

Original Paper

# Risk Factors Associated With Nonfatal Opioid Overdose Leading to Intensive Care Unit Admission: A Cross-sectional Study

Avijit Mitra<sup>1</sup>, MSc; Hiba Ahsan<sup>1</sup>, BTech; Wenjun Li<sup>2,3</sup>, PhD; Weisong Liu<sup>4</sup>, PhD; Robert D Kerns<sup>5,6,7,8</sup>, PhD; Jack Tsai<sup>9,10</sup>, PhD; William Becker<sup>8,11</sup>, MD; David A Smelson<sup>3,12</sup>, PsyD; Hong Yu<sup>1,3,4,13</sup>, PhD

<sup>1</sup>College of Information and Computer Sciences, University of Massachusetts Amherst, Amherst, MA, United States

<sup>2</sup>Department of Public Health, University of Massachusetts Lowell, Lowell, MA, United States

<sup>3</sup>Center for Healthcare Organization and Implementation Research, Veterans Affairs Bedford Healthcare System, Bedford, MA, United States

<sup>4</sup>Department of Computer Science, University of Massachusetts Lowell, Lowell, MA, United States

<sup>5</sup>Department of Psychiatry, Yale University School of Medicine, New Haven, CT, United States

<sup>6</sup>Department of Neurology, Yale University School of Medicine, New Haven, CT, United States

<sup>7</sup>Department of Psychology, Yale University School of Medicine, New Haven, CT, United States

<sup>8</sup>Pain Research, Informatics, Multimorbidities and Education Center, Veterans Affairs Connecticut Healthcare System, West Haven, CT, United States

<sup>9</sup>School of Public Health, University of Texas Health Science Center at Houston, Houston, TX, United States

<sup>10</sup>National Center on Homelessness Among Veterans, United States Department of Veterans Affairs, Tampa, FL, United States

<sup>11</sup>Department of Internal Medicine, Yale University School of Medicine, New Haven, CT, United States

<sup>12</sup>Department of Psychiatry, University of Massachusetts Chan Medical School, Worcester, MA, United States

<sup>13</sup>Department of Medicine, University of Massachusetts Chan Medical School, Worcester, MA, United States

**Corresponding Author:**

Hong Yu, PhD

Department of Computer Science

University of Massachusetts Lowell

1 University Avenue

Lowell, MA, 01854

United States

Phone: 1 508 612 7292

Email: [Hong\\_Yu@uml.edu](mailto:Hong_Yu@uml.edu)

## Abstract

**Background:** Opioid overdose (OD) and related deaths have significantly increased in the United States over the last 2 decades. Existing studies have mostly focused on demographic and clinical risk factors in noncritical care settings. Social and behavioral determinants of health (SBDH) are infrequently coded in the electronic health record (EHR) and usually buried in unstructured EHR notes, reflecting possible gaps in clinical care and observational research. Therefore, SBDH often receive less attention despite being important risk factors for OD. Natural language processing (NLP) can alleviate this problem.

**Objective:** The objectives of this study were two-fold: First, we examined the usefulness of NLP for SBDH extraction from unstructured EHR text, and second, for intensive care unit (ICU) admissions, we investigated risk factors including SBDH for nonfatal OD.

**Methods:** We performed a cross-sectional analysis of admission data from the EHR of patients in the ICU of Beth Israel Deaconess Medical Center between 2001 and 2012. We used patient admission data and International Classification of Diseases, Ninth Revision (ICD-9) diagnoses to extract demographics, nonfatal OD, SBDH, and other clinical variables. In addition to obtaining SBDH information from the ICD codes, an NLP model was developed to extract 6 SBDH variables from EHR notes, namely, housing insecurity, unemployment, social isolation, alcohol use, smoking, and illicit drug use. We adopted a sequential forward selection process to select relevant clinical variables. Multivariable logistic regression analysis was used to evaluate the associations with nonfatal OD, and relative risks were quantified as covariate-adjusted odds ratios (aOR).

**Results:** The strongest association with nonfatal OD was found to be drug use disorder (aOR 8.17, 95% CI 5.44-12.27), followed by bipolar disorder (aOR 2.69, 95% CI 1.68-4.29). Among others, major depressive disorder (aOR 2.57, 95% CI 1.12-5.88), being on a Medicaid health insurance program (aOR 2.26, 95% CI 1.43-3.58), history of illicit drug use (aOR 2.09, 95% CI 1.15-3.79), and current use of illicit drugs (aOR 2.06, 95% CI 1.20-3.55) were strongly associated with increased risk of nonfatal

OD. Conversely, Blacks (aOR 0.51, 95% CI 0.28-0.94), older age groups (40-64 years: aOR 0.65, 95% CI 0.44-0.96; >64 years: aOR 0.16, 95% CI 0.08-0.34) and those with tobacco use disorder (aOR 0.53, 95% CI 0.32-0.89) or alcohol use disorder (aOR 0.64, 95% CI 0.42-1.00) had decreased risk of nonfatal OD. Moreover, 99.82% of all SBDH information was identified by the NLP model, in contrast to only 0.18% identified by the ICD codes.

**Conclusions:** This is the first study to analyze the risk factors for nonfatal OD in an ICU setting using NLP-extracted SBDH from EHR notes. We found several risk factors associated with nonfatal OD including SBDH. SBDH are richly described in EHR notes, supporting the importance of integrating NLP-derived SBDH into OD risk assessment. More studies in ICU settings can help health care systems better understand and respond to the opioid epidemic.

(*JMIR Med Inform* 2021;9(11):e32851) doi: [10.2196/32851](https://doi.org/10.2196/32851)

## KEYWORDS

opioids; overdose; risk factors; electronic health records; social and behavioral determinants of health; natural language processing; intensive care unit

## Introduction

The opioid epidemic in the United States is one of the most severe public health emergencies in recent times, with opioid overdose (OD) deaths quadrupling from 1999 to 2019 [1]. Almost 50,000 OD-related deaths occurred in 2019 alone [2], and the estimated economic burden including opioid use disorder and fatal OD totaled US \$1021 billion during 2017 [3]. The sharp rise in opioid fatalit is responsible for a decline in the US life expectancy [4] and a surge in “deaths of despair” [5]. The opioid crisis is a complex situation involving a broad range of contributing factors including social determinants of health (SDOH) [6,7].

SDOH are the conditions in which people are born, live, work, and age [8]. Adverse SDOH can affect health through various means. For example, social or familial disruptions are well-known precipitants of suicide attempt [9-11]. Behavioral determinants include alcohol consumption, tobacco usage, and use of illicit drugs, among others. Together, adverse social and behavioral determinants of health (SBDH) can be defined as those variables that can hinder an individual’s disease management and negatively impact existing medical conditions [12]. Multiple prior studies suggested strong correlations between OD and a number of SBDH [6,7,13]. Analyzing SBDH in relation to OD can help us better address the OD crisis.

Prior studies found that lack of SBDH information can significantly decrease health care quality [14,15]. Realizing the impact of SBDH on health outcomes, many prior studies focused on extracting SBDH from structured data (eg, diagnosis codes, medications) and/or unstructured data (eg, discharge summaries, progress notes) [11,12,16-18]. However, existing electronic health records (EHRs) often lack the necessary SBDH information in a structured format, undermining its use in clinical care and research settings. On the other hand, EHR notes often describe SBDH [19], for example, financial insecurity (eg, “\$807 SSI and \$16/month food stamps”) and risky alcohol consumption (eg, “Drinking >4 drinks on one occasion or >14 drinks per week”). In addition, EHR notes describe change of status (eg, “recently lost job” or “recently purchased a gun”) that may more precisely identify the current state of a patient. As a consequence, we can take advantage of the rich information provided by unstructured EHR notes via natural language processing (NLP) [20]. NLP has already been

successfully utilized for essential information extraction from EHR text to examine various clinical problems, including opioid use and risk assessment [21,22].

With nonfatal ODs increasing, there is a growing need for critical care of these patients in the United States [23]. Although a relatively high proportion of nonfatal OD cases leads to intensive care unit (ICU) admission, little is known about the risk factors of OD for ICU admissions. [24]. This is essential to understand the severity of the opioid epidemic and anticipate critical care needs for patients with OD. There has been inadequate work on assessing risk factors associated with OD leading to ICU admission, which may be important in comprehensively preventing the public health problem of ODs.

In this study, we specifically focused on the ICU setting to address the aforementioned issues. To mitigate the scarcity of structured SBDH information, we used an NLP system to automatically extract SBDH information from EHR notes and integrated that with available structured SBDH data entered upon admission. Then, we investigated the associations of various demographic, SBDH, and clinical variables with nonfatal OD for eligible ICU admissions. To date, none of the studies on OD utilized the EHR text for extracting SBDH information and few focused on the ICU setting. We bridge this gap by (1) showing that NLP systems can help extract SBDH information when structured data are inadequate and (2) identifying the risk factors that are crucial to the characterization of nonfatal OD leading to ICU admission.

## Methods

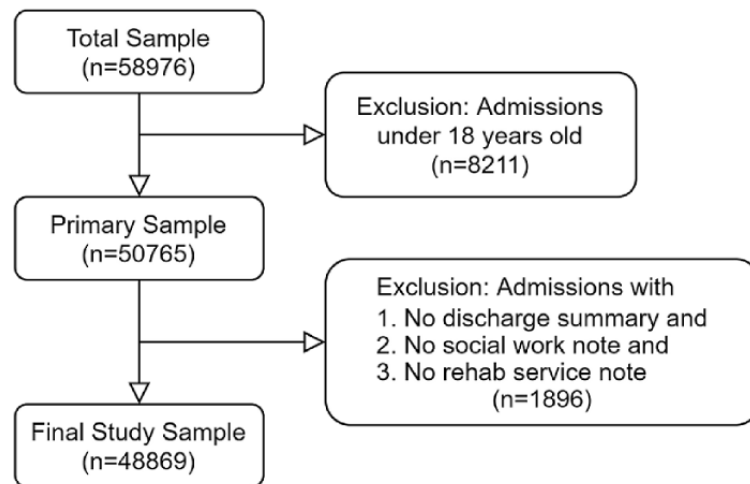
### Dataset

Our primary data source is MIMIC-III [25], one of the largest publicly available ICU databases encompassing 12 years of data (2001-2012) from Beth Israel Deaconess Medical Center. First, we excluded admission data from patients who were less than 18 years old at the time of admission. For inclusion, admissions were also required to have at least one note from any of these 3 categories: discharge summary, social work note, or rehabilitation service note. We selected these 3 types of notes to maximize the use of social and behavioral information for SBDH extraction: Discharge summaries are a comprehensive summary of a patient’s hospital stay, social work notes focus specifically on the social nature of a patient’s life, and

rehabilitation service notes focus on improving patients' function and mobility to stabilize them for discharge. The final sample consisted of 48,869 hospital admissions from 37,361 patients.

An overview of the data selection process is shown in [Figure 1](#).

**Figure 1.** Data selection process.



## Variables

All baseline variables were grouped into 3 categories: demographic, clinical, and SBDH. The demographic variables included age (18-39 years, 40-64 years, >64 years), gender (male or female), race/ethnicity (White, Black, Hispanic, or others), and marital status (married, divorced, widowed, single, or unknown marital status). As clinical variables, we considered drug use disorder, bipolar disorder, tobacco use disorder, major depressive disorder, alcohol use disorder, cirrhosis, chronic obstructive pulmonary disease (COPD), and renal insufficiency. This comprehensive list was made based on earlier studies related to OD [26-29], clinical judgment, and statistical analyses (see the “Statistical Analysis” section for further details). All clinical variables were detected using the International Classification of Disease, 9th Revision (ICD-9) codes from the admission diagnosis chart and included as dichotomous variables. The list of ICD-9 codes is available in [Multimedia Appendix 1](#).

For SBDH variables, we used NLP to analyze the unstructured text data available in MIMIC-III. For each type of note, we chose the most relevant sections to extract the SBDH information: (1) discharge summaries: “Social History” sections; (2) social worker notes: “Patient/Family Assessment,” “Past Addictions History,” “Past Medical History” sections; (3) rehabilitation services: “Sexual and Social History” section.

We used the popular clinical NLP tool medSpaCy [30] to extract these sections from a note. We randomly chose a note, extracted the relevant sections as mentioned, and annotated for 6 categories of SBDH information. This process was repeatedly followed until we reached 1000 notes with at least one SBDH annotation. This annotated subset was later used to train a Bidirectional Encoder Representations from Transformers (BERT) model to extract SBDH at the word level. BERT [31] is a state-of-the-art language representation model that has

successfully outperformed many other NLP systems across a wide range of tasks. We used the trained model to predict the SBDH information for the remaining notes. For an admission with multiple notes of the same type, we took the last note as representative of that admission as it typically includes the content of all the previous notes.

The 6 SBDH variables we chose were (1) housing insecurity, (2) unemployment, (3) social isolation, (4) alcohol use, (5) tobacco use, and (6) illicit drug use. The first 3 are social determinants and were selected based on the list of well-accepted social determinants provided by the Kaiser Family Foundation [32]. The rest were substance use-related health risk behaviors (ie, behavioral determinants) that were chosen for their clinical significance and relevance to OD. Details about the annotation process, NLP model development, and SBDH variable extraction procedures are provided in [Multimedia Appendix 2](#).

In addition to the NLP-derived SBDH variables, we also identified social determinants from the structured data. We used the ICD-9 codes from patient diagnoses [33] to construct these 3 SBDH variables: (1) housing insecurity, (2) unemployment, and (3) social isolation. These were later integrated with the NLP-derived SBDH variables and prioritized in case of any mismatch. For example, if the NLP system detected “housing insecurity” as “No” for an admission and we obtained “Yes” from that admission’s diagnoses codes, we considered “Yes” as the correct value. In the end, there were 41,669 admissions (41,669/48,869, 85.27%) with at least one SBDH variable. [Table 1](#) illustrates the 6 SBDH variables with brief descriptions and examples. If an admission had no mention of SBDH information, SBDH variables were coded as “unknown.” For instance, if an admission had no mention of patient housing status in the corresponding notes, homelessness was considered “unknown.” Other than these 3 SBDH variables, we also extracted insurance provider (private, Medicaid, Medicare, other government, or self-pay) information using ICD-9 codes.

**Table 1.** Descriptions and examples of social and behavioral determinants of health (SBDH) variables.

SBDH Variable	Description and example
<b>Housing insecurity</b>	
Yes	Lack housing or stable shelter. Example: <i>homeless</i> , living with friends.
No	Has access to housing. Example: <i>lives in [**location**]</i> by himself.
<b>Unemployment</b>	
Yes	Patient has no source of income or lost job. Example: Patient used to work for the state lottery system, currently <i>unemployed</i> .
No	Patient has employment or some source of income. Example: He works for [**Company**].
<b>Social isolation</b>	
Yes	Lack of social support or community engagement. Example: Lives <i>alone</i> in [**Location**].
No	Presence of social support. Example: He is <i>married</i> and <i>lives</i> with his wife.
<b>Alcohol use</b>	
Current	Patient currently consumes alcohol. Example: two glasses of <i>wine</i> per night and 3 bottles over the weekend.
Former	Patient has a history of alcohol consumption. Example: He has a history of <i>alcohol</i> abuse.
None	Patient never consumed alcohol. Example: She denies any <i>alcohol use</i> .
<b>Smoking</b>	
Current	Patient currently smokes. Example: He <i>smokes one pack of cigarettes</i> per week.
Former	Patient has a history of tobacco usage. Example: The patient has a past history of <i>smoking</i> .
None	Patient never consumed tobacco. Example: She is a <i>nonsmoker</i> .
<b>Illicit drug use</b>	
Current	Patient uses non-prescribed controlled substance. Example: occasional <i>marijuana</i> use.
Former	Patient has a history of using non-prescribed controlled substance. Example: Has a h/o <sup>a</sup> of <i>cocaine</i> and <i>marijuana</i> abuse.
None	Patient never used non-prescribed controlled substance, e.g., cocaine, marijuana. Example: Does not drink alcohol or use recreational <i>drugs</i> .

<sup>a</sup>h/o: history of.

## Outcome

The outcome was nonfatal OD, which was identified using ICD-9 codes [34].

## Statistical Analysis

First, we performed correlation and collinearity analyses for all the variables. The correlation plot and variance inflation factor [35] did not show multicollinearity among the variables. For the clinical variables, based on earlier work and task relevance, we chose 14 comorbidities: posttraumatic stress disorder, major depressive disorder, bipolar disorder, schizophrenia, alcohol use disorder, drug use disorder, tobacco use disorder, hepatitis C, diabetes, congestive heart failure, obstructive sleep apnea, COPD, cirrhosis, and renal insufficiency. We built logistic regression models and employed the sequential forward selection procedure [36] to identify the most essential clinical variables related to OD. The final list included 8 clinical variables: drug use disorder, bipolar disorder, tobacco use disorder, major depressive disorder, alcohol use disorder, cirrhosis, COPD, and renal insufficiency.

We used a logistic regression model to examine the associations of nonfatal OD with demographic, SBDH, and clinical variables.

This was assessed in terms of adjusted odds ratios (aOR) with 95% CIs. We also evaluated the crude odds ratio (OR) with 95% CIs. The statistical significance was measured at  $P < .05$ . Hosmer-Lemeshow test was conducted and indicated a sufficient fit for our model ( $\chi_8 = 10.39$ ;  $P = .24$ ). All statistical analyses in this study were conducted in R (version 4.0.2).

## Results

### Descriptive Analysis

Table 2 presents the characteristics of our cohort ( $n = 48,869$ ). Our sample was comprised of mostly men (27,436/48,869, 56.14%) and white (35,058/48,869, 71.74%) adults. The majority of patients were aged 64 years or older (25,276/48,869, 51.72%). Of the clinical variables, renal insufficiency was the most prevalent (8158/48,869, 16.69%), followed by COPD (5674/48,869, 11.61%) and alcohol use disorder (4121/48,869, 8.43%). In our cohort, we observed that 7.28% (3559/48,869) of the patients were unemployed, 13.35% (6523/48,869) were socially isolated, and 0.82% (402/48,869) had housing insecurity. We found 171 (171/48,869, 0.35%) admissions with nonfatal OD.

**Table 2.** Prevalence of demographic, clinical, and social and behavioral determinants of health (SBDH) variables in MIMIC-III.

Variables	Overall (n=48,869)	With OD <sup>a</sup> (n=171)	Without OD (n=48,698)
<b>Age<sup>b</sup> (years), n (%)</b>			
<40	4715 (9.65)	62 (36.26)	4653 (9.55)
40-64	18,878 (38.63)	92 (53.80)	18,786 (38.58)
>64	25,276 (51.72)	17 (9.94)	25,259 (51.87)
<b>Gender,<sup>b</sup> n (%)</b>			
Male	27,436 (56.14)	100 (58.48)	27,336 (56.13)
Female	21,433 (43.86)	71 (41.52)	21,362 (43.87)
<b>Race/ethnicity,<sup>b</sup> n (%)</b>			
White	35,058 (71.74)	127 (74.27)	34,931 (71.73)
Black	4694 (9.61)	13 (7.60)	4681 (9.61)
Hispanic	1664 (3.40)	8 (4.68)	1656 (3.40)
Other	7453 (15.25)	23 (13.45)	7430 (15.26)
<b>Marital status,<sup>b</sup> n (%)</b>			
Married	23,378 (47.84)	42 (24.56)	23,336 (47.92)
Divorced	3664 (7.50)	22 (12.87)	3642 (7.48)
Widowed	7018 (14.36)	6 (3.51)	7012 (14.40)
Single	12,329 (25.23)	78 (45.61)	12,251 (25.16)
Unknown	2480 (5.07)	23 (13.45)	2457 (5.04)
<b>Clinical variables,<sup>b</sup> n (%)</b>			
Drug use disorder	1493 (3.06)	80 (46.78)	1413 (2.90)
Bipolar disorder	1009 (2.06)	28 (16.37)	981 (2.01)
Tobacco use disorder	3274 (6.70)	20 (11.70)	3254 (6.68)
Major depressive disorder	298 (0.61)	7 (4.09)	291 (0.60)
Alcohol use disorder	4121 (8.43)	37 (21.64)	4084 (8.39)
Cirrhosis	2431 (4.97)	19 (11.11)	2412 (4.95)
COPD <sup>c</sup>	5674 (11.61)	18 (10.53)	5656 (11.61)
Renal insufficiency	8158 (16.69)	12 (7.02)	8146 (16.73)
<b>Social determinant<sup>d</sup>: insurance provider, n (%)</b>			
Private	15,371 (31.45)	43 (25.15)	15,328 (31.48)
Medicaid	4307 (8.81)	60 (35.09)	4247 (8.72)
Medicare	27,365 (56.00)	48 (28.07)	27,317 (56.09)
Government (others)	1324 (2.71)	14 (8.19)	1310 (2.69)
Self-pay	502 (1.03)	6 (3.50)	496 (1.02)
<b>Social determinant<sup>d</sup>: housing insecurity, n (%)</b>			
Yes	402 (0.82)	10 (5.85)	392 (0.80)
No	27,119 (55.49)	92 (53.80)	27,027 (55.50)
Unknown	21,348 (43.69)	69 (40.35)	21,279 (43.70)
<b>Social determinant<sup>d</sup>: unemployment, n (%)</b>			
Yes	3559 (7.28)	37 (21.64)	3522 (7.22)
No	12,671 (25.93)	31 (18.13)	12,640 (25.96)

Variables	Overall (n=48,869)	With OD <sup>a</sup> (n=171)	Without OD (n=48,698)
Unknown	32,639 (66.79)	103 (60.23)	32,536 (66.82)
<b>Social determinant<sup>d</sup>: social isolation, n (%)</b>			
Yes	6523 (13.35)	23 (13.45)	6500 (13.35)
No	24,001 (49.11)	86 (50.29)	23,915 (49.11)
Unknown	18,345 (37.54)	62 (36.26)	18,283 (37.54)
<b>Substance use<sup>e</sup>: alcohol use, n (%)</b>			
Current	14,150 (28.96)	70 (40.94)	14,080 (28.91)
Former	2333 (4.77)	9 (5.26)	2324 (4.77)
None	15,378 (31.47)	40 (23.39)	15,338 (31.50)
Unknown	17,008 (34.80)	52 (30.41)	16,956 (34.82)
<b>Substance use<sup>e</sup>: smoking, n (%)</b>			
Current	6954 (14.23)	62 (36.26)	6892 (14.15)
Former	12,032 (24.62)	23 (13.45)	12,009 (24.66)
None	13,963 (28.57)	30 (17.54)	13,933 (28.61)
Unknown	15,920 (32.58)	56 (32.75)	15,864 (32.58)
<b>Substance use<sup>e</sup>: illicit drug use, n (%)</b>			
Current	1796 (3.67)	49 (28.65)	1747 (3.59)
Former	1362 (2.79)	26 (15.20)	1336 (2.74)
None	13,908 (28.46)	31 (18.13)	13,877 (28.50)
Unknown	31,803 (65.08)	65 (38.02)	31,738 (65.17)

<sup>a</sup>OD: opioid overdose.

<sup>b</sup>Variables extracted from structured data.

<sup>c</sup>COPD: chronic obstructive pulmonary disease.

<sup>d</sup>Variables extracted from only structured data (insurance provider) or both structured data and unstructured text notes (natural language processing [NLP]).

<sup>e</sup>Variables extracted from unstructured text notes (NLP).

Of the 6 NLP-derived SBDH variables, only housing insecurity, unemployment, and social isolation had associated ICD-9 diagnostic codes. Compared with their NLP-derived counterparts, these structured variables were coded infrequently. For example, using ICD-9 codes, we found 258 admissions with “housing insecurity,” whereas the NLP system detected 402 admissions. For “unemployment,” it was 20 for the ICD-9 codes and 10,876 for the NLP system. And more striking, for “social isolation,” only 4 admissions had relevant ICD-9 codes in their diagnosis compared to 6523 admissions found by the NLP system. Due to the substantial prevalence gap, we did not compare the quality of these 2 types of SBDH variables side by side. In all, structured SBDH variables accounted for only 0.18% of the SBDH variables. This clearly shows that NLP can be useful to extract SBDH information from EHR notes when structured data are not enough. This also helps reduce bias from the use of structured data only.

### Multivariable Logistic Regression Analysis

Several factors were strongly associated with nonfatal OD (Table 3). Among the demographic risk factors, Blacks (aOR 0.51, 95% CI 0.28-0.94) and older age groups (40-64 years: aOR 0.65, 95% CI 0.44-0.96; >64 years: aOR 0.16, 95% CI 0.08-0.34) had lower odds compared with White and younger patients. Among the 8 clinical variables, 5 were strong risk factors for nonfatal OD. We observed increased odds of overdose among individuals with drug use disorder (aOR 8.17, 95% CI 5.44-12.27), bipolar disorder (aOR 2.69, 95% CI 1.68-4.29), and major depressive disorder (aOR 2.57, 95% CI 1.12-5.88). Interestingly, tobacco use disorder (aOR 0.53, 95% CI 0.32-0.89) and alcohol use disorder (aOR 0.64, 95% CI 0.42-1.00) had decreased odds. Among the SBDH variables, individuals with Medicaid had increased odds compared with those with private medical insurance (aOR 2.26, 95% CI 1.43-3.58). History of (aOR 2.09, 95% CI 1.15-3.79) and current (aOR 2.06, 95% CI 1.20-3.55) use of illicit drugs were also strongly associated with the outcome.

**Table 3.** Multivariable logistic regression analysis for the factors associated with nonfatal opioid overdose (OD).

Variables	Crude OR <sup>a</sup>	95% CI	aOR <sup>b</sup>	95% CI
<b>Age (years)</b>				
<40	Ref <sup>c</sup>	Ref	Ref	Ref
40-64	0.37	0.27-0.51	0.65	0.44-0.96
>64	0.05	0.03-0.08	0.16	0.08-0.34
<b>Gender</b>				
Male	Ref	Ref	Ref	Ref
Female	0.91	0.67-1.23	1.13	0.81-1.58
<b>Race/ethnicity</b>				
White	Ref	Ref	Ref	Ref
Black	0.76	0.41-1.30	0.51	0.28-0.94
Hispanic	1.33	0.60-2.55	0.69	0.33-1.45
Others	0.85	0.53-1.30	0.59	0.35-0.98
<b>Marital status</b>				
Married	Ref	Ref	Ref	Ref
Divorced	3.36	1.97-5.57	1.56	0.89-2.74
Widowed	0.48	0.18-1.04	0.76	0.30-1.88
Single	3.54	2.44-5.19	1.03	0.65-1.61
Unknown	5.20	3.08-8.58	2.85	1.55-5.24
<b>Clinical variables</b>				
Drug use disorder	29.42	21.65-39.90	8.17	5.44-12.27
Bipolar disorder	9.52	6.20-14.12	2.69	1.68-4.29
Tobacco use disorder	1.85	1.12-2.88	0.53	0.32-0.89
Major depressive disorder	7.10	2.99-14.16	2.57	1.12-5.88
Alcohol use disorder	3.02	2.06-4.30	0.64	0.42-1.00
Cirrhosis	2.40	1.44-3.77	1.65	0.97-2.82
COPD <sup>d</sup>	7.10	2.99-14.16	1.65	0.97-2.81
Renal insufficiency	3.02	2.06-4.30	0.62	0.33-1.15
<b>Social determinant: insurance type</b>				
Private	Ref	Ref	Ref	Ref
Medicaid	5.04	3.41-7.50	2.26	1.43-3.58
Medicare	0.63	0.41-0.95	1.34	0.81-2.23
Government (others)	3.81	2.01-6.80	1.90	0.99-3.65
Self-paid	4.31	1.64-9.42	1.83	0.73-4.56
<b>Social determinant: housing insecurity</b>				
No	Ref	Ref	Ref	Ref
Yes	7.49	3.63-13.80	0.98	0.47-2.06
Unknown	0.95	0.69-1.30	0.89	0.60-1.33
<b>Social determinant: unemployment</b>				
No	Ref	Ref	Ref	Ref
Yes	4.28	2.66-6.95	1.10	0.65-1.87
Unknown	1.29	0.87-1.96	0.73	0.47-1.14

Variables	Crude OR <sup>a</sup>	95% CI	aOR <sup>b</sup>	95% CI
<b>Social determinant: social isolation</b>				
No	Ref	Ref	Ref	Ref
Yes	0.98	0.61-1.53	0.97	0.59-1.60
Unknown	0.94	0.68-1.31	1.01	0.66-1.53
<b>Substance use: alcohol use</b>				
None	Ref	Ref	Ref	Ref
Former	1.48	0.67-2.92	0.66	0.30-1.44
Current	1.91	1.30-2.84	1.11	0.71-1.72
Unknown	1.18	0.78-1.79	1.05	0.62-1.78
<b>Substance use: smoking</b>				
None	Ref	Ref	Ref	Ref
Former	0.89	0.51-1.53	0.92	0.52-1.65
Current	4.18	2.72-6.55	1.40	0.84-2.33
Unknown	1.64	1.06-2.59	1.12	0.64-1.96
<b>Substance use: illicit drug use</b>				
None	Ref	Ref	Ref	Ref
Former	8.71	5.12-14.7	2.09	1.15-3.79
Current	12.56	8.03-19.93	2.06	1.20-3.55
Unknown	0.92	0.60-1.42	1.05	0.65-1.70

<sup>a</sup>OR: odds ratio.

<sup>b</sup>aOR: adjusted odds ratio.

<sup>c</sup>Ref: Reference.

<sup>d</sup>COPD: chronic obstructive pulmonary disorder.

## Discussion

### Principal Findings

To our knowledge, this is the first study to examine the risk factors associated with nonfatal OD leading to ICU admission. In the United States, the need for characterizing critical care patients with OD is rising [23,24], and this study partially addressed that by identifying the risk factors for nonfatal OD from a large ICU database. The novelty also lies in the use of a state-of-the-art NLP system that utilized unstructured EHR notes for essential SBDH extraction due to inadequate representation from structured data. There is a growing body of literature showing that SBDH can strongly influence patient health and outcomes [12]. For example, SBDH variables have been shown to be strongly associated with suicide attempt [11], mortality [17], and mental health diagnosis [18]. The challenges here for the health care systems are to set up methods that can identify SBDH and use them at the point of care to inform clinical action [37,38]. Our work demonstrated that using NLP to detect SBDH information from EHR text can be a viable option in this regard.

According to our analysis, multiple SBDH variables were significantly associated with nonfatal OD in ICU settings. We observed that patients with economic instability (unemployed) were more likely to have an overdose, but homelessness and

social isolation conferred little additional risk. Among behavioral determinants, *current* alcohol users and smokers had higher odds of overdose, whereas *former* users had decreased odds. Illicit drug use was strongly associated with nonfatal OD for both *former* and *current* users. Among clinical variables, tobacco use disorder and alcohol use disorder had strong negative associations with nonfatal OD. We hypothesize that the majority of the patients diagnosed with such disorders were already receiving additional social counseling or clinical support, which helped them build better health and behavioral practices. However, we did not have enough relevant admission data in MIMIC-III to validate this hypothesis; future research is needed to identify the reasons for this observation.

### Limitations

There are several limitations of our study. EHR data are prone to variability by provider documentation and may contain incomplete SBDH information [39]. Additionally, using only ICD-9 codes to identify different medical conditions may lead to inaccurate or misleading values for the corresponding variable. However, structured data often significantly lack SBDH information (only 0.18% for this study), making an NLP-based approach a valuable integration for population studies. Finally, our data had a very low prevalence of nonfatal OD cases (171/48,869, 0.35%), and the MIMIC (ICU) database might not characterize the general outpatient/inpatient hospital setting.



While our study describes an important methodological process that can identify important SBDH factors to consider, which is a necessary first step, further research is needed on subsequent steps on how best to share and translate this information to providers so that they can effectively and actionably use the findings. As our future work, we would like to work on modeling the NLP system predictions for SBDH extraction and how they can be better tied with predictor assessment metrics (eg, OR).

## Conclusions

This is the first work to evaluate the risk factors associated with nonfatal OD leading to ICU admissions. Our work concluded that data-driven NLP systems can be largely beneficial in the automatic extraction of SBDH information from unstructured EHR text data. We also showed that analyzing critical care admissions is crucial to better understand the opioid epidemic. Utilizing NLP to leverage the rich EHR notes and more epidemiological studies in critical care settings could be useful for deeper analysis of the OD crisis, leading to the development of better risk assessment tools and effective prevention systems.

## Acknowledgments

We thank Minhee Sung, Jimin Kim, and Chen Kun for their valuable comments. This work was supported in part by the grant R01DA045816 from the National Institutes of Health (NIH). The contents of this paper do not represent the views of the NIH.

## Conflicts of Interest

None declared.

## Multimedia Appendix 1

International Classification of Disease, 9th Revision codes for clinical variables.

[\[DOCX File , 36 KB-Multimedia Appendix 1\]](#)

## Multimedia Appendix 2

Natural language processing model training and evaluation.

[\[DOCX File , 23 KB-Multimedia Appendix 2\]](#)

## References

1. Overdose Death Rates. National Institute on Drug Abuse (NIDA). 2017. URL: <https://www.drugabuse.gov/drug-topics/trends-statistics/overdose-death-rates> [accessed 2021-10-16]
2. Opioid Overdose Crisis. National Institute on Drug Abuse (NIDA). URL: <https://www.drugabuse.gov/drug-topics/opioids/opioid-overdose-crisis> [accessed 2021-10-16]
3. Luo F, Li M, Florence C. State-level economic costs of opioid use disorder and fatal opioid overdose - United States, 2017. *MMWR Morb Mortal Wkly Rep* 2021 Apr 16;70(15):541-546 [FREE Full text] [doi: [10.15585/mmwr.mm7015a1](https://doi.org/10.15585/mmwr.mm7015a1)] [Medline: [33857070](https://pubmed.ncbi.nlm.nih.gov/33857070/)]
4. Dowell D, Arias E, Kochanek K, Anderson R, Guy GP, Losby JL, et al. Contribution of opioid-involved poisoning to the change in life expectancy in the United States, 2000-2015. *JAMA* 2017 Sep 19;318(11):1065-1067 [FREE Full text] [doi: [10.1001/jama.2017.9308](https://doi.org/10.1001/jama.2017.9308)] [Medline: [28975295](https://pubmed.ncbi.nlm.nih.gov/28975295/)]
5. Case A, Deaton A. Mortality and morbidity in the 21 century. *Brookings Pap Econ Act* 2017;2017:397-476 [FREE Full text] [doi: [10.1353/eca.2017.0005](https://doi.org/10.1353/eca.2017.0005)] [Medline: [29033460](https://pubmed.ncbi.nlm.nih.gov/29033460/)]
6. Dasgupta N, Beletsky L, Ciccarone D. Opioid crisis: no easy fix to its social and economic determinants. *Am J Public Health* 2018 Feb;108(2):182-186. [doi: [10.2105/ajph.2017.304187](https://doi.org/10.2105/ajph.2017.304187)]
7. Volkow ND, Blanco C. The changing opioid crisis: development, challenges and opportunities. *Mol Psychiatry* 2021 Jan;26(1):218-233 [FREE Full text] [doi: [10.1038/s41380-020-0661-4](https://doi.org/10.1038/s41380-020-0661-4)] [Medline: [32020048](https://pubmed.ncbi.nlm.nih.gov/32020048/)]
8. Cole BL, Fielding JE. Health impact assessment: a tool to help policy makers understand health beyond health care. *Annu Rev Public Health* 2007;28:393-412. [doi: [10.1146/annurev.publhealth.28.083006.131942](https://doi.org/10.1146/annurev.publhealth.28.083006.131942)] [Medline: [17173539](https://pubmed.ncbi.nlm.nih.gov/17173539/)]
9. Kposowa AJ. Unemployment and suicide: a cohort analysis of social factors predicting suicide in the US National Longitudinal Mortality Study. *Psychol Med* 2001 Jan;31(1):127-138. [doi: [10.1017/s0033291799002925](https://doi.org/10.1017/s0033291799002925)] [Medline: [11200951](https://pubmed.ncbi.nlm.nih.gov/11200951/)]
10. Dube SR, Anda RF, Felitti VJ, Chapman DP, Williamson DF, Giles WH. Childhood abuse, household dysfunction, and the risk of attempted suicide throughout the life span: findings from the Adverse Childhood Experiences Study. *JAMA* 2001 Dec 26;286(24):3089-3096. [doi: [10.1001/jama.286.24.3089](https://doi.org/10.1001/jama.286.24.3089)] [Medline: [11754674](https://pubmed.ncbi.nlm.nih.gov/11754674/)]
11. Blossnich JR, Montgomery AE, Dichter ME, Gordon AJ, Kavalieratos D, Taylor L, 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 Jun;35(6):1759-1767 [FREE Full text] [doi: [10.1007/s11606-019-05447-z](https://doi.org/10.1007/s11606-019-05447-z)] [Medline: [31745856](https://pubmed.ncbi.nlm.nih.gov/31745856/)]

12. Feller DJ, Bear Don't Walk Iv OJ, Zucker J, Yin MT, Gordon P, Elhadad N. Detecting social and behavioral determinants of health with structured and free-text clinical data. *Appl Clin Inform* 2020 Jan 04;11(1):172-181 [FREE Full text] [doi: [10.1055/s-0040-1702214](https://doi.org/10.1055/s-0040-1702214)] [Medline: [32131117](https://pubmed.ncbi.nlm.nih.gov/32131117/)]
13. Cantu R, Fields-Johnson D, Savannah S. Applying a social determinants of health approach to the opioid epidemic. *Health Promot Pract* 2020 Jul 26;1524839920943207. [doi: [10.1177/1524839920943207](https://doi.org/10.1177/1524839920943207)] [Medline: [32713219](https://pubmed.ncbi.nlm.nih.gov/32713219/)]
14. Gottlieb LM, Tirozzi KJ, Manchanda R, Burns AR, Sandel MT. Moving electronic medical records upstream: incorporating social determinants of health. *Am J Prev Med* 2015 Feb;48(2):215-218. [doi: [10.1016/j.amepre.2014.07.009](https://doi.org/10.1016/j.amepre.2014.07.009)] [Medline: [25217095](https://pubmed.ncbi.nlm.nih.gov/25217095/)]
15. Weir CR, Staggers N, Gibson B, Doing-Harris K, Barrus R, Dunlea R. A qualitative evaluation of the crucial attributes of contextual information necessary in EHR design to support patient-centered medical home care. *BMC Med Inform Decis Mak* 2015 Apr 16;15:30 [FREE Full text] [doi: [10.1186/s12911-015-0150-x](https://doi.org/10.1186/s12911-015-0150-x)] [Medline: [25881181](https://pubmed.ncbi.nlm.nih.gov/25881181/)]
16. Gottlieb L, Sandel M, Adler NE. Collecting and applying data on social determinants of health in health care settings. *JAMA Intern Med* 2013 Jun 10;173(11):1017-1020. [doi: [10.1001/jamainternmed.2013.560](https://doi.org/10.1001/jamainternmed.2013.560)] [Medline: [23699778](https://pubmed.ncbi.nlm.nih.gov/23699778/)]
17. Blossnich JR, Montgomery AE, Taylor LD, Dichter ME. Adverse social factors and all-cause mortality among male and female patients receiving care in the Veterans Health Administration. *Prev Med* 2020 Dec;141:106272. [doi: [10.1016/j.yjmed.2020.106272](https://doi.org/10.1016/j.yjmed.2020.106272)] [Medline: [33022319](https://pubmed.ncbi.nlm.nih.gov/33022319/)]
18. Blossnich JR, Marsiglio MC, Dichter ME, Gao S, Gordon AJ, Shipherd JC, et al. Impact of social determinants of health on medical conditions among transgender veterans. *Am J Prev Med* 2017 Apr;52(4):491-498 [FREE Full text] [doi: [10.1016/j.amepre.2016.12.019](https://doi.org/10.1016/j.amepre.2016.12.019)] [Medline: [28161034](https://pubmed.ncbi.nlm.nih.gov/28161034/)]
19. Dorr D, Bejan CA, Pizzimenti C, Singh S, Storer M, Quinones A. Identifying patients with significant problems related to social determinants of health with natural language processing. *Stud Health Technol Inform* 2019 Aug 21;264:1456-1457. [doi: [10.3233/SHTI190482](https://doi.org/10.3233/SHTI190482)] [Medline: [31438179](https://pubmed.ncbi.nlm.nih.gov/31438179/)]
20. Liu F, Weng C, Yu H. Natural language processing, electronic health records, and clinical research. In: Richesson R, Andrews JE, editors. *Clinical Research Informatics*. New York City, NY: Springer Publishing Company; 2012:293-310.
21. Carrell DS, Cronkite D, Palmer RE, Saunders K, Gross DE, Masters ET, et al. Using natural language processing to identify problem usage of prescription opioids. *Int J Med Inform* 2015 Dec;84(12):1057-1064. [doi: [10.1016/j.ijmedinf.2015.09.002](https://doi.org/10.1016/j.ijmedinf.2015.09.002)] [Medline: [26456569](https://pubmed.ncbi.nlm.nih.gov/26456569/)]
22. Haller IV, Renier CM, Hitz P, Palcher JA, Elliott TE. Validation of the automated diagnosis, intractability, risk, efficacy (DIRE) opioid risk assessment tool. *Journal of Patient-Centered Research and Reviews* 2016 Aug 15;3(3):227. [doi: [10.17294/2330-0698.1403](https://doi.org/10.17294/2330-0698.1403)]
23. Stevens JP, Wall MJ, Novack L, Marshall J, Hsu DJ, Howell MD. The critical care crisis of opioid overdoses in the United States. *Annals ATS* 2017 Dec;14(12):1803-1809. [doi: [10.1513/annalsats.201701-022oc](https://doi.org/10.1513/annalsats.201701-022oc)]
24. Pfister GJ, Burkes RM, Guinn B, Steele J, Kelley RR, Wiemken TL, et al. Opioid overdose leading to intensive care unit admission: Epidemiology and outcomes. *J Crit Care* 2016 Oct;35:29-32. [doi: [10.1016/j.jcrc.2016.04.022](https://doi.org/10.1016/j.jcrc.2016.04.022)] [Medline: [27481733](https://pubmed.ncbi.nlm.nih.gov/27481733/)]
25. Johnson AEW, Pollard TJ, Shen L, Lehman LWH, Feng M, Ghassemi M, et al. MIMIC-III, a freely accessible critical care database. *Sci Data* 2016 May 24;3(1):160035 [FREE Full text] [doi: [10.1038/sdata.2016.35](https://doi.org/10.1038/sdata.2016.35)] [Medline: [27219127](https://pubmed.ncbi.nlm.nih.gov/27219127/)]
26. Bohnert ASB, Valenstein M, Bair MJ, Ganoczy D, McCarthy JF, Ilgen MA, et al. Association between opioid prescribing patterns and opioid overdose-related deaths. *JAMA* 2011 Apr 06;305(13):1315-1321. [doi: [10.1001/jama.2011.370](https://doi.org/10.1001/jama.2011.370)] [Medline: [21467284](https://pubmed.ncbi.nlm.nih.gov/21467284/)]
27. Dunn KM, Saunders KW, Rutter CM, Banta-Green CJ, Merrill JO, Sullivan MD, et al. Opioid prescriptions for chronic pain and overdose: a cohort study. *Ann Intern Med* 2010 Jan 19;152(2):85-92 [FREE Full text] [doi: [10.7326/0003-4819-152-2-201001190-00006](https://doi.org/10.7326/0003-4819-152-2-201001190-00006)] [Medline: [20083827](https://pubmed.ncbi.nlm.nih.gov/20083827/)]
28. Campbell CI, Bahorik AL, VanVeldhuisen P, Weisner C, Rubinstein AL, Ray GT. Use of a prescription opioid registry to examine opioid misuse and overdose in an integrated health system. *Prev Med* 2018 May;110:31-37 [FREE Full text] [doi: [10.1016/j.yjmed.2018.01.019](https://doi.org/10.1016/j.yjmed.2018.01.019)] [Medline: [29410132](https://pubmed.ncbi.nlm.nih.gov/29410132/)]
29. Glanz JM, Narwaney KJ, Mueller SR, Gardner EM, Calcaterra SL, Xu S, et al. Prediction model for two-year risk of opioid overdose among patients prescribed chronic opioid therapy. *J Gen Intern Med* 2018 Oct 29;33(10):1646-1653 [FREE Full text] [doi: [10.1007/s11606-017-4288-3](https://doi.org/10.1007/s11606-017-4288-3)] [Medline: [29380216](https://pubmed.ncbi.nlm.nih.gov/29380216/)]
30. medspacy. GitHub. URL: <https://github.com/medspacy/medspacy> [accessed 2021-10-16]
31. Devlin J, Chang MW, Lee K, Toutanova K. BERT: Pre-training of deep bidirectional transformers for language understanding. *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)* 2019:4171-4186 [FREE Full text] [doi: [10.18653/v1/N19-1423](https://doi.org/10.18653/v1/N19-1423)]
32. Heiman HJ, Artiga S. Beyond Health Care: The Role of Social Determinants in Promoting Health and Health Equity. Kaiser Family Foundation. 2018 May 10. URL: <https://www.kff.org/racial-equity-and-health-policy/issue-brief/beyond-health-care-the-role-of-social-determinants-in-promoting-health-and-health-equity/> [accessed 2021-10-16]

33. Kessler RC, Bauer MS, Bishop TM, Demler OV, Dobscha SK, Gildea SM, et al. Using administrative data to predict suicide after psychiatric hospitalization in the Veterans Health Administration System. *Front Psychiatry* 2020 May 6;11:390 [FREE Full text] [doi: [10.3389/fpsy.2020.00390](https://doi.org/10.3389/fpsy.2020.00390)] [Medline: [32435212](https://pubmed.ncbi.nlm.nih.gov/32435212/)]
34. Glanz J, Binswanger I, Shetterly SM, Narwaney KJ, Xu S. Association between opioid dose variability and opioid overdose among adults prescribed long-term opioid therapy. *JAMA Netw Open* 2019 Apr 05;2(4):e192613 [FREE Full text] [doi: [10.1001/jamanetworkopen.2019.2613](https://doi.org/10.1001/jamanetworkopen.2019.2613)] [Medline: [31002325](https://pubmed.ncbi.nlm.nih.gov/31002325/)]
35. O'Brien RM. A caution regarding rules of thumb for variance inflation factors. *Qual Quant* 2007 Mar 13;41(5):673-690. [doi: [10.1007/s11135-006-9018-6](https://doi.org/10.1007/s11135-006-9018-6)]
36. Ferri FJ, Pudil P, Hatef M, Kittler J. Comparative study of techniques for large-scale feature selection. *Machine Intelligence and Pattern Recognition* 1994;16:403-413. [doi: [10.1016/B978-0-444-81892-8.50040-7](https://doi.org/10.1016/B978-0-444-81892-8.50040-7)]
37. Schickedanz A, Hamity C, Rogers A, Sharp A, Jackson A. Clinician experiences and attitudes regarding screening for social determinants of health in a large integrated health system. *Med Care* 2019 Jun;57 Suppl 6 Suppl 2:S197-S201 [FREE Full text] [doi: [10.1097/MLR.0000000000001051](https://doi.org/10.1097/MLR.0000000000001051)] [Medline: [31095061](https://pubmed.ncbi.nlm.nih.gov/31095061/)]
38. Horwitz LI, Chang C, Arcilla HN, Knickman JR. Quantifying health systems' investment in social determinants of health, by sector, 2017-19. *Health Aff (Millwood)* 2020 Feb;39(2):192-198. [doi: [10.1377/hlthaff.2019.01246](https://doi.org/10.1377/hlthaff.2019.01246)] [Medline: [32011928](https://pubmed.ncbi.nlm.nih.gov/32011928/)]
39. Weiskopf NG, Bakken S, Hripcsak G, Weng C. A data quality assessment guideline for electronic health record data reuse. *EGEMS (Wash DC)* 2017 Sep 04;5(1):14 [FREE Full text] [doi: [10.5334/egems.218](https://doi.org/10.5334/egems.218)] [Medline: [29881734](https://pubmed.ncbi.nlm.nih.gov/29881734/)]

## Abbreviations

- aOR:** adjusted odds ratio  
**BERT:** Bidirectional Encoder Representations from Transformers  
**COPD:** chronic obstructive pulmonary disease  
**EHR:** Electronic health record  
**ICD-9:** International Classification of Disease, 9th Revision  
**ICU:** Intensive care unit  
**NIH:** National Institutes of Health  
**NLP:** natural language processing  
**OD:** opioid overdose  
**OR:** odds ratio  
**SBDH:** social and behavioral determinants of health  
**SDOH:** social determinants of health

*Edited by G Eysenbach; submitted 12.08.21; peer-reviewed by J Coquet; comments to author 05.09.21; revised version received 23.09.21; accepted 26.09.21; published 08.11.21*

*Please cite as:*

Mitra A, Ahsan H, Li W, Liu W, Kerns RD, Tsai J, Becker W, Smelson DA, Yu H  
*Risk Factors Associated With Nonfatal Opioid Overdose Leading to Intensive Care Unit Admission: A Cross-sectional Study*  
*JMIR Med Inform* 2021;9(11):e32851  
URL: <https://medinform.jmir.org/2021/11/e32851>  
doi: [10.2196/32851](https://doi.org/10.2196/32851)  
PMID:

©Avijit Mitra, Hiba Ahsan, Wenjun Li, Weisong Liu, Robert D Kerns, Jack Tsai, William Becker, David A Smelson, Hong Yu. Originally published in *JMIR Medical Informatics* (<https://medinform.jmir.org/>), 08.11.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR Medical Informatics*, is properly cited. The complete bibliographic information, a link to the original publication on <https://medinform.jmir.org/>, as well as this copyright and license information must be included.