Identifying Suicidal Ideation and Attempt From Clinical Notes Within a Large Integrated Health Care System

Fagen Xie, PhD; Deborah S Ling Grant, PhD, MPH, MBA; John Chang, MPH; Britta I Amundsen, LMFT; Rulin C Hechter, MD, PhD
Perm J 2022;26:21.102 • E-pub: 04/05/2022 • https://doi.org/10.7812/TPP/21.102

Abstract

PURPOSE: The purpose of this study was to develop a natural language processing algorithm to identify suicidal ideation/attempt from free-text clinical notes.

METHODS: Clinical notes containing prespecified keywords related to suicidal ideation/attempts from 2010 to 2018 were extracted from our organization’s electronic health record system. A random sample of 864 clinical notes was selected and equally divided into 4 subsets. These subsets were reviewed and classified as 1 of the following 3 suicidal ideation/attempt categories (current, historical, and no) by experienced research chart abstractors. The first 3 data sets were used to develop the rule-based computerized algorithm sequentially and the fourth data set was used to evaluate the algorithm’s performance. The validated algorithm was then applied to the entire study sample of clinical notes.

RESULTS: The computerized algorithm correctly identified 23 of the 26 confirmed current suicidal ideation/attempts and all 10 confirmed historical suicidal ideation/attempts in the validation data set. It produced an 88.5% sensitivity and a 100.0% positive predictive value for current suicidal ideation/attempts, and a 100.0% sensitivity and positive predictive value for historical suicidal ideation/attempts. After applying the computerized algorithm to the entire set of study notes, we identified a total of 1,050,287 current ideation/attempt events and 293,037 historical ideation/attempt events documented in clinical notes. Those for which current ideation/attempt events were documented were more likely to be female (59.5%), 25–44 years old (28.3%), and White (43.4%).

CONCLUSION: Our study demonstrated that a computerized algorithm can effectively identify suicidal ideation/attempts from clinical notes. This algorithm can be utilized in support of suicide prevention research programs and patient care quality improvement initiatives.

Corresponding Author
Fagen Xie, PhD
fagen.xie@kp.org

Author Affiliations
Department of Research and Evaluation, Kaiser Permanente Southern California, Pasadena, CA, USA

Author Contributions:
Fagen Xie, PhD, Deborah S Ling Grant, PhD, MPH, MBA, and Rulin C Hechter, MD, PhD, conceived and led the design of the study and drafted the manuscript; Fagen Xie extracted the data and conducted all of the analysis. John Chang, MPH, and Britta I Amundsen, LMFT, conducted chart reviews of the training and validation data sets. Deborah S Ling Grant validated and confirmed the chart review results. All authors participated in interpreting the study results, critically reviewed the manuscript, and have given final approval of the manuscript.

Disclosures
Conflicts of Interest: None declared.
Funding: This research study was funded by Kaiser Permanente’s Office of Community Health as part of its Firearm Injury Prevention Program. The views and opinions expressed in this article are the responsibility of the authors and do not necessarily represent the official views of Kaiser Permanente.

Introduction

Suicide is a devastating outcome of psychiatric illness and a critical public health problem worldwide. It is the 10th leading cause of loss of life in the United States and the second leading cause of death among individuals aged 15–24 years and members of the US military. Despite extensive...
suicide prevention research efforts, suicide rates in the United States have continued to increase in recent years.5,6 Suicide is a complex process that involves a series of pathways and mechanisms from thoughts of killing oneself, planning, and finally attempt.7 Suicidal behavior differs between genders, age groups, geographic regions, and sociopolitical settings.8 In addition to increasing health care clinician training and restricting access to lethal means, evidence-based screening using validated, population- and setting-specific assessment tools is recommended for patients in all medical settings to detect early suicidal thoughts or preparatory behaviors before an actual attempt.9,10

Previous studies have used survey instruments or clinical diagnoses from health care administrative or claims data to identify suicidal behaviors.11,12 The recent exponential growth in the utilization of electronic health records (EHRs) has provided unprecedented opportunities for researching documented suicidal behaviors using an automated health care database.13 However, a systematic review has reported a low sensitivity (13.8%–65%) for identifying suicide or suicidal ideation using structured administrative or claims data, because suicidal behaviors are often undercoded.14 Utilizing information other than diagnostic codes or questionnaires to identify populations at risk of suicide can mitigate the misclassification bias and facilitate population-based mental health research studies that examine the risk factors for suicide and inform the development of targeted suicide prevention programs.14 To fully utilize the information on suicidal behaviors documented in unstructured free-text clinical notes and improve the efficiency of chart review, natural language processing (NLP) can automate the examination15 and convert information residing in free-text natural language into a more structured format for analysis.16

NLP techniques have been recently utilized to detect suicidality and suicidal behaviors from suicide notes,17 psychiatric clinical databases,18 social media,19 and clinical notes describing US veterans,20,21 adolescents with autism spectrum disorders,22 and pregnant women.23 The combination of advanced machine learning and NLP techniques has also been explored to classify and predict suicidal thoughts and behaviors from both structured24–26 and unstructured data.19,21,27–29 However, studies focused on detecting suicide attempts and suicidal ideation among a large, diverse population in a general medical setting are limited. In this study, we sought to develop and validate a rule-based computerized algorithm and automated process to correctly identify suicidal behaviors (ideation/attempt) from clinical notes in the comprehensive EHR of Kaiser Permanente Southern California, a large integrated health care system.

Methods

STUDY SETTING AND POPULATION
With 15 hospitals and more than 230 affiliated medical offices, Kaiser Permanente Southern California is an integrated health care system that provides complete medical services to over 4.7 million members, including ambulatory primary and specialty care, inpatient care, and emergency care, as well as laboratory, immunization, and pharmacy services. The diverse demographic characteristics of Kaiser Permanente Southern California members are largely representative of the residents in the Southern California region.30 Kaiser Permanente Southern California’s extensive EHR system captures both structured and unstructured data documented by health care clinicians at various care settings, including non-Kaiser Permanente facilities. The structured data contain diagnosis codes, procedure codes, information on pharmacy dispenses, immunizations, laboratory results, pregnancy episodes, and outcomes. The unstructured data include free-text clinical notes, radiology reports, pathology reports, imaging, videos, and more. The protocol for this study was reviewed and approved by the Kaiser Permanente Southern California Institutional Review Board with a waiver of the requirement for informed consent.

SUICIDAL THOUGHTS AND BEHAVIORS KEYWORD SELECTION
The Kaiser Permanente Southern California EHR system captures billions of clinical notes created by clinicians during clinical care. To reduce the sheer volume of notes and process them efficiently, our study only extracted the clinical notes indicating suicidal behavior for analysis. To retrieve relevant clinical notes for processing, keywords and phrases related to suicidal ideation and behaviors were compiled through consultations with mental health clinicians, diagnosis definitions, and ontologies in the Unified Medical Language System.31 A suicidal ideation/attempt term was composed of a compound suicide-related term and thought- or behavior-related term, such as “suicide attempt”, “suicide ideas,” and “suicide plan.” The terms associated with suicide included
Identifying Suicidal Ideation and Attempt From Clinical Notes Within a Large Integrated Health Care System

<table>
<thead>
<tr>
<th>Term category</th>
<th>Keywords and phrases</th>
</tr>
</thead>
<tbody>
<tr>
<td>Suicide</td>
<td>suicide, suicidal, suicidality, parasuicide, parasuicidal, parasuicidality, kill self, kill myself, kill yourself, kill herself, kill himself</td>
</tr>
<tr>
<td>Behavior</td>
<td>behavior, action, attempt, idea, ideation, thought, think, tendency, plan, intent, intention, urge, overdose</td>
</tr>
<tr>
<td>Past history section</td>
<td>past medical history, previous medical history, patient medical history, past psychiatric history, past psychiatric treatment, past psych history, past psych treatment, past medical, past psychotropics, medical history, pmh, psh, pmhx, phx</td>
</tr>
<tr>
<td>Diagnosis section</td>
<td>diagnosis, diagnose, dx, primary diagnosis, primary encounter diagnosis, primary problem, psychiatric problem</td>
</tr>
</tbody>
</table>

Table 1: Keywords or phrases associated with suicidal ideation or behaviors

*The algorithm also searched abbreviations, misspellings, and variants (plural form and verb tenses) of these words, where applicable.

*Used for algorithm development.

“suicide,” “suicidality,” and “kill self,” while the terms associated with suicidal thoughts or behaviors included “ideation,” “ideas,” and “think,” among others. The complete list of the associated terms used for the study is summarized in Table 1. Potential lemma variants, abbreviations, and misspellings identified through manual chart review and algorithm development were also included in searches. For example, “ideation” can be abbreviated as “id,” or “attempt” might be misspelled as “attemt.” The detailed list of variants is summarized in Supplemental Material Table S1.

CLINICAL NOTE EXTRACTION AND PREPROCESSING
Clinical notes containing at least 1 of the suicide-related terms listed in Table 1 were extracted from the EHR data for members who received care at Kaiser Permanente Southern California between January 1, 2010, and December 31, 2018, via the SAS PROC SQL. The extracted clinical notes contained the detailed descriptions written by health care clinicians (eg, mental health clinicians, primary care physicians, social workers) regarding the patient’s condition at the time of the visit. The length of the clinical notes varied according to specific medical conditions. Any patient instructional notes were excluded, as these notes usually provide general warnings, instructions, or recommendations regarding medical conditions rather than make diagnoses or describe the patient’s symptoms. A total of 16,642,120 clinical notes with suicide-related terms were retrieved during the study period. These extracted clinical notes were then transferred to a Linux server and preprocessed through sentence separation and tokenization (ie, segmenting text into linguistic units such as words and punctuations) via Python programming, the Natural Language Toolkit, and a customized sentence boundary detection algorithm. For example, the special symbols “¶” in the clinical notes indicated the end of sections and sentences in the EHR system. The clinical notes were further standardized using the word tokens listed in Supplemental Material Table S1.

TRAINING DATA SET, VALIDATION DATA SET, AND REFERENCE STANDARD
A representative sample of 864 clinical notes was randomly selected from the entire set of preprocessing clinical notes and equally divided into 4 subsets, each containing 216 notes. These subsets were first sequentially reviewed and each note was classified as 1 of the following 3 suicidal ideation/attempt categories by experienced research chart abstractors: current, historical, and no. If evidence of both current and historical suicidal ideation was presented in the same note, then the note was classified as current to identify the more severe suicide risk situation. The chart review results were then validated and confirmed by another independent, trained chart abstractor. Cases that could not be classified by the abstractors were further reviewed and adjudicated by the study’s principal investigator and project manager with additional consultations from physician experts. Current episodes included notes that mentioned a suicidal ideation/attempt event within 2 weeks of the note date. Historical episodes were defined as documented suicidal ideation/attempt events that occurred more than 2 weeks before the note date. Events were classified as no if there was no evidence of an actual suicidal ideation/attempt event in the note. The manual review results were deemed the reference standard. The results of the first 3 subsets were used for algorithm development, while the rest of the clinical notes were used to evaluate the algorithm performance.

NLP COMPUTERIZED ALGORITHM
A rule-based computerized algorithm was developed through an iterative process to ascertain the status of suicidal ideation/attempt for each clinical note using the training datasets. With each iteration, the algorithm was refined to match the results of the reference standards that were derived through manual review and adjudication of the 3 training subsets sequentially. The process first searched suicidal ideation/attempt at the sentence level and then aggregated at the note and date levels. More specifically, the steps were to detect...
suicidal ideation/attempts within each clinical note as shown in Figure 1.

Sentences were excluded for the following:

- No suicide-related term from the list in Table 1
- Someone other than the patient had suicidal ideation/attempts. For example, “patient’s daughter has been acting out and voicing suicidal thoughts.” Terms used for searching nonpatient self-description are summarized in Supplemental Material Table S2.
- Negated or uncertain description of suicidal ideation/attempts. Examples include “patient denies any suicidal ideation,” “negative for suicide ideas,” “no suicidal thinking,” and “may have suicidal ideation.” The list of negated or uncertain terms in the pyConTextNLP36 was applied and subsequently enhanced by previous studies34,35 as well as during the training process. Terms used in searching for negated or uncertain descriptions are summarized in Supplemental Material Table S2.
- Suicidal ideation/attempts description did not refer to an actual situation, such as “discussed risks, benefits, and side effects, including but not limited to suicidal thinking,” and “in case of having any suicidal ideas.” Terms used in searching for general descriptions are summarized in Supplemental Material Table S2.

The status of suicidal ideation/attempts was classified as historical for any of the following situations:

- Sentence contained a question about a “history of suicidal thoughts,” and the corresponding answer was “yes”
- Terms related to suicidal ideation/attempts appeared in the history section of a clinical note, such as “past medical history: suicidal attempt.” A complete list of terms applicable to the history section is presented in Table 1.
- Sentence contained terms indicating historical events (events that occurred more than 2 weeks before the clinical note date). Clinical notes typically contained information describing the patient’s history of medical conditions. The temporality terms regarding historical events included the lists in the pyConTextNLP36 which was enriched based on previous studies34,35 and the training data sets. Examples include “one suicide attempt in teen years,” “tried to commit suicide several years ago.” Terms used in searching the history description are summarized in Supplemental Material Table S2.
- Sentence contained a specific time description (i.e., date, year, month, weeks, days) associated with a term indicative of suicidal ideation/attempt, and this date was 2 weeks before the clinical note date. This would apply to a note that read “patient mentioned having suicidal thoughts 4 weeks ago,” for example.

---

**Figure 1:** Flow diagram of the natural language process to identify suicidal ideation/attempts events.

### Table 1: Excluded Sentences

- Sentence contains a definite term prior to a suicidal ideation/attempt term
- Sentence contains a definite term after a suicidal ideation/attempt term
- Sentence contains at least 1 of the following diagnosis codes related to suicide ideation/attempts: V62.84, R45.851, or E958.9
- Sentence contains a specific date associated with suicidal ideation/attempts which is at least 2 weeks prior to the clinical note date
- Sentence indicates suicidal ideation/attempts within 2 weeks
- Suicide risk assessment included in sentence and risk was moderate or above
- Sentence indicates event was historical
- Sentence contains at least 1 of the following diagnosis codes related to suicide ideation/attempts: V62.84, R45.851, or E958.9
- Sentence contains a definite term prior to other a suicidal ideation/attempt term
- Any sentences with selected keywords or terms but not classified as “current” or “historical” are classified as “no”

### Table 2: Excluded Sentences

- Sentence contains a definite term prior to a suicidal ideation/attempt term
- Sentence contains a definite term after a suicidal ideation/attempt term
- Sentence contains at least 1 of the following diagnosis codes related to suicide ideation/attempts: V62.84, R45.851, or E958.9
- Sentence contains a specific date associated with suicidal ideation/attempts which is at least 2 weeks prior to the clinical note date
- Sentence indicates suicidal ideation/attempts within 2 weeks
- Suicide risk assessment included in sentence and risk was moderate or above
- Sentence indicates event was historical
- Sentence contains at least 1 of the following diagnosis codes related to suicide ideation/attempts: V62.84, R45.851, or E958.9
- Sentence contains a definite term prior to other a suicidal ideation/attempt term
- Any sentences with selected keywords or terms but not classified as “current” or “historical” are classified as “no”

---

**Figure 1:** Flow diagram of the natural language process to identify suicidal ideation/attempts events.

---

**Table 1:**

<table>
<thead>
<tr>
<th>Step 1: Processing at sentence level</th>
<th>Step 2: Processing at note level</th>
<th>Step 3: Processing at date level</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Exclude/Classify as “No” suicidal ideation/attempts event</strong></td>
<td><strong>Exclude/Classify as “No” suicidal ideation/attempts event</strong></td>
<td><strong>Exclude/Classify as “No” suicidal ideation/attempts event</strong></td>
</tr>
<tr>
<td>Not detected</td>
<td>Not detected</td>
<td>Not detected</td>
</tr>
<tr>
<td><strong>Classify as “Historical” suicidal ideation/attempts event</strong></td>
<td><strong>Classify as “Current” suicidal ideation/attempts event</strong></td>
<td><strong>Classify as “Historical” suicidal ideation/attempts event</strong></td>
</tr>
<tr>
<td>Not detected</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td><strong>Classify as “Current” suicidal ideation/attempts event</strong></td>
<td><strong>Classify as “Current” suicidal ideation/attempts event</strong></td>
<td><strong>Classify as “Current” suicidal ideation/attempts event</strong></td>
</tr>
<tr>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td><strong>Classify as “No” suicidal ideation/attempts event</strong></td>
<td><strong>Classify as “No” suicidal ideation/attempts event</strong></td>
<td><strong>Classify as “No” suicidal ideation/attempts event</strong></td>
</tr>
<tr>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>
A suicidal ideation/attempt was classified as current if 2 conditions were met: 1) a sentence contained phrases or questions asking about “suicidal thoughts within the last 2 weeks” or “thoughts of death/self-harm within the last 2 weeks” with a corresponding answer of “yes,” “some,” “several days,” “more than half of the days,” or “nearly every day” and 2) one of the following situations applied:

- Sentence was part of a suicide risk assessment in which the risk was deemed moderate or above
- Sentence contained a diagnosis code of “V62.84,” “R45.851,” or “E958.9”
- Terms related to suicidal ideation/attempt appeared in the diagnosis, chief complaint, likely self-harm, active problem list, or patient safety sections. Examples include “diagnosis: depression, suicidal ideation, drug overdose,” and “chief complaint: depression, suicidal behavior.” The detailed list of relevant terms in the diagnosis section is presented in Table 1.
- Sentence contained a definitive term before or after a term related to a suicidal ideation/attempt or suicide only. Examples include “positive for depression and suicidal ideas,” “has some suicidal ideation, on and off for about 2 weeks,” and “patient tried to commit suicide 2 days ago.” Terms used in searching for definite descriptions are summarized in Supplemental Material Table S2.

Suicidal ideation/attempts at the note level were prioritized using the following rubric:

1. classify the status of suicidal ideation/attempt as current if at least 1 sentence was deemed current, regardless of the historical status
2. categorize the status of suicidal ideation/attempt as historical if at least 1 of the sentences was considered historical
3. if not 1 or 2, categorize the status of this visit date as no

APPLICATION OF THE COMPUTERIZED ALGORITHM
The developed algorithm was implemented through Python programming on a Linux server to process the entire set of clinical notes. The algorithm created outputs of the suicidal ideation/attempt status at the sentence, note, and visit date levels for further analysis.

DATA ANALYSIS
The computerized algorithm developed from the training data sets was applied to the validation data set to identify suicidal ideation/attempt events at the note level. The results from the computerized algorithm and manual chart review were compared for the validation data set. Sensitivity, positive predictive value (PPV), and F-score were calculated and reported for both the current and historical categories. Sensitivity was defined as the proportion of suicidal ideation/attempt events (current or historical) correctly assigned by the computerized algorithm among all cases confirmed by manual review. PPV was defined as the proportion of correctly determined suicidal ideation/attempt events (current or historical) among all those identified by the computerized algorithm. The F-score was calculated as $2 \times \text{PPV} \times \text{Sensitivity} / (\text{PPV} + \text{Sensitivity})$. The number and percentage of suicidal ideation/attempt events (current or historical) ascertained by the computerized algorithm at the visit date level were summarized by year. The distribution of patients by the total number of current suicidal ideation/attempt cases per patient was reported for the entire study period. Finally, the demographic characteristics of individuals identified with current suicidal ideation/attempt events were examined during the study period.

Results

The classification of suicidal ideation/attempt events by manual review of the training and validation data sets is summarized in Table 2. Among the 216 clinical notes in each data set, about 8.3% to 10.6% were classified as current suicidal ideation/attempt events, and 2.8% to 4.6% were categorized as historical suicidal ideation/attempt events. Among the 216 clinical notes in the validation data set, the computerized algorithm correctly identified 23 of the 26 current suicidal ideation/attempt events and all 10 of the historical suicidal ideation/attempt
Identifying Suicidal Ideation and Attempt From Clinical Notes Within a Large Integrated Health Care System

The algorithm produced a PPV of 100.0%, a sensitivity of 88.5%, and an F-score of 0.94 for current events and a PPV of 100.0%, a sensitivity of 100.0%, and an F-score of 1.00 for historical events (Table 3).

A total of 13,946,374 encounters with clinical notes containing suicide-related terms between 2010 and 2018 were extracted. Of them, 1,050,289 encounters were identified by the validated computerized algorithm as having a current suicidal ideation/attempt event, and 293,038 were identified as having a historical event documented (Table 4). The number and percentage of identified events confirmed by manual review. The algorithm produced a PPV of 100.0%, a sensitivity of 88.5%, and an F-score of 0.94 for current events and a PPV of 100.0%, a sensitivity of 100.0%, and an F-score of 1.00 for historical events (Table 3).

A first current suicidal ideation/attempt event was more likely to be noted for patients who were female (58.7%), aged 15–24 years (28.8%), and Hispanic (41.7%). Taking into account all current events, suicidal ideation/attempt events were still most common among females (59.5%), but they were more likely to be documented for patients who were aged 25–44 years (28.3%) and White (43.4%) (Table 5). In comparison, female patients, White patients, and those 25 years or older were more likely to have a historical event than a current event documented.

Discussion

In this study, we developed a rule-based computerized algorithm for identifying current and historical suicidal ideation/attempt events from free-text clinical notes, with a high PPV and sensitivity. Our findings suggest it is feasible to apply NLP technology to identify individuals at risk for suicide using automated EHR data, and the use of the NLP algorithm can mitigate misclassification bias in research studies caused by undercoding suicidal ideation/attempts and support accurate evaluation of the association between risk factors and suicide outcomes.

Historical suicidal ideation has been revealed as a risk factor for recurrent/repeated suicidal ideation.\(^{37}\)
Previous studies reported around 30% to 40% of individuals presenting with current suicidal ideation also had a history of suicidal ideation.\textsuperscript{38,39} Our NLP algorithm identified more than 400,000 individuals with 1 or more current suicidal ideation/attempts within our study period, and 37.6% of them had repeated or recurrent events after the first event. This was consistent with the findings from previous studies.\textsuperscript{38,39}

Our computerized algorithm misclassified 3 current suicidal ideation/attempts in the validation data set. This was because the suicide-related term representing phrases “not be alive” (2 cases) and “wants to jump off bridge” (1 case) did not appear in the training data sets. Refinement to include additional suicide-related linguistics and relevant phrases can be carried out in future work to improve the algorithm’s performance. This could also indicate a need to improve clinical note drafting practices at the organizational level. For example, adding a clinical impression of current suicidal ideation would allow the computerized algorithm to pick up a case even if the clinician also documents the patient’s own idiosyncratic words.

Conducting a full manual chart review is labor intensive and time consuming. Our study included a huge volume of clinical notes that contained predefined suicidal ideation/attempt keywords and phrases. It was not feasible to manually review the entire set of study notes for the algorithm development and validation. To develop an effective algorithm, a small proportional sample (864 notes) was randomly selected from the total set of study notes. This sample was randomly and equally divided into 4 subsets. Such randomization ensures that each subset has similar characteristics and generally represents the study notes. In addition, the iterative training process tuned and produced the robust performance of the computerized algorithm. However, further validation may be considered when the developed algorithm is applied to different study settings.

It is worthwhile to point out that this algorithm is a computerized tool to identify documented suicidal ideation/attempts from free text automatically, which is markedly different than a clinical screening instrument intended to detect true cases.\textsuperscript{9,10} The performance of this tool relied on the completeness and accuracy of the clinical notes. However, clinicians may sometimes copy text descriptions from previous progress notes and apply them to the current encounter summary. It is challenging to distinguish and omit notes that were

### Table 5: Demographic characteristics of individuals at the first current, all current, and historical suicide ideation/attempt events during the study period

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>First current event (N = 400,436)</th>
<th>All current events (N = 1,050,287)</th>
<th>Historical event (N = 293,037)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>234,924 (58.7%)</td>
<td>625,382 (59.5%)</td>
<td>186,539 (63.6%)</td>
</tr>
<tr>
<td>Male</td>
<td>165,475 (41.3%)</td>
<td>424,826 (40.5%)</td>
<td>105,165 (35.9%)</td>
</tr>
<tr>
<td>Unknown</td>
<td>37 (0.0%)</td>
<td>79 (0.0%)</td>
<td>1,333 (0.5%)</td>
</tr>
<tr>
<td>Race/ethnicity</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>151,264 (37.8%)</td>
<td>455,854 (43.4%)</td>
<td>136,183 (46.5%)</td>
</tr>
<tr>
<td>Black</td>
<td>38,785 (9.7%)</td>
<td>109,375 (10.4%)</td>
<td>28,596 (9.8%)</td>
</tr>
<tr>
<td>Hispanic</td>
<td>165,084 (41.2%)</td>
<td>380,782 (36.3%)</td>
<td>101,115 (34.5%)</td>
</tr>
<tr>
<td>Asian American Pacific Islander</td>
<td>26,572 (6.7%)</td>
<td>65,557 (6.2%)</td>
<td>17,075 (5.8%)</td>
</tr>
<tr>
<td>Native American</td>
<td>1,040 (0.3%)</td>
<td>2,722 (0.3%)</td>
<td>724 (0.3%)</td>
</tr>
<tr>
<td>Hispanic</td>
<td>2,286 (0.6%)</td>
<td>6,095 (0.6%)</td>
<td>1,639 (0.6%)</td>
</tr>
<tr>
<td>Other</td>
<td>4,393 (1.1%)</td>
<td>10,552 (1.0%)</td>
<td>2,918 (1.0%)</td>
</tr>
<tr>
<td>Unknown</td>
<td>11,012 (2.8%)</td>
<td>19,350 (1.8%)</td>
<td>4,787 (1.6%)</td>
</tr>
<tr>
<td>Age (year)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt; 15</td>
<td>67,853 (16.9%)</td>
<td>102,723 (9.8%)</td>
<td>15,717 (5.4%)</td>
</tr>
<tr>
<td>15-24</td>
<td>115,197 (28.8%)</td>
<td>284,282 (27.1%)</td>
<td>68,262 (23.3%)</td>
</tr>
<tr>
<td>25-44</td>
<td>102,409 (25.6%)</td>
<td>296,893 (28.3%)</td>
<td>92,357 (31.5%)</td>
</tr>
<tr>
<td>45-64</td>
<td>81,204 (20.3%)</td>
<td>268,898 (25.6%)</td>
<td>82,762 (28.2%)</td>
</tr>
<tr>
<td>≥ 65</td>
<td>33,773 (8.4%)</td>
<td>97,491 (9.3%)</td>
<td>33,939 (11.6%)</td>
</tr>
</tbody>
</table>
Identifying Suicidal Ideation and Attempt From Clinical Notes Within a Large Integrated Health Care System

carried over if there are no specific indicators, such as the encounter date associated with the previous event. Therefore, some historical events might be misclassified as current due to the imprecision of the clinical notes.

In recent years, machine learning techniques used alone or in combination with rule-based NLP have been applied to identify suicidal behaviors for specialized populations. The performance of these algorithms varies widely, with a PPV ranging from 30% to 92% and a sensitivity ranging from 56% to 98%. The performance of our rule-based computerized algorithm was reasonable and close to the top range of these existing models. However, further research on the combination of rule-based and advanced machine learning approaches, such as random forest and deep neural network learning, may increase the sensitivity of identification of suicidal ideation/attempts in clinical notes among large, diverse populations. In addition, the NLP algorithm can be further enhanced to stratify the level of severity of suicidal ideation. Such enhancements can strengthen the clinical value of this algorithm and facilitate the development and implementation of targeted suicide prevention programs in clinical practice.

This study acknowledges several limitations. First, similar to most existing methods of detecting suicide risk, our study relies on patient self-reports of suicidal thoughts or behavior to a health care professional, and this discussion must have been accurately documented in the EHR system. Second, our computerized algorithm was limited by the prespecified search terms of interest in screening for potentially relevant clinical notes. We used this approach to increase efficiency and to avoid having to process an enormous number of clinical notes that had no mention of any terms related to suicide. Although misclassification may occur, as it is impossible to include all possible terms used by clinicians to document such events, the performance of our algorithm was satisfactory with a high sensitivity level. Third, applying this computerized algorithm in other scenarios or health care settings may yield slightly different results, and thus it may need some modification to account for variations in the format and presentation of clinical notes in different health care settings.

In this study, our accuracy remained consistent, since the rule-based algorithm was robust and not specifically limited to any fixed or strictly formatted notes. Lastly, our current study algorithm identified suicidal thoughts/ideation and suicide attempt as a single event instead of distinguishing them as 2 separate events. This distinction could be made with a future algorithm or with manual review by the algorithm’s users if needed. It is worth pointing out the nonlinear relationship between suicidal ideation and completed suicide. Suicidal thoughts can be understood as a symptom and risk factor of completed suicide. Not everyone presenting with suicidal ideation will carry out suicide. Suicidal thinkers and suicide completers were treated as 2 separate groups but can overlap over time.

Despite these limitations, we successfully developed an effective computerized algorithm to identify individuals with documented suicidal ideation/attempts from free-text clinical notes covering a large, diverse population. The algorithm was able to distinguish clinical notes addressing a current event versus a historical event. This computerized algorithm can be automated to regularly process clinical notes to build a near real-time patient database to support suicide prevention research as well as to evaluate the effectiveness of intervention programs intended to reduce suicidal ideation and suicide.

Supplementary Materials

REFERENCES


