Hill *et al. BMC Public Health* (2025) 25:1453
https://doi.org/10.1186/s12889-025-22490-0

BMC Public Health

RESEARCH

Open Access

The geographic identification of elevated suicide risk model: evaluating a method for examining suicide-related behaviors at the neighborhood level in Harris County, Texas

Ryan M. Hill[1,5*], Sam C. Cacace[2], Austin Starkey[1], Tyler Hendley[1], Jackson Bolner[1], Angela Cummings[3], Angie Hayes[4] and Robert J. Cramer[2]

Abstract

Background The 2024 National Strategy for Suicide Prevention has called for the development of community-based suicide prevention resources, and improved existing prevention efforts. In line with such efforts, Hill and colleagues developed the Geospatial Identification of Elevated Suicide Risk model that estimates the relative prevalence of adolescent suicide risk within specific geographical areas. The current study seeks to further evaluate and refine the model for use as a tool to evaluate risk and protective factors at the neighborhood level.

Method Drawing from multiple sources, data was collected detailing adolescent suicidal ideation, suicide attempts, suicide fatalities, and census tract characteristics. Utilizing data resulting from an initial pool of 74,883 suicidal ideation and attempt screens found in electronic health records, suicidal ideation and attempt rates were calculated, described, and mapped onto relevant census tracts via the Census Geocoder. Once mapped, a total of 1,098 census tracts were examined for criterion validity and minimum data evaluations.

Results Data indicate that rates of positive suicide risk screens are relatively normally distributed when using a minimum cell size of at least $n=5$, with additional improvements at $n=10$ screens per census tract. Of 48,928 records with completed screens and patient address data listed in the electronic health record, 44,776 addresses (91.5%) were matched to U.S. census tracts via the Census Geocoder database. When evaluating criterion validity, the simultaneous multivariate logistic regression revealed that the model did not fit well to the data, and suicide attempts and suicidal ideation only predicted 0.02% of the variance in the probability of suicide fatality. Finally, a classification tree revealed that a minimum of 10 data points were required to delineate between high and low-risk census tracts.

Conclusion The refined model may act as a helpful tool to evaluate neighborhood level risk and protective factors. Findings suggest a prevention-oriented, as opposed to risk prediction, approach to suicide risk management at the community level may be needed; such an approach would prioritize community connectedness, adequate mental health support services, and reduction of community-level risk factors (e.g., substance misuse), among others.

*Correspondence:
Ryan M. Hill
ryanhill@lsu.edu

Full list of author information is available at the end of the article

Keywords Suicide, Adolescents, Geographic identification, Neighborhood

Suicide rates in the United States (U.S.) have increased dramatically over the past quarter century, from 29,199 deaths in 1999 to 48,344 deaths in 2018 [1]. Suicide is the second leading cause of death among adolescents, and rates continue to rise across all racial and demographic subgroups [1]. In addition, data from the 2021 nationally-representative Youth Risk Behavior Survey indicate that 22.2% of U.S. high school students seriously considered suicide in the previous 12 months and 10.2% reported making a suicide attempt (SA) over that same period [2]. To combat the rising rates of suicide-related behaviors, the 2024 National Strategy for Suicide Prevention has called to "support the development of comprehensive community-based suicide prevention resources for states and communities, and improve the effectiveness of existing community-based suicide prevention efforts" [3].

Suicide is a complex human phenomenon that can be affected by a multitude of varying interactions of psychological, biological, and social risk and protective factors that are both static and dynamic [4, 5]. Due to the complexities of the network of factors that influence an individual's or population's risk of suicide, the Center for Disease Control and Prevention (CDC) [6] proposed that suicide prevention methods follow a four-level social-ecological model that encompasses the interplay of individual, relationship, community, and societal level factors. To date, much of the research dedicated to suicide prevention has focused on the individual and relationship levels of the social-ecological model [7].

To address the need to apply a social-ecological framework to suicide prevention, Cramer and Kapusta [8] developed the Social-Ecological Suicide Prevention Model (SESPM). Through the integration of macro- and micro-level risk and protective factors, the model provides a basis for a more comprehensive understanding of suicide risk. The SESPM further highlighted the lack of macro-level suicide research and prevention efforts, especially at the neighborhood level [8]. Speaking to the promise of addressing macro-level factors, Aytur and colleagues [9] examined contextual risk (e.g., food insecurity) and protective (e.g., community service) factors for youth suicidal thinking in the presence of adverse childhood events. Also, as seen in other areas of health research (e.g., Covid-19 morbidity rates, cardiovascular disease, anxiety disorders), neighborhood level factors can have substantial impacts on a population's health outcomes, especially in marginalized communities [10–12]. From a public health programming perspective, a better understanding of factors among marginalized communities that contribute to poorer health outcomes like suicide would allow for the efficient allocation of limited community resources.

The examination of risk and protective factors at larger socioecological levels and the development of community and neighborhood-level suicide prevention programs targeting identified risk factors requires the ability to evaluate and monitor suicide risk at this level of spatial granularity. Large data repositories and surveillance systems, such as the Youth Risk Behavior Surveillance Survey and the CDC's Web-based Injury Statistics Query and Reporting System, provide state and national data. However, these data sources either accumulate data too slowly to inform community level examination or are conducted in limited areas [1, 2]. The development and validation of brief suicide screening instruments, such as the Ask Suicide Screening Questions Toolkit [13, 14], Columbia-Suicide Severity Rating Scale Screen Version (C-SSRS Screen) [15–17], and Patient Health Questionnaire [18, 19], has led to an increase in universal suicide risk screening within healthcare institutions [20–25]. This screening data, which may already exist within many electronic health records (EHRs), has the potential to identify variations in suicide risk within hospital catchment areas and local communities.

The geographic identification of elevated suicide risk (GIESR) method

In previous work, a model was developed for estimating the relative prevalence of adolescent suicide risk within discrete geographic areas [26]. This method, to be further evaluated in this study, is now titled the Geographic Identification of Suicide Risk (GIESR). The GIESR method used EHR data to examine rates of positive screens for suicidal ideation (SI) and SA among adolescents presenting to the Emergency Department (ED) at the level of U.S. Postal ZIP Codes [26]. Using data from more than 12,000 suicide risk screens, the GIESR model demonstrated that the rate of positive suicide risk screens for 96 ZIP codes in the catchment area of a pediatric children's hospital ranged from 6.17 to 31.03% ($M = 18.33$, $SD = 5.14$) and approximated a normal distribution (skew = 0.19, kurtosis = 0.13). The authors concluded that, at the level of U.S. Postal ZIP Codes, rates of positive screens on the C-SSRS Screen were approximately normally distributed, providing data suitable for potential analysis at a relatively small geographic level [26]. However, a few aspects of that initial work limited the utility of the GIESR method for use in research: (a) U.S. Postal ZIP Codes vary widely in size and may cross neighborhoods with different socioeconomic statuses and characteristics, and (b) U.S. Postal ZIP Codes do not align with other data collected at the national level, such as data from the U.S. Census Bureau, which utilizes the census tract as the level of analysis. Additionally, the initial study by Hill and colleagues [26] included only limited efforts to examine the validity of the proposed model.

The present study

The present study further evaluated and refined the GIESR model for use in research settings by using U.S. census tracts as the level of analysis. Specifically, this study aimed to: (1) examine the rates of positive screens and distributions of the data; (2) to document the ability of the GIESR methodology to utilize EHR data and match it to U.S. census tracts, (3) to examine the criterion validity of the model, in comparison with suicide fatality data, and (4) to examine the least amount of data needed to derive accurate estimates of the rate of positive screens for SI or SA at the level of U.S. census tracts. For Aim 1, we hypothesized that rates of positive screens would be normally distributed across census tracts. Aims 2–4 were considered exploratory and no a priori hypotheses were made. Successful application of the GIESR model at the level of U.S. census tracts would allow for more detailed examination of neighborhood level variation in suicide-related behaviors and the macro-level risk and protective factors that may impact rates of suicide-related behaviors in youth.

Method

Participants and procedures

This project utilized SI and SA data drawn from the EHR of a large pediatric ED system in Texas, which included EDs in three hospitals. Adolescents, ages 11 years and older, presenting to the pediatric ED were asked to complete the C-SSRS Screen as part of standard ED practice. Positive risk screens were handled in accordance with hospital protocols, including the provision of mental health resources and contact with social work and/or psychiatry staff as indicated. Data were drawn from ED visits occurring between January 2018 and December 2022. Respondents ($N = 41{,}296$) ranged in age from 11 to 18 years of age ($M = 14.05$, $SD = 1.94$) and were 56.3% female (43.7% male). Respondents' self-identified race and ethnicity, as identified in the EHR, were: 48.9% Hispanic, 26.1% non-Hispanic White, 18.4% non-Hispanic Black/African American, 3.0% Asian, 0.1% Native Hawaiian or Pacific Islander, 0.1% Native American or Alaskan Native, 3.5% multiracial, unidentified, or other race/ethnicity. This records review study was approved by Louisiana State University's Institutional Review Board and the study was conducted in accordance with the Common Rule. Of note, the pediatric ED system included in this study is one of several hospitals that serve pediatric patients in a major metropolitan area. Consequently, youth from across the geographic region may not be equally represented in this sample.

Measures

Suicidal ideation and suicide attempts

The C-SSRS Screen [27] is a brief three to seven-item screen for past-month SI and recent/lifetime SA. A positive SI screen was defined as a "yes" response to any of the first five items assessing recent ideation. A positive SA screen was defined as a "yes" response to item six (lifetime history of SA). The C-SSRS Screen has been extensively validated in adolescent and adult samples and successfully integrated into large health systems [16, 28]. Of note, both active and passive SI were considered as a positive screen in this analysis as research has suggested that there is a lack of differential association with known risk factors across active and passive SI and because passive SI is significantly associated with suicide attempts [29].

Suicide fatality data

Data on child and adolescent suicide deaths were provided by the Harris County Child Fatality Review Team, a multidisciplinary group that reviews non-natural (homicide, suicide, accident, and undetermined) deaths of children in Harris County, Texas. The team gathers data and information about each death from multiple sources, including the district attorney's office, the medical examiner's office, emergency responders, hospitals, and a range of community providers or resources. The medical examiner's office determines the cause of death. Location data were provided for all pediatric suicide deaths occurring in Harris County, Texas, between 2008 and 2020, totaling 269 suicide fatalities among children under 18 years. Location data was used to identify the census tract where the suicide death occurred; census tracts were coded 1 *"fatality present"* or 0 *"no fatality."*

Census tract characteristics

Data describing census tracts within Houston and Harris County, Texas, were informed by publicly available data from the U.S. Census Bureau, accessed via the Houston State of Health website which compiles data from the U.S. Census Bureau and other national data sources [30]. Data within the U.S. Census Bureau databases were collected via the 2021 American Community Survey, a yearly survey that provides social, economic, housing, and demographic characteristics using 5-year estimates for specific geographic areas) [31]. Crime data were gathered via publicly available reports from the Houston Police Department website [32].

Data analysis

Data were screened and merged using Microsoft Excel and SPSS, version 28. The initial data set included records from the EHR and REDCap databases for the ED from January 2018 to December 2022. The initial data comprised 74,883 records with completed C-SSRS Screens. First, duplicates resulting from the dual data sources were removed, resulting in 72,791 records. To prevent individuals with multiple ED visits from biasing results, only the first ED visit with a completed C-SSRS Screen was retained, resulting in 52,097 records. Next, records with no address data were removed from the data set, resulting in 48,928 complete records. Finally, the remaining records were run through the U.S. Census Bureau's Census Geocoder database [33] to identify census tracts based on the address data provided in the EHR for each record. Census records provided matches for 44,776 records (91.5% of those with address data). Finally, to support comparisons with pediatric suicide fatality data, the analysis was limited to those ages 11–17, reducing the number of records to 41,296 records.

To evaluate Aim 1 (replication of the GIESR model), we first identified rates of positive SI and SA screens for each census tract, computed as the number of positive screens for a given census tract divided by the total number of screens for that census tract. We then examined means and standard deviations, skewness, and kurtosis of the rates of positive SI and SA screens across census tracts. To evaluate Aim 2 (matching EHR data to census tracts), EHR address data were matched to U.S. census tracts using Census Geocoder [33]. After matching, we randomly selected 100 non-matched addresses for further examination by the research team. Unmatched addresses were examined by two team members, both visually and via entry into Google Maps [34], to identify potential factors resulting in a failure to match. Several potential sources of mismatch were identified and coded (e.g., typographical errors, incomplete data, use of Post Office Boxes, etc.).

To evaluate Aim 3 (criterion validity), we conducted a simultaneous multivariate logistic regression, where the percentage of positive screens for SA and the percentage of positive screens for SI were entered as predictors of suicide fatalities. Prior studies indicate the suicide-related behaviors, especially suicide attempts, serve as indicators of risk for suicide, thus offering a means for estimating criterion validity [35]. Stata SE v. 18.0 [36] was used for all analyses. A total of 1,098 census tracts were used to evaluate the association between SA, SI, and suicide fatalities in Harris County, TX. Prior to analysis, assumptions of multivariate logistic regression were assessed, including checking for extreme outliers using scatterplots, examining for multicollinearity of predictors using Pearson's r, and checking for a linear association between predictors and the logit of the outcome using the Box-Tidwell test. Further, multicollinearity was assessed again after running the multivariate logistic regression by examining variance inflation factors. Of these, there was some multicollinearity between SA and SI ($r = .54$, $p < .001$) and clear linear associations between counts of SA and SI within census tracts (see Fig. 1), though it was within acceptable range (VIF = 3.59 for each predictor). Further, ten extreme outliers were identified, which were iteratively excluded from analyses to ensure that results remained unchanged. All other assumptions were met.

To evaluate Aim 4 (minimum data), we created classification trees, where risk for SA or SI was entered into a machine learning model with census tract-level predictors. Risk was calculated by the percentage of positive SA screens in any given census tract, where high risk was deemed greater than 14.3% positive screens, the mean percentage across all sampled census tracts. The data were split such that 70% of the sample was used to train the model ($n = 768$), and 30% was used to test the model ($n = 330$), and sensitivity, specificity, and area under the curve (*AUC*) calculated for model accuracy on the test set [37, 38]. The minimum number of required data points was tuned using the training set using a random search with 10 draws of minimum data points [39] and trained by resampling the training set using ten folds with replacement [40]. RStudio v. 2024.04.1 [41] was used for analysis.

Results

Aim 1. Examine the rates of positive screens and distributions of the data at various minimum cell sizes

A total of 2,388 census tracts were represented in the screening data. Results are displayed in Table 1. Of 2,388 census tracts with at least a single screen, 1,574 had at least five screens, 1,240 had at least ten screens, 768 had at least 20 screens, 469 had at least 30 screens, and 250 had at least 40 completed screens. Of note, estimates of the mean rate of positive screens for SI and SA were similar across all minimum cell sizes. Notably, the standard deviations for the mean rates of positive screens were largest when examining census tracts with at least one screen (compared to all other minimum cell sizes). They were substantially reduced for census tracts with at least 5 and 10 (or more) completed screens. The shift from 1 completed screen to at least 10 (or more) completed screens also substantially reduced the range (e.g., removing census tracts with 100% screening rates), skew, and kurtosis. At all minimum cell values, the data remain somewhat non-normally distributed, as indicated by significant Kolmogorov-Smirnov and Shapiro-Wilk tests, except for the distribution of rates of positive SI screens when a minimum of 40 completed screens were present. Taken together, data indicate that rates of positive SI

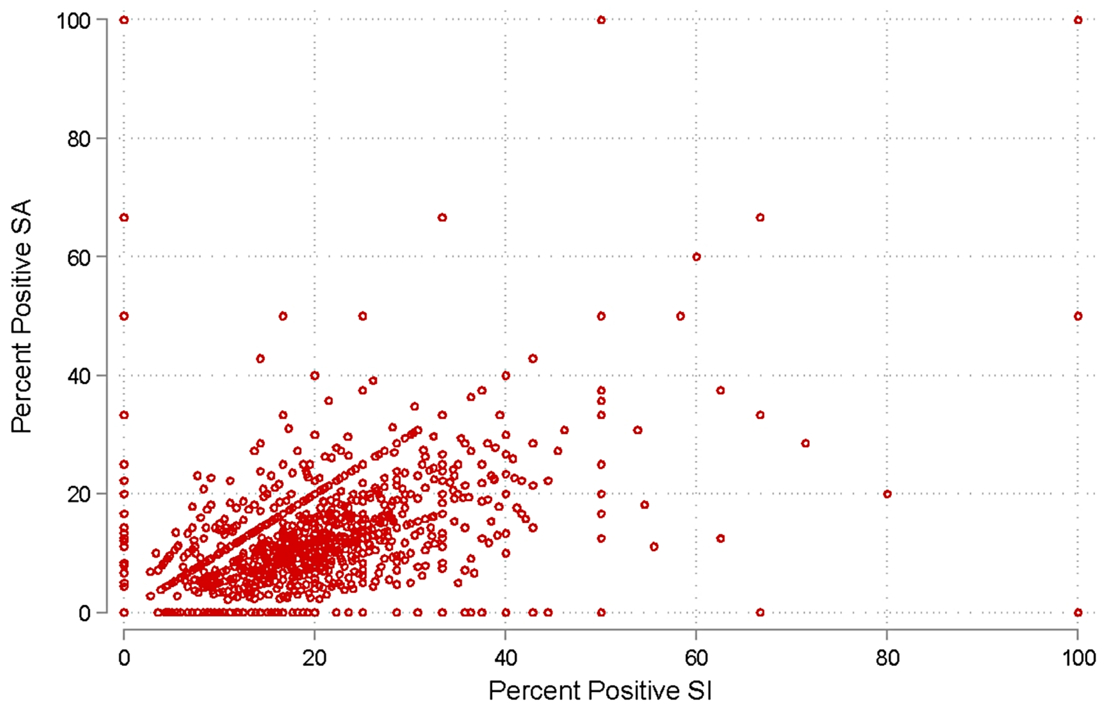

Fig. 1 Evidence of Multicollinearity and Extreme Outliers. *Note.* SA = suicide attempts; SI = suicidal ideation

Table 1 Normality of rates of positive suicidal ideation and suicide attempt screens by minimum cell size

	N	Mean (SD)	Range Min, Max	Skew Est. (SE)	Kurtosis Est. (SE)	Kolmogorov-Smirnov test	Shapiro-Wilk
Census Tracts ≥ 1 screen	2388						
Suicidal ideation		0.19 (*0.21*)	0.00, 1.00	2.15 (*0.05*)	5.69 (*0.10*)	0.19***	0.76***
Suicide attempt		0.12 (*0.18*)	0.00, 1.00	3.24 (*0.05*)	12.62 (*0.10*)	0.26***	0.62***
Census Tracts ≥ 5 screens	1574						
Suicidal ideation		0.19 (*0.11*)	0.00, 0.71	0.55 (*0.06*)	1.16 (*0.12*)	0.07***	0.97***
Suicide attempt		0.12 (*0.09*)	0.00, 0.60	0.92 (*0.06*)	1.55 (*0.12*)	0.09***	0.93***
Census Tracts ≥ 10 screens	1240						
Suicidal ideation		0.19 (*0.09*)	0.00, 0.58	0.44 (*0.07*)	0.81 (*0.14*)	0.06***	0.98***
Suicide attempt		0.12 (*0.07*)	0.00, 0.50	0.69 (*0.07*)	0.89 (*0.14*)	0.07***	0.97***
Census Tracts ≥ 20 screens	768						
Suicidal ideation		0.19 (*0.08*)	0.00, 0.52	0.33 (*0.09*)	0.52 (*0.17*)	0.05***	0.99***
Suicide attempt		0.11 (*0.06*)	0.00, 0.41	0.73 (*0.95*)	0.95 (*0.17*)	0.06***	0.97***
Census Tracts ≥ 30 screens	469						
Suicidal ideation		0.18 (*0.07*)	0.00, 0.41	0.31 (*0.11*)	0.44 (*0.21*)	0.05**	0.99***
Suicide attempt		0.11 (*0.06*)	0.00, 0.41	0.95 (*0.11*)	1.87 (*0.21*)	0.08***	0.95***
Census Tracts ≥ 40 screens	250						
Suicidal ideation		0.18 (*0.06*)	0.03, 0.39	0.08 (*0.14*)	0.29 (*0.28*)	0.03	0.99
Suicide attempt		0.11 (*0.05*)	0.00, 0.41	1.44 (*0.14*)	4.52 (*0.28*)	0.08***	0.92***

Note. N = total number of census tracts; SD = standard deviation; SE = standard error; ** $p < .01$; *** $p < .001$

and SA screens are relatively normally distributed when using a minimum cell size of at least $n=5$, with additional improvements at $n=10$ screens per census tract.

Aim 2. Examine match between EHR data and U.S. Census tracts

Of 48,928 records with completed screens and patient address data listed in the EHR, 44,776 addresses (91.5%) were matched to U.S. census tracts via the Census Geocoder database. To examine reasons for the non-match of addresses in the Census Geocoder database, the research team reviewed a randomly selected subset of 100 non-matched addresses. Visual inspection and internet searches were conducted on the 100 randomly selected addresses. Examination revealed that 33% appeared to have street name, city, or ZIP code errors in the EHR data, likely resulting in a lack of exact matches; 14% were Post Office boxes and, therefore, could not be matched to census tracts, and 12% appeared to be potentially incomplete (e.g., lacking a key address element). The remaining 41% appeared to be complete and accurate, with no apparent reason for failing to match in the Census Geocoder database. This could indicate errors within Census Geocoder itself or may reflect unusual issues (e.g., newer homes not yet listed in Census Geocoder). This relatively low number of non-matches (< 10%) has been common, and sometimes greater, across previous studies examining health data at the census tract level [41–44]. Non-matched data were not included in any study analyses.

Aim 3. Examine the criterion validity of the model

The simultaneous multivariate logistic regression revealed that the model did not fit well to the data, and SA and SI only predicted 0.02% of the variance in the probability of suicide fatality ($LR\ \chi^2(2)=0.19$, $p=.900$, McFadden's $R^2=0.00$, McFadden's Adj$R^2=0.00$). Removing each extreme outlier (see Fig. 1) iteratively resulted in the same model fit and logit coefficients. The final model suggests that there is no increase in probability for suicide fatalities with each percentage point increase in SA (OR = 1.34; $b=0.28$; $z=0.36$, $p=.722$) or SI (OR = 0.74; b = -0.30; z = -0.40, $p=.686$). Repeating the analysis with SI ($LR\chi^2(1)=0.03$, $p=.871$, McFadden's $R^2=0.00$, McFadden's AdjR^2 = -0.00) and SA ($LR\chi^2(1)=0.03$, $p=.871$, McFadden's $R^2=0.00$, McFadden's AdjR^2 = -0.00) as predictors in separate models, we discovered the same result; neither SA nor SI were predictive of suicide fatality.

Aim 4. Examine the amount of data needed to differentiate risk level for SI or SA at the level of U.S. Census tracts

The classification tree revealed that a minimum of 16 data points were required to delineate between high (SI > 33.8%; SA > 24.1%) and low-risk (SI < 33.8%; SA < 24.1%) census tracts. To predict SI, the percentage of positive SA screens was most important to the model, contributing about 27.4% of the predictive capacity of the model (see Fig. 2). The total number of screens was also a valuable predictor, contributing about 7.8% of the predictive capacity. In predicting SA, the percentage of positive SI screens was the most valuable, contributing about 36% of the predictive capacity of the model, though total screened had substantially less influence at 0.5% (see Fig. 3). However, neither models were especially accurate for predicting high-risk census tracts. The model predicting SI risk could only successfully classify 23.3% of high-risk census tracts, while the model predicting SA risk could only successfully classify 17.1% of high-risk census tracts. Both models were much better at predicting low-risk counties (95.9% for low-risk SA; 96.3% for low-risk SI). This differentiation is likely due to the comparative rarity of actual SA and SI outcomes, which would require a much larger sample for model training and evaluation.

Discussion

The present study sought to evaluate and refine the Geospatial Identification of Elevated Suicide Risk (GIESR) model for use as a tool to evaluate neighborhood level risk and protective factors for SI and SA [26]. This study sought to extend the prior model, which utilized U.S. Postal ZIP Codes, for use with U.S. census tracts, thereby opening additional opportunities for integration with a range of data collected by the U.S. Census Bureau. Specifically, this study aimed to: (1) examine the rates of positive screens and distributions of the data at various minimum cell sizes; (2) to document the ability of the GIESR methodology to utilize EHR data and match it to U.S. census tracts, (3) to examine the criterion validity of the model, in comparison with suicide fatality data, and (4) to examine the minimal amount of data needed to derive accurate estimates of the rate of positive screens for SI or SA at the level of U.S. census tracts.

Aim 1 examined the rates of positive SI and SA screens, with a focus on the distribution of the data when different minimum cell sizes (i.e., minimum number of screens per census tract) are included. Results indicated that restricting the data to census tracts with at least ten completed screens yields data that, while not entirely normally distributed, has a more realistic range (i.e., removing census tracts with 100% positive SI and SA screening rates), with acceptable skew and kurtosis values. Further cell size restrictions resulted in relatively little change in range, skew, and kurtosis, but did reduce the number of census tracts with adequate data, reducing sample size for subsequent analyses. While the original GIESR model used a conservative cutoff of at least 50 screens per geographic unit, this revised model, created with a much larger sample, appears to stabilize with a smaller minimum cell size of ten [26].

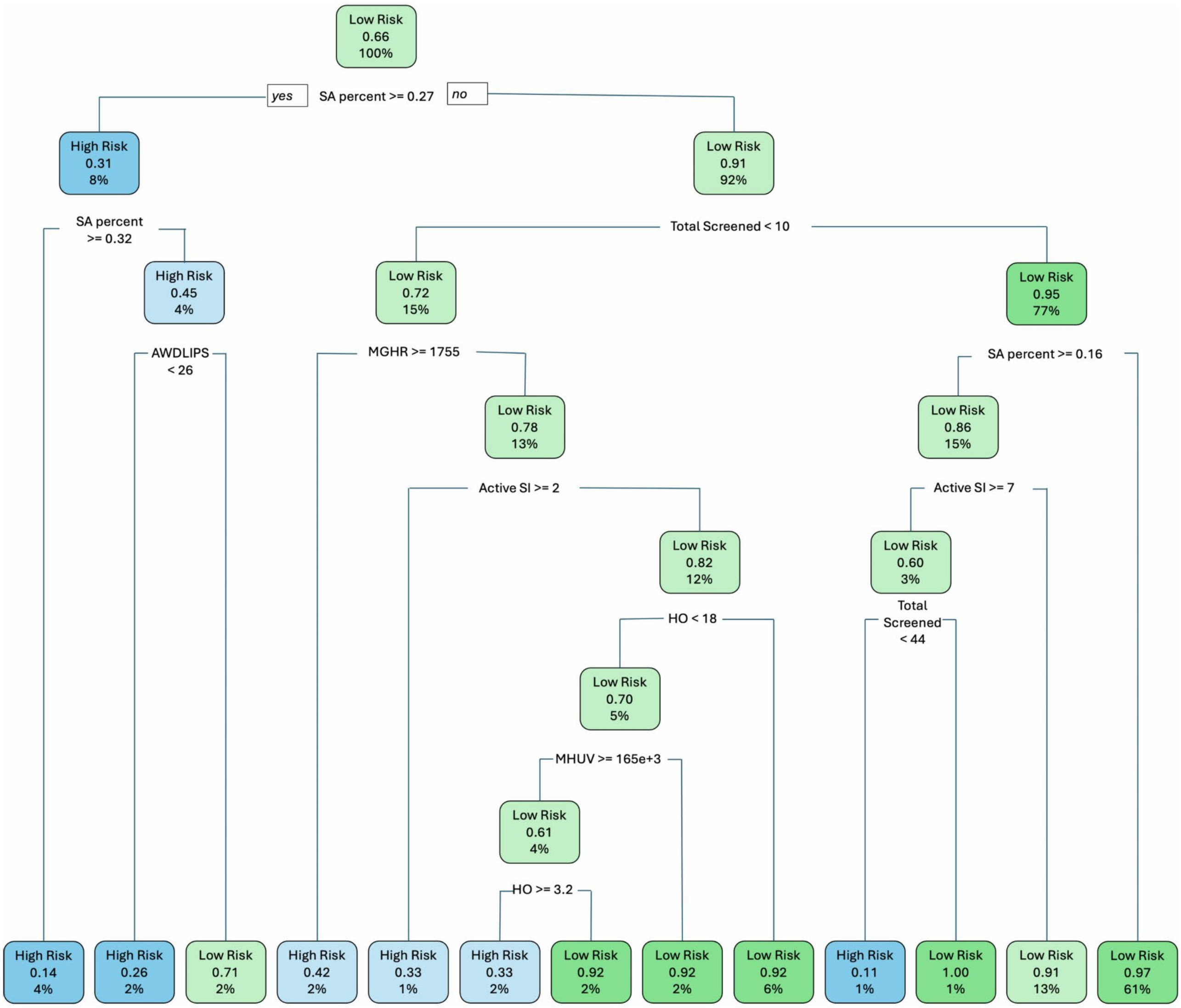

Fig. 2 Classification Tree for Predicting High and Low Suicidal Ideation Risk Census Tracts. *Note.* SApercent = percentage of positive suicide attempt screens; TotalScreened = total number of suicide attempt or ideation screens in census tract; AWDLIP5 = 5 year average percentage of adults aged 20–64 living with a disability and below poverty level; MHGR = median household gross rent; HO = percentage of all housing units (i.e., occupied and unoccupied) that are occupied by homeowners; MHUV = median housing unit value

Aim 2 evaluated the accuracy of the Census Geocoder tool and the extent of data loss due to the conversion of address data to census tract-level data. The switch to census tracts from U.S. Postal ZIP codes used in the previous study produces more detailed information about specific locations in the analysis [26]. However, the additional step of deriving census tracts based on addresses is required. This conversion process resulted in a loss of approximately 8.5% of the data. While it proved difficult to identify why some addresses did not match the Census Geocoder database, possible reasons included the use of post office boxes as addresses, errors in the medical record (e.g., misspelling street names), and potential issues with the Census Geocoder tool (e.g., new addresses that may not have been added to the database). Overall, the loss of a small percentage of data for a substantial increase in spatial granularity in the results may be beneficial, provided a large enough initial data set is available.

Aim 3 evaluated the criterion validity of the GIESR model, comparing rates of positive SI and SA screens with the location of pediatric suicide fatalities over the previous decade. Results indicated that rates of positive screens were not significantly associated with the locations of suicide deaths. The rate of positive screens was calculated using those who visited the ED as the denominator, a group more likely to be at risk for many health outcomes, including suicide. As such, this rate would not be a true population rate of SI or SA, rather, most likely, an overestimation of the true rate. Consequently, it may

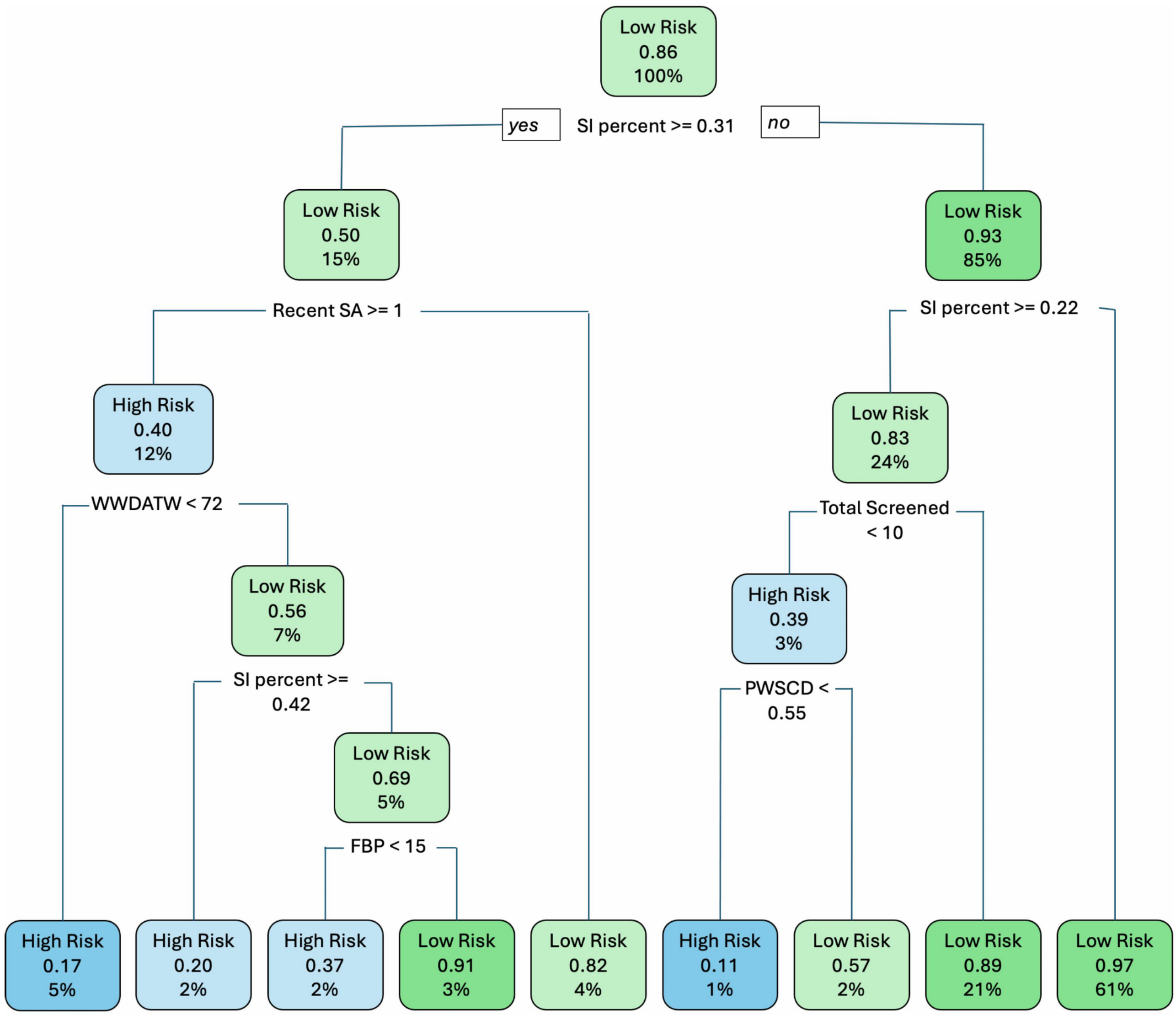

Fig. 3 Classification Tree for Predicting High and Low Suicide Attempt Risk Census Tracts. *Note*. SIpercent = percentage of positive suicidal ideation screens; RecentSA = number of recent suicide attempts reported; WWDATW = percentage of workers aged 16 years and over who get to work by driving alone in a car, truck, or van; PWSCD = percentage of the population with a self-care difficulty; FBP = percentage of the population who are a foreign born (persons not born in the United States, including all foreign born persons regardless of whether they are naturalized U.S. citizens

be unlikely to observe a screen positive rate highly associated with the suicide mortality rate as screen rates are not measures of population risk; a possible explanation for the model's low criterion validity. Of note, research on individual level prediction of suicide-related behaviors has proven particularly ineffective, with most risk and protective factors unable to predict risk much better than chance [45]. The present study results, therefore, replicate the individual level data, indicating poor predictive value with respect to suicide deaths. Findings suggest a prevention-oriented, as opposed to risk prediction, approach to suicide risk management at the community level may be needed; such an approach would prioritize community connectedness, adequate mental health support services, and reduction of community-level risk factors (e.g., substance misuse), among others [46].

Aim 4 examined the minimum amount of data required to predict high and low-risk census tracts. Results suggested that a minimum of 10 data points would be required to predict risk level by location and that at least two suicide risk screens are required to determine risk. The resulting model was overly sensitive, misclassifying a significant percentage of low-risk census tracts as high-risk. The accuracy of identifying high-risk census tracts was limited, and future research is needed with a larger overall volume of positive screens may help to improve the sensitivity of these models. As part of the process of identifying the minimum data required, we also discovered that two variables related to speaking a language other than English contributed substantially to the predictive power of suicide risk, including language isolation and a larger percentage of the census tract struggling to speak English, echoing prior work showing the heightened risk of suicidal behavior among ethnic minorities and immigrants [47].

Contributions to theory and practice

As noted by Cramer and Kapusta's [8] SESPM, in comparison with the volume of literature examining suicide prevention at the individual and interpersonal levels, very little research has examined suicide-related thoughts and behaviors at the neighborhood level. This is, perhaps, due in part to the lack of available measurement of suicide-related thoughts and behaviors at the neighborhood level, as current surveillance systems do not function at the level of U.S. census tracts. Suicide fatalities are sufficiently rare so as to preclude nuanced spatial analysis. However, in the newly released *2024 National Strategy for Suicide Prevention*, strategic directions 1 and 3, Community-Based Suicide Prevention; and Surveillance, Quality Improvement, and Research, respectively, highlight the importance of high-quality data and community-level interventions to mitigate suicide-related risk factors and strengthen protective factors at the individual and neighborhood levels [48]. The GIESR model aims to address this gap, utilizing EHR data derived from routine suicide risk screening to promote suicide prevention research at the neighborhood or community level. The addition of empirical research on suicide-related thoughts and behaviors at the neighborhood level is needed to develop and refine theoretical models of suicide that incorporate a broad range of contextual factors. That is, while existing models of suicide-related thoughts and behaviors consider individual and interpersonal factors that contribute to suicide risk, the GIESR model may help identify high risk catchment areas characterized by certain suicide- and culturally-related factors.

Identification of culturally-related factors defining high suicide risk census tracts fits with existing theoretical and empirical work concerning multicultural perspectives on suicide. For instance, Chu and colleagues' cultural model of suicide includes unique ethnic/racial minority expressions of suicide (that may or may not be adequately captured by the C-SSRS Screen) and social discord (e.g., intergenerational conflict stemming from linguistic barriers) [49, 50]. Likewise, evidence links acculturative stress and acculturation to suicide risk among ethnic and racial minority youth [51, 52]. Our findings regarding languages other than English spoken in homes as a potential community level risk factor for suicide offers indirect affirmation of the role of such cultural factors. Indeed, this pattern of findings raises a valuable next future research direction in the form of examining associations between community-level indicators of structural stigma or acculturative stress and high-risk geographic areas. Such work would also respond to recent calls for increased focus on intersectionality and cultural factors in suicide research [53, 54].

Future research should evaluate potential risk and protective factors at the neighborhood level, such as social determinants of health that may impact the overall rate of suicide-related thoughts and behaviors across census tracts. Better understanding neighborhood level risk and protective factors may lead to public health interventions targeted toward these risk and protective factors, with the aim of creating neighborhoods and spaces that contribute to reduced risk of suicide-related thoughts and behaviors. As recommended by the social-ecological model, health outcomes must be approached from all levels of society [6]. If research continues to skew toward individual and interpersonal factors, efforts may miss opportunities for prevention and intervention of suicide-related thoughts and behaviors spanning across all social-ecological levels [8].

Limitations and future directions

The findings of this study should be considered in the context of the study's limitations. Data were drawn from the EHR of a large pediatric ED system and, because this ED system is not the only provider within the greater metropolitan area, the sample may not be representative of the region's general population. Adolescents with higher levels of risky behavior, histories of abuse, and higher depression scores, as well as adolescents with fewer financial resources, may be more likely to use the ED compared to their peers and thus may be overrepresented in this sample [55]. Additionally, while past research indicates that youth do not systematically alter their responses to suicide screening in the ED when follow-up may occur, some youth may have withheld reporting SI to avoid potential psychiatric hospitalization [56, 57]. Consequently, estimates of rates of positive SI and SA screens may not reflect true population rates and may not align with anonymous studies of youth, such as the Youth Risk Behavior Survey [58]. If possible, future research may consider gathering data (i.e., screens) universally from all relevant providers (e.g., all hospitals in a given catchment area using the same screening tool) to estimate rates in relation to the total population of the given area. This method of estimation may provide a better estimate of true population rates.

Universal suicide screening in emergency centers is feasible and acceptable among pediatric patient populations [28, 59] and has become widespread in hospitals due to the 2019 Joint Commission requirement to screen all patients with a primary behavioral health concern

[60]. Recently, the American Academy of Pediatrics recommended universal suicide screening among pediatric patients ≥ 12 years in primary care, regardless of behavioral health concerns [61]. As universal suicide-specific screening in primary care becomes more common, the GIESR method can be applied to a more general pediatric population more representative of the community surrounding primary care clinics. Finally, the use of suicide fatality data from the preceding decade was also a limitation, as the dependent variable for Aim 3 primarily occurred prior to the onset of routine ED screening and historical changes (e.g., rising rates of suicide risk) may have impacted the analysis [62]. However, an extensive data range for suicide fatality data was necessary to provide sufficient data for examination. Further, to better assess important contributions to risk and predict future risk, researchers may wish to use more sensitive machine learning tools, such as zero-inflated Poisson regression models [63] to effectively establish predictive capacity for available data.

Additional research is needed to evaluate and refine the GIESR model in different regions and based on different suicide screening tools [14, 19]. A similar model could also be utilized with other standardized screens, such as depression screens. As suicide risk screening becomes routine in a greater number of hospital systems, additional research should also consider the possibility of using GIESR to evaluate changes in positive screen rates over time and the amount of data necessary to detect temporal changes. If successful, the use of the GIESR model to detect a change in positive screen rates over time would provide a means for ongoing monitoring and, potentially, an avenue for evaluating the effectiveness of target public health suicide prevention programs. Future research should also consider the possibility of using the GIESR approach to predict positive screen rates for SI and SA. For public health programs to effectively address SI and SA, it is important to evaluate factors affecting risk across all social-ecological levels. While the majority of research examining risk and protective factors for SI and SA has focused on the intrapersonal and interpersonal levels [8], understanding community level impacts on SI and SA may lead to novel preventive interventions. To understand risk and protective factors for suicide-related thoughts and behaviors beyond individual and interpersonal levels, future research should consider allocating efforts toward factors at broader granularities, such as census tracts.

Abbreviations

AUC	Area under the curve
C-SSRS Screen	Columbia-Suicide Severity Rating Scale Screen Version
CDC	Centers for Disease Control and Prevention
ED	Emergency department
EHRs	Electronic health records
GIESR	Geographic Identification of Elevated Suicide Risk
SA	Suicide attempt
SESPM	Social-Ecological Suicide Prevention Model
SI	Suicidal ideation
U.S.	United States
VIF	Variance inflation factors

Acknowledgements
Not applicable.

Author contributions
RMH was responsible for conceptualization, funding acquisition, data curation, formal analysis, investigation, supervision, methodology, project administration, and writing (original draft and review/editing); SCC was responsible for data curation, formal analysis, methodology, visualization, and writing (original draft and review/editing); AS was responsible for data curation, formal analysis, and writing (original draft and review/editing); TH was responsible for data curation and writing (original draft and review/editing); JB was responsible for data curation and writing (original draft and review/editing); AC was responsible for writing (original draft and review/editing); AH was responsible for writing (original draft and review/editing); RJC was responsible for writing (original draft and review/editing).

Funding
This work was funded through R21MH128557-01.

Data availability
Data will be available through the National Institute of Mental Health Data Archive.

Declarations

Ethics approval and consent to participate
This chart review study was approved by Louisiana State University's Institutional Review Board and conducted in accordance with the Common Rule.

Consent for publication
Per Louisiana State University's Institutional Review Board, consent was not required for this study as it was conducted as a chart review.

Competing interests
The authors declare no competing interests.

Author details
[1]Department of Psychology, Louisiana State University, Baton Rouge, LA, USA
[2]Department of Epidemiology and Community Health, University of North Carolina at Charlotte, Charlotte, NC, USA
[3]Department of Pediatrics, Baylor College of Medicine, Houston, TX, USA
[4]Department of Healthcare Transformation Initiatives, University of Texas Health Science Center Houston, Houston, TX, USA
[5]236 Audubon Hall, 1105 Field House Drive, Baton Rouge, LA 70802, USA

Received: 9 September 2024 / Accepted: 25 March 2025
Published online: 17 April 2025

References

1. Center for Disease Control and Prevention. WISQARS (Web-based Injury Statistics Query and Reporting System), Injury Center [Internet]. CDC; 2020 [cited 2024]. Available from: https://www.cdc.gov/injury/wisqars/index.html
2. Centers for Disease Control and Prevention. Youth Risk Behavior Survey 2021 Results [Internet]. CDC; 2023 [cited 2024]. Available from: https://www.cdc.gov/healthyyouth/data/yrbs/results.htm
3. National Action Alliance for Suicide Prevention. The nation's public-private partnership for suicide prevention [Internet]. [cited 2024]. Available from: https://theactionalliance.org/sites/default/files/action_alliance_marketing_handout_2.pdf
4. de Beurs D, Bockting C, Kerkhof A, Scheepers F, O'Connor R, Penninx B, et al. A network perspective on suicidal behavior: Understanding suicidality as a

complex system. Suicide Life-Threatening Behav. 2021;51(1):115–26. https://doi.org/10.1111/sltb.12676.
5. Steele IH, Thrower N, Noroian P, Saleh FM. Understanding suicide across the lifespan: A united States perspective of suicide risk factors, assessment & management. J Forensic Sci. 2017;63(1):162–71. https://doi.org/10.1111/1556-4029.13519.
6. Center for Disease Control and Prevention. The social-ecological model: A framework for prevention. [Internet]. CDC; 2022 [cited 2024]. Available from: https://www.cdc.gov/violenceprevention/about/social-ecologicalmodel.html
7. White JH, Marsh I, Kral MJ, Morris J. Critical suicidology: transforming suicide research and prevention for the 21st century. Vancouver: UBC; 2016.
8. Cramer RJ, Kapusta ND. A social-ecological framework of theory, assessment, and prevention of suicide. Front Psychol. 2017;8. https://doi.org/10.3389/fpsyg.2017.01756.
9. Aytur SA, Carlino S, Bernard F, West K, Dobrzycki V, Malik R. Social-ecological theory, substance misuse, adverse childhood experiences, and adolescent suicidal ideation: applications for community–academic partnerships. J Community Psychol. 2021;50(1):265–84. https://doi.org/10.1002/jcop.22560.
10. Isasi F, Naylor MD, Skorton D, Grabowski D, Hernandez S, Montgomery Rice V. Patients, families, and communities COVID-19 impact assessment: lessons learned and compelling needs. NAM Perspect. 2021;11. https://doi.org/10.31478/202111c.
11. Kreatsoulas C, Anand SS. The impact of social determinants on cardiovascular disease. Can J Cardiol. 2010;26:8 C-13 C. https://doi.org/10.1016/s0828-282x(10)71075-8.
12. Nunes JC, Carroll MK, Mahaffey KW, Califf RM, Doraiswamy PM, Short S, et al. General anxiety disorder-7 questionnaire as a marker of low socioeconomic status and inequity. J Affect Disord. 2022;317:287–97. https://doi.org/10.1016/j.jad.2022.08.085.
13. Horowitz LM, Snyder DJ, Boudreaux ED, He J-P, Harrington CJ, Cai J, et al. Validation of the ask suicide-screening questions for adult medical inpatients: A brief tool for all ages. Psychosomatics. 2020;61(6):713–22. https://doi.org/10.1016/j.psym.2020.04.008.
14. NIMH. Ask suicide-screening questions (ASQ) toolkit [Internet]. U.S. Department of Health and Human Services; Available from: https://www.nimh.nih.gov/research/research-conducted-at-nimh/asq-toolkit-materials/index.shtml
15. Chappell P, Feltner DE, Makumi C, Stewart M. Initial validity and reliability data on the Columbia–Suicide severity rating scale. Am J Psychiatry. 2012;169(6):662–3. https://doi.org/10.1176/appi.ajp.2012.12010123.
16. Gipson PY, Agarwala P, Opperman KJ, Horwitz A, King CA. Columbia-Suicide severity rating scale. Pediatr Emerg Care. 2015;31(2):88–94. https://doi.org/10.1097/pec.0000000000000225.
17. Posner K, Brown GK, Stanley B, Brent DA, Yershova KV, Oquendo MA, et al. The Columbia–Suicide severity rating scale: initial validity and internal consistency findings from three multisite studies with adolescents and adults. Am J Psychiatry. 2011;168(12):1266–77. https://doi.org/10.1176/appi.ajp.2011.10111704.
18. Gilbody S, Richards D, Brealey S, Hewitt C. Screening for depression in medical settings with the patient health questionnaire (PHQ): A diagnostic meta-analysis. J Gen Intern Med 2007 Sept 14;22(11):1596–602. https://doi.org/10.1007/s11606-007-0333-y
19. Johnson JG, Harris ES, Spitzer RL, Williams JBW. The patient health questionnaire for adolescents: validation of an instrument for the assessment of mental disorders among adolescent primary care patients. J Adolesc Health. 2002;30(3):196. https://doi.org/10.1016/s1054-139x(01)00333-0.
20. Bahraini N, Brenner LA, Barry C, Hostetter T, Keusch J, Post EP, et al. Assessment of rates of suicide risk screening and prevalence of positive screening results among US veterans after implementation of the veterans affairs suicide risk identification strategy. JAMA Netw Open. 2020;3(10). https://doi.org/10.1001/jamanetworkopen.2020.22531.
21. Betz ME, Arias SA, Segal DL, Miller I, Camargo CA, Boudreaux ED. Screening for suicidal thoughts and behaviors in older adults in the emergency department. J Am Geriatr Soc. 2016;64(10):e72–7. https://doi.org/10.1111/jgs.14529
22. Diamond GS, Herres JL, Krauthamer Ewing ES, Atte TO, Scott SW, Wintersteen MB, et al. Comprehensive screening for suicide risk in primary care. Am J Prev Med. 2017;53(1):48–54. https://doi.org/10.1016/j.amepre.2017.02.020.
23. Heyland M, Delaney KR, Shattell M. Steps to achieve universal suicide screening in emergency departments: A call to action. J PsychoSoc Nurs Ment Health Serv. 2018;56(10):21–6. https://doi.org/10.3928/02793695-20180503-03.
24. Horowitz LM, Snyder D, Ludi E, Rosenstein DL, Kohn-Godbout J, Lee L, et al. Ask suicide-screening questions to everyone in medical settings: the Asq'em quality improvement project. Psychosomatics. 2013;54(3):239–47. https://doi.org/10.1016/j.psym.2013.01.002.
25. Horowitz LM, Boudreaux ED, Schoenbaum M, Pao M, Bridge JA. Universal suicide risk screening in the hospital setting: still a Pandora's box? Joint Comm J Qual Patient Saf. 2018;44(1):1–3. https://doi.org/10.1016/j.jcjq.2017.11.001.
26. Hill RM, Gushanas KL, Alvis L, Dodd CG, Kurian S, Saxena J, et al. Geospatial identification of high youth suicide risk areas via electronic health records: avenues for research and prevention efforts. Suicide Life-Threatening Behav. 2021;51(2):255–62. https://doi.org/10.1111/sltb.12701.
27. Posner K, Brown GK, Stanley B, Brent DA, Yershova KV, Oquendo MA, Currier GW, Melvin GA, Greenhill L, Shen S, Mann JJ. The Columbia–Suicide severity rating scale: initial validity and internal consistency findings from three multisite studies with adolescents and adults. Am J Psychiatry. 2011;168(12):1266–77.
28. Roaten K, Johnson C, Genzel R, Khan F, North CS. Development and implementation of a universal suicide risk screening program in a safety-net hospital system. Joint Comm J Qual Patient Saf. 2018;44(1):4–11. https://doi.org/10.1016/j.jcjq.2017.07.006.
29. Liu RT, Bettis AH, Burke TA. Characterizing the phenomenology of passive suicidal ideation: A systematic review and meta-analysis of its prevalence, psychiatric comorbidity, correlates, and comparisons with active suicidal ideation. Psychol Med. 2020;50(3):367–83. https://doi.org/10.1017/S003329171900391X.
30. Houston State of Health [Internet]. Houston Public Health Data Portal; [cited 2024]. Available from: https://www.houstonstateofhealth.com/
31. US Census Bureau. American Community Survey (ACS) [Internet]. 2024 [cited 2024]. Available from: https://www.census.gov/programs-surveys/acs/
32. City of Houston [Internet]. Houston Police Department; Available from: https://www.houstontx.gov/police/
33. US Census Bureau. Census geocoder documentation [Internet]. 2024 [cited 2024]. Available from: https://www.census.gov/programs-surveys/geography/technical-documentation/complete-technical-documentation/census-geocoder.html
34. Google [Internet]. Google maps; [cited 2024]. Available from: http://www.google.com/maps
35. Probert-Lindström S, Berge J, Westrin Å, Öjehagen A, Skogman Pavulans K. Long-term risk factors for suicide in suicide attempters examined at a medical emergency in patient unit: results from a 32-year follow-up study. BMJ Open. 2020;10(10). https://doi.org/10.1136/bmjopen-2020-038794.
36. Stata statistical software. Release 18. College station. Tex: Stata; 2003.
37. James G, Witten D, Hastie T, Tibshirani R. An introduction to statistical learning [Internet]. SAGE Publications Ltd; 2013. Available from: https://www.statlearning.com/
38. Raschka S. Model evaluation, model selection, and algorithm selection in machine learning [Internet]. Cornell University; 2020. Available from: https://arxiv.org/abs/1811.12808
39. Bergstra J, Bengio Y. Random search for hyper-parameter optimization. Journal of Machine Learning Research. 2012;13(2). Available at: https://www.jmlr.org/papers/volume13/bergstra12a/bergstra12a.pdf?ref=broutonlab.com
40. Chobtham K, Constantinou AC. Tuning structure learning algorithms with out-of-sample and resampling strategies. Knowl Inf Syst. 2024;66(8). https://doi.org/10.1007/s10115-024-02111-9.
41. RStudio Team. RStudio: An integrated development environment for R [Internet]. 2024. Available from: https://www.rstudio.com/
42. DeMass R, Gupta D, Self S, Thomas D, Rudisill C. Emergency department use and Geospatial variation in social determinants of health: A pilot study from South Carolina. BMC Public Health. 2023;23(1). https://doi.org/10.1186/s12889-023-16136-2.
43. Koziatek CA, Bohart I, Caldwell R, Swartz J, Rosen P, Desai S, et al. Neighborhood-level risk factors for severe hyperglycemia among emergency department patients without a prior diabetes diagnosis. J Urb Health. 2023;100(4):802–10. https://doi.org/10.1007/s11524-023-00771-6.
44. Samuels EA, Taylor RA, Pendyal A, Shojaee A, Mainardi AS, Lemire ER, et al. Mapping emergency department asthma visits to identify poor-quality housing in new Haven, CT, USA: A retrospective cohort study. Lancet Public Health. 2022;7(8). https://doi.org/10.1016/s2468-2667(22)00143-8.
45. Franklin JC, Ribeiro JD, Fox KR, Bentley KH, Kleiman EM, Huang X, et al. Risk factors for suicidal thoughts and behaviors: A meta-analysis of 50 years of research. Psychol Bull. 2017;143(2):187–232. https://doi.org/10.1037/bul0000084.

46. Center for Disease Control and Prevention. Suicide prevention resource for action [Internet]. CDC; 2024 [cited 2024]. Available from: https://www.cdc.gov/suicide/resources/prevention.html
47. Forte A, Trobia F, Gualtieri F, Lamis DA, Cardamone G, Giallonardo V, et al. Suicide risk among immigrants and ethnic minorities: A literature overview. Int J Environ Res Public Health. 2018;15(7):1438. https://doi.org/10.3390/ijerph15071438.
48. U.S. Department of Health and Human Services (HHS), National Strategy for Suicide Prevention. Washington, DC: HHS. 2024. Available at: http://hdl.handle.net/10713/22111
49. Chu JP, Goldblum P, Floyd R, Bongar B. The cultural theory and model of suicide. Appl Prev Psychol. 2010;14(1–4):25–40. https://doi.org/10.1016/j.appsy.2011.11.001.
50. Chu J, O'Neill SE, Ng JF, Khoury O. The cultural theory and model of suicide for youth. Springer Int Publishing. 2022;99–106. https://doi.org/10.1007/978-3-031-06127-1_11.
51. Cho Y-B, Haslam N. Suicidal ideation and distress among immigrant adolescents: the role of acculturation, life stress, and social support. J Youth Adolesc. 2009;39(4):370–9. https://doi.org/10.1007/s10964-009-9415-y.
52. Polanco-Roman L, Ebrahimi CT, Mafnas KS, Hausmann-Stabile C, Meca A, Mazzula SL, et al. Acculturation and suicide-related risk in ethnoracially minoritized youth in the US: A scoping review and content analysis of the empirical evidence. Soc Psychiatry Psychiatr Epidemiol. 2023;58(8):1121–37. https://doi.org/10.1007/s00127-023-02494-0.
53. Molock SD, Boyd RC, Alvarez K, Cha C, Denton E, Glenn CR, et al. Culturally responsive assessment of suicidal thoughts and behaviors in youth of color. Am Psychol. 2023. https://doi.org/10.1037/amp0001140.
54. Standley CJ. Expanding our paradigms: intersectional and socioecological approaches to suicide prevention. Death Stud. 2020;46(1):224–32. https://doi.org/10.1080/07481187.2020.1725934.
55. Wilson KM, Klein JD. Adolescents who use the emergency department as their usual source of care. Arch Pediatr Adolesc Med. 2000;154(4):361–5. https://doi.org/10.1001/archpedi.154.4.361.
56. King CA, Hill RM, Wynne HA, Cunningham RM. Adolescent suicide risk screening: the effect of communication about type of follow-up on adolescents' screening responses. J Clin Child Adolesc Psychol. 2012;41(4):508–15. https://doi.org/10.1080/15374416.2012.680188.
57. Jones N, Gius BK, Shields M, Collings S, Rosen C, Munson M. Investigating the impact of involuntary psychiatric hospitalization on youth and young adult trust and help-seeking in pathways to care. Soc Psychiatry Psychiatr Epidemiol. 2021;56(11):2017–27. https://doi.org/10.1007/s00127-021-02048-2.
58. Gaylor EM, Krause KH, Welder LE, Cooper AC, Ashley C, Mack KA, et al. Suicidal thoughts and behaviors among high school students — youth risk behavior survey, united States, 2021. MMWR Supplements. 2023;72(1):45–54. https://doi.org/10.15585/mmwr.su7201a6.
59. Rufino KA, Kerr T, Beyene H, Hill RM, Saxena J, Kurian S et al. Suicide screening in a large pediatric emergency department. Pediatr Emerg Care 2021 Sept 17;38(3):e1127–32. https://doi.org/10.1097/pec.0000000000002530
60. Joint Commission. R3 report: Requirement, rationale, reference: National patient safety goal for suicide prevention [Internet]. Joint Commission. 2019 [cited 2020 Apr 13]. Available from: https://www.jointcommission.org/-/media/tjc/documents/standards/r3-reports/r3_18_suicide_prevention_hap_bhc_cah_11_4_19_final1.pdf
61. American Academy of Pediatrics. Suicide: Blueprint for youth suicide prevention [Internet]. [cited 2024 Jun 5]. Available from: https://www.aap.org/en/patient-care/blueprint-for-youth-suicide-prevention/
62. Plemmons G, Hall M, Doupnik S, Gay J, Brown C, Browning W, et al. Hospitalization for suicide ideation or attempt: 2008–2015. Pediatrics. 2018;141(6):1–10. https://doi.org/10.1542/peds.2017-2426.
63. Lee SK, Jin S. Decision tree approaches for zero-inflated count data. J Applied Statistics. 2006;33(8):853–65. https://doi.org/10.1080/02664760600743613.

Publisher's note