Differences in overdose deaths by intent: unintentional & suicide drug poisonings in North Carolina, 2015–2019. *Prev Med* 2022; 163: 107217.
14. Blizinsky KD, Bonham VL. Leveraging the learning health care model to improve equity in the age of genomic medicine. *Learn Health Syst* 2018; 2 (1): e10046.
15. Ertl A, Sheats KJ, Petrosky E, et al. Surveillance for violent deaths—National Violent Death Reporting System, 32 States, 2016. *MMWR Surveill Summ* 2019; 68 (9): 1–36.
16. Friday JC. National Violent Death Reporting System coding manual. *stacks.cdc.gov*. 2008. <https://stacks.cdc.gov/view/cdc/13114>. Accessed December 5, 2022.
17. Khurana D, Koli A, Khatter K, et al. Natural language processing: state of the art, current trends and challenges. *Multimed Tools Appl* 2023; 82 (3): 3713–44.
18. Patra BG, Sharma MM, Vekaria V, et al. Extracting social determinants of health from electronic health records using natural language processing: a systematic review. *J Am Med Inform Assoc* 2021; 28 (12): 2716–27.
19. Bejan CA, Angiolillo J, Conway D, et al. Mining 100 million notes to find homelessness and adverse childhood experiences: 2 case studies of rare and severe social determinants of health in electronic health records. *J Am Med Inform Assoc* 2018; 25 (1): 61–71.
20. Greenwald JL, Cronin PR, Carballo V, et al. A novel model for predicting rehospitalization risk incorporating physical function, cognitive status, and psychosocial support using natural language processing. *Med Care* 2017; 55 (3): 261–6.
21. Blosnich JR, Montgomery AE, Dichter ME, et al. Social Determinants and Military Veterans' suicide ideation and attempt: a cross-sectional analysis of electronic health record data. *J Gen Intern Med* 2020; 35 (6): 1759–67.
22. Hatlef E, Rouhizadeh M, Tia I, et al. Assessing the availability of data on social and behavioral determinants in structured and unstructured electronic health records: a retrospective analysis of a multilevel health care system. *JMIR Med Inform* 2019; 7 (3): e13802.
23. Kim HM, Smith EG, Ganoczy D, et al. Predictors of suicide in patient charts among patients with depression in the Veterans Health Administration Health System: importance of prescription drug and alcohol abuse. *J Clin Psychiatry* 2012; 73 (10): e1269–75.
24. Mowery D, South B, Patterson O, et al. Investigating the documentation of electronic cigarette use in the Veteran Affairs Electronic Health Record: a pilot study. In: *BioNLP 2017*. Vancouver, Canada: Association for Computational Linguistics; 2017: 282–6.
25. Hollister BM, Restrepo NA, Farber-Eger E, et al. Development and performance of text-mining algorithms to extract socioeconomic status from de-identified electronic health records. *Pac Symp Biocomput* 2017; 22: 230–41.
26. Feller DJ, Bear Don't Walk Iv OJ, Zucker J, et al. Detecting social and behavioral determinants of health with structured and free-text clinical data. *Appl Clin Inform* 2020; 11 (1): 172–81.
27. Lybarger K, Ostendorf M, Yetisgen M. Annotating social determinants of health using active learning, and characterizing determinants using neural event extraction. *J Biomed Inform* 2021; 113: 103631.
28. Lindemann EA, Chen ES, Wang Y, et al. Representation of social history factors across age groups: a topic analysis of free-text social documentation. *AMIA Annu Symp Proc* 2017; 2017: 1169–78.
29. Wang L, Lakin J, Riley C, et al. Disease trajectories and end-of-life care for dementias: latent topic modeling and trend analysis using clinical notes. *AMIA Annu Symp Proc* 2018; 2018: 1056–65.
30. Afshar M, Joyce C, Dligach D, et al. Subtypes in patients with opioid misuse: a prognostic enrichment strategy using electronic health record data in hospitalized patients. *PLoS One* 2019; 14 (7): e0219717.
31. Dang Y, Li F, Hu X, et al. Systematic design and evaluation of social determinants of health ontology (SDoHO). *arXiv* [cs.CY], 2022. <http://arxiv.org/abs/2212.01941>, preprint: not peer reviewed.
32. Lee J, Yoon W, Kim S, et al. BioBERT: a pre-trained biomedical language representation model for biomedical text mining. *Bioinformatics* 2020; 36 (4): 1234–40.
33. Kingma DP, Ba J. Adam: a method for stochastic optimization. In: *Proceedings of the 3rd International Conference on Learning Representations, ICLR 2015*; San Diego, CA; 2015.
34. Devlin J, Chang M-W, Lee K, et al. BERT: pre-training of deep bidirectional transformers for language understanding. In: *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*. Minneapolis, MN: Association for Computational Linguistics; 2019: 4171–86. doi: [10.18653/v1/N19-1423](https://doi.org/10.18653/v1/N19-1423).
35. Gu Y, Tinn R, Cheng H, et al. Domain-specific language model pretraining for biomedical natural language processing. *ACM Trans Comput Healthc* 2021; 3: 1–23.
36. Lewis G, Sloggett A. Suicide, deprivation, and unemployment: record linkage study. *BMJ* 1998; 317 (7168): 1283–6.
37. Reeves A, McKee M, Stuckler D. Economic suicides in the great recession in Europe and North America. *Br J Psychiatry* 2014; 205 (3): 246–7.
38. Wang S, Tang L, Majety A, et al. Trustworthy assertion classification through prompting. *J Biomed Inform* 2022; 132: 104139.
39. Liu P, Yuan W, Fu J, Jiang Z, Hayashi H, Neubig G. Pre-train, prompt, and predict: a systematic survey of prompting methods in natural language processing. *ACM Comput Surv* 2023; 55 (9): 1–35. doi: [10.1145/3560815](https://doi.org/10.1145/3560815).