UC Berkeley
ISSI Fellows Working Papers

Title
Predicting Suicidal Ideation among Native American High Schoolers in California

Permalink
https://escholarship.org/uc/item/5z1751rw

Author
Sierra, Valentín Q

Publication Date
2023-08-29

Data Availability
The data associated with this publication are available at: https://osf.io/4dpwt/

Predicting Suicidal Ideation among Native American High Schoolers in California

by Valentín Quiroz Sierra

School of Social Welfare
University of California, Berkeley
August 21, 2023

Predicting Suicidal Ideation among Native American High Schoolers in California

Valentín Quiroz Sierra

School of Social Welfare
University of California, Berkeley

vqsierra@berkeley.edu

Author Note

This study was registered via the Open Science Framework and can be accessed at https://osf.io/4dpwt/

Predicting Suicidal Ideation among Native American High Schoolers in California

Table of Contents

Abstract

Suicide is the leading cause of non-accidental death for Native American young people ages 15-24 years old. Concerningly, suicide rates have continued to rise over the past decade despite a myriad of prevention efforts. This shortcoming has urged some scholars to (re)examine key theoretical constructs to better direct suicide prevention efforts in tribal communities. Using Indigenous Wholistic Theory, an algorithmic approach was employed to identify a broader set of factors that may influence suicidal ideation among Native American high schoolers in California (n = 2,609). Lasso penalized regression was used to select the most accurate predictors of suicidal ideation. Ten out of the 17 input predictors were significant including: depressive symptoms; school-based victimization; sexual and gender minority status; lifetime use of alcohol, vapes, and cannabis; breakfast consumption; access to alcohol and other drugs; and parent education level. The study found that a combination of factors across individual, emotional-social, mental-political, and physical-economic domains could be used to predict the individualized risk of experiencing suicidal ideation. I argue that this multi-level wholistic model is more appropriate and useful, especially for Native American youth. The study highlights the need for a more comprehensive understanding of suicide-related behavior among Native American youth and points to new directions in suicide screening.

Introduction

Death by suicide among Native American peoples is an alarming, racialized health disparity (Warne & Lajimodiere, 2015; Sarche & Spicer, 2008). According to the Office of Minority Health, Native Americans young people ages 15 to 24 years old die by suicide at a disproportionately higher rate compared to white youth of the same age (39.1 vs. 17.0 per 100,000; 2021). Within California, suicide-related behavioral health outcomes for Native American youth are also worse than white youth. From 2017-2019, on average 20.0% of Native American high school students reported experiencing suicidal ideation compared to 17.6% of white youth (WestEd, 2020).

Attempting to address this crisis, the US Department of Health and Human Services Substance Abuse and Mental Health Services Administration spent approximately $20,000,000.00 per year on tribal-specific suicide prevention efforts from 2018 to 2023 (SAMHSA, 2023). Despite these efforts, suicide rates have continued to rise over the past decade perpetuating suicide as the leading cause of non-accidental death for Native American young people ages 15 to 24 years old (CDC, 2023; Curtin & Hedegaard, 2019). This shortcoming has urged some scholars to rethink Native American youth suicide by urging a (re)examination of key theoretical constructs to better direct suicide prevention efforts in tribal communities (Ansloos, 2017; 2018; Elliott-Groves, 2018; Wexler & Gone, 2016; White, 2017). These critics of mainstream suicidology argue that the domination of a psycho-centric conceptualization of suicide-related behavior overemphasizes the risk factors related to psychopathologies, maladaptive cognitive schema, and other mentally disordered characteristics (Ansloos, 2018; White, 2017). As such, the invitation to rethink Native American youth suicide emphasizes a

need to examine suicide and suicide-related behavior beyond this narrow psycho-centric focus to be more engaged with social, structural, and cultural dimensions of tribal health.

This paper uses an Indigenous-based theoretical approach to guide a (re)examination of the multi-level, wholistic factors that may influence suicide-related behavior (i.e., suicidal ideation) among Native American young people in California. Indigenous Wholistic Theory constructs health as multifaceted, amalgamating spiritual, emotional, mental, and physical domains (Absolon, 2010). Aligned with tribal constructions of health and wellness, positive health outcomes occur when these domains align with responsive historical, social, political, and economic contexts (See Figure 1). Subsequently, maladaptive health outcomes—like the exhibition of suicidal ideation— can arise from dysfunction influenced by determinants across any of these domains and related contexts. In this paper, Indigenous Wholistic Theory provides a framework to introduce and situate a broader array of factors that may influence suicide-related behavior in a manner culturally relevant to the lived experienced of Native American youth.
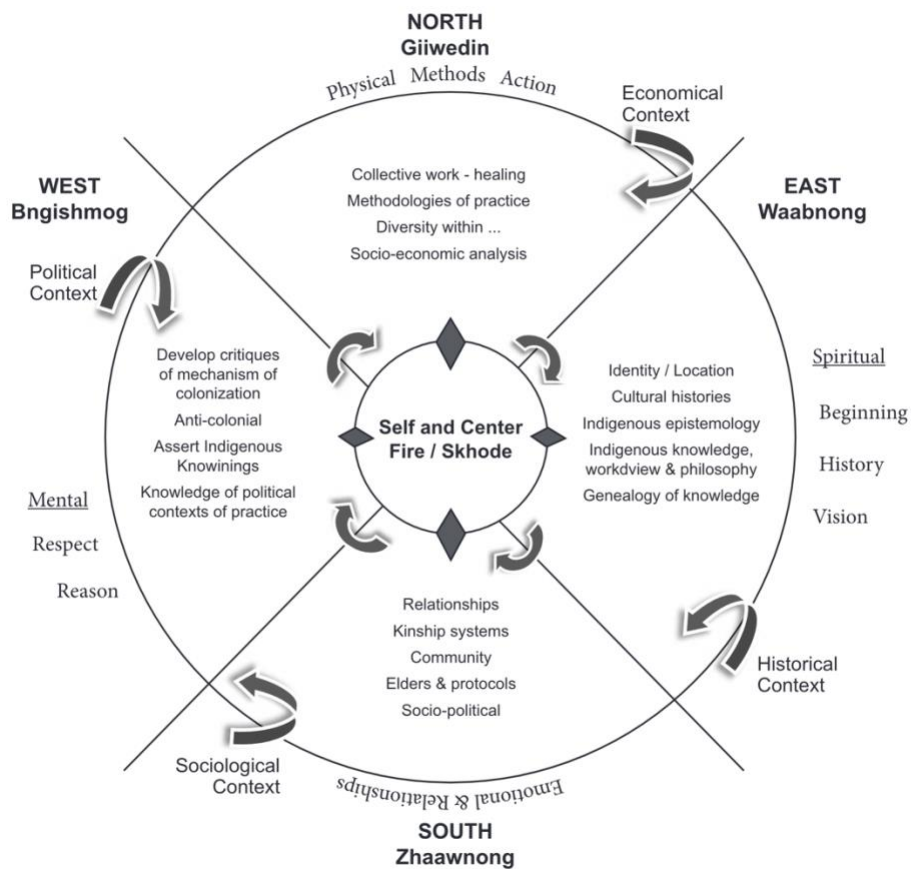
Figure 1 – Indigenous Wholistic Theory Diagram (Absolon, 2010, p. 77)

A computational, algorithmic approach is employed to understand how a more wholistic array of factors from various levels can predict the individualized presence of suicidal ideation among Native American youth in California (Breiman, 2001). This approach diverges from traditional regression methods that inquire about the statistical associations between risk factors and health outcomes at the population level. This paper incorporates prediction modeling techniques for three purposes relevant to Native American youth in California: 1) identify a model comprising a selection of factors that combine to best predict the individualized presence of suicidal ideation; 2) evaluate how the selected model can discriminate between the presence and absence of suicidal ideation; and 3) assess the agreement between the model's predicted

outcomes with the actual outcomes of the input dataset (Brathwaite et al., 2021; Moons et al., 2009). Additionally, this paper leverages a substantiation of the Indigenous Computational Approach to guide its inquiry (Sierra, 2023a). As such, this paper asks the following research question: which combination of factors can predict suicidal ideation among Native American high schoolers in California?

Methods

Indigenous Computational Approach

The Indigenous Computational Approach is a response to the inherent limitations of Western research methodologies that employ computational methods within the domain of social science research. By establishing a framework rooted in Indigenous ways of knowing and being, the Indigenous Computational Approach underscores the imperative of recognizing and valuing Indigenous knowledge systems that center tribal voices and perspectives. Integral to this approach is the development of an Indigenous Statistical Space (Walter & Andersen, 2013), which comprises three essential components: 1) standpoint, which encompasses the situated and reflexive positionality of the researcher within the context of the empirical work; 2) theoretical framework, which provides a conceptual lens through which to interpret and analyze data from an Indigenous perspective; and 3) data analysis technique, which for this approach leverages the use of computer-based algorithms, models, or techniques to analyze, process, and interpret data. Figure 2 is a visual representation of the Indigenous Statistical Space constructed for this paper, culminating in the Indigenous Computational Approach guiding the empirical investigation. The sequential components iteratively combine to foster a research framework that aligns with a

decolonial and Indigenous-centered approach relevant for suicide prevention research with Native American young people.

Figure 2 – Indigenous Statistical Space Diagram adapted from Walter and Andersen (2013, p. 45)

First, the standpoint adopted for this paper draws from the author's worldview as a Yoeme person. In particular, the cultural protocol of "en tui hiapsimake" translated as "with good heart" (Evers & Molina, 1987) guides the author's engagement with suicide prevention research, reflecting the assertion that advocating for the well-being and comprehensive development of Native American young people is fundamentally tied to the larger project of decolonialization. Advocating for the continuity of Indigenous life, rather than actively (or apathetically) upholding the forces fostering Indigenous death, challenges and aims to dismantle the persisting systems of colonial oppression and violence that continue marginalize Native American communities.

Second, Indigenous Wholistic Theory, as previously mentioned, forms the underpinning theoretical framework. This theoretical framework provides a comprehensive lens through which to situate and analyze data, acknowledging the interconnectedness of various factors from multiple levels impacting the lives of Native American high schoolers in California. Third, this paper leverages machine learning to structure a computational, algorithmic analysis of the collected data. The data analysis technique is further described in this section.

Study Setting and Study Participants

The California Healthy Kids Survey (CHKS) is a school-based, multi-modular survey administered annually to students across California (CA Dept. of Education, 2022; CalSCHLS, n.d.). The survey asks questions examining different aspects concerning youth health outcomes and school climate using a variety of different Likert-scaled, multiple-choice questions. CHKS is comprised of its "Core Module" which is administered to all participants. School districts have the option to opt-in to various other modules tailored to measure additional aspects and features. Data was drawn from self-identified Native American high schoolers in grades nine through twelve from the 2019-2020 school year. For analysis, participants who completed all questions from the "Core Module" were included (n = 2,609).

Measure of Suicidal Ideation

The presence of suicidal ideation was measured using a self-reported response. Participants were asked "During the last 12 months, did you ever seriously consider attempting suicide?" Responses were dichotomously coded (0 = No; 1 = Yes). The use of this self-reported measure mirrors other empirical inquiries on suicidal ideation among young people in California from the CHKS dataset (Cederbaum et al., 2014; Perez-Brumer et al., 2017).

Predictors of Suicidal Ideation

The selected predictors include factors previously associated with suicide-related behavior among Native American youth; most have been included in screening tools, but subsequent analyses rely on a singular, psychocentric conceptualization of suicide (Espel, 2015). These factors were motivated and (re)oriented by the integration of Indigenous Wholistic Theory across the following domains: 1) individual; 2) spiritual-historical; 3) emotional-social; 4)

mental-political; 5) physical-economic (See Table 1). For the individual domain, demographic

factors included: gender (Borowsky et al., 1999); gender minority status and sexual minority

status (The Trevor Project, 2020). For this same domain, behavioral factors included: academic

performance (Bolton et al., 2014; Kaplan et al., 2013); and substance use stratified according to

lifetime use of alcohol, tobaccos and vapes, cannabis and use of other drugs (Balis & Postolache,

2010; Bolton et al., 2014; Cwik et al., 2015; Manzo et al., 2015; Pavkov et al., 2010; Yoder et

al., 2006).

Potential Predictors included in the Model		Categories and Coding for Categorical variables / Descriptions of continuous variables	n (%) / [Range] Median (IQR)
Individual (Demographic)			
1	Gender	Male = 0	1,288 (49.31%)
		Female =1	1,321 (50.63%)
2	Gender Minority Status	Cisgender = 0	2,532 (97.05%)
		Transgender = 1	77 (2.95%)
3	Sexual Minority Status	Heterosexual = 0	2,166 (83.02%)
		Non-Heterosexual = 1	443 (16.98%)
Individual (Behavioral)			
4	Academic Performance	Higher scores represent greater academic performance	[-2.54 to 1.30] 0.20 (-0.34, 0.75)
5	Lifetime use of Alcohol	Higher scores represent higher frequency of substance use	[-0.52 to 2.46] -0.52 (-0.52, 0.67)
6	Lifetime use of Cigarettes	Higher scores represent higher frequency of substance use	[-0.23 to 5.31] -0.23 (-0.23, -0.23)
7	Lifetime use of Vapes	Higher scores represent higher frequency of substance use	[-0.46 to 2.70] -0.46 (-0.46, 0.17)
8	Lifetime use of Cannabis	Higher scores represent higher frequency of substance use	[-0.49 to 2.42] -0.49 (-0.49, 0.10)
9	Lifetime use of other Drugs	Higher scores represent higher frequency of substance use	[-0.26 to 6.47] -0.26 (-0.26, -0.26)
Emotional-Social			
10	Foster Care Placement	No = 0	2,600 (99.66%)
		Yes = 1	9 (0.34%)
11	Parent Education Level	Higher scores represent higher level of education	[-1.25 to 1.30] -0.40 (-1.25, 1.30)

12	Positive Substance Use Education	Higher scores represent higher level of internalized positive substance use education	[-1.81 to 1.03] 0.32 (-0.27, 0.91)
13	School-Based Victimization	Higher scores represent higher frequency of victimization	[-0.64 to 4.48] -0.43 (-0.64, 0.17)
14	Access to Drugs and Alcohol	Higher score represents greater ease of access to drugs and/or alcohol	[-2.24 to 1.00] 0.19 (-0.35, 1.00)
Mental-Political			
15	Depressive Symptoms	No = 0	1,566 (60.02%)
		Yes = 1	1,043 (39.98%)
Physical-Economic			
16	Homeless	No = 0	2,543 (97.47%)
		Yes = 1	66 (2.53%)
17	Breakfast Consumption	No = 0	1,175 (45.04%)
		Yes = 1	1,434 (54.96%)

Table 1 - Distribution of Predictors in Sample

For the spiritual-historical domain, no factors were included, because there were no suitable questions included in CHKS. For the emotional-social domain, factors included: foster care placement (Taussig et al., 2014) and school-based victimization (Pavkov et al., 2010). Additional factors for this domain included parent education level, internalization of positive substance use education, and access to alcohol and other drugs. For the mental-political domain, factors included depressive symptoms (e.g., feeling sad or hopeless) (Balis & Postolache, 2010; Cwik et al., 2015; Yoder et al., 2006). For the physical-economic, factors included: homelessness (Holleran & Poon, 2018) and breakfast consumption.

The construction of the selected predictors was motivated by the Indigenous Computational Approach to construct an Indigenous statistical space better reflective of Native American conceptions of health and wellness (Sierra, 2023a). As previously mentioned, Indigenous Wholistic Theory guided this process. The Indigenous Computational Approach is

built on the understanding that Native American health outcomes can be better understood with an alignment linking Indigenous-centered theories with Indigenous research methodologies to guide purposeful data analysis. Please review supplemental materials (Appendix A) for a more detailed description concerning the construction of the selected predictors.

Ethics and Data Transparency

This study received an Institutional Review Board exemption from the University of California, Berkeley Office for the Protection of Human Subjects noting the study's secondary use of data. Reproducible data documentation and code scripts can be found online (See: Sierra, 2023b).

Statistical Analysis

The data analysis was performed using STATA version 17.0. Descriptive statistics were reported in terms of numbers and percentages to depict the sample characteristics and the distribution of categorical predictors. Continuous predictors were standardized to have a mean of zero and a standard deviation of one, and the analyses were reported using median and interquartile ranges. To select the optimal predictors of suicidal ideation, least absolute shrinkage and selection operator (lasso) penalized regression was employed. Lasso regression shrinks the coefficients of the least influential variables to zero, effectively removing them from the model to identify the best subset of predictors (Hastie et al., 2016).

The selection of the tuning parameter lambda was accomplished using 10-fold cross-validation (Hastie et al., 2016). This approach determines both the level of coefficient shrinkage and provides a reliable estimate of the predictive performance of the final model for new cases. The use of ten-fold cross-validation is preferred over split sample validation to avoid

overestimating the predictive performance in unseen cases (Steyerberg & Vergouwe, 2014). To evaluate the ability of the model to distinguish between the presence and absence of suicidal ideation, the area under the curve (AUC) was estimated within the discriminative quality thresholds ranging from 50% indicating an inability to discriminate between individuals with the outcome or not; 70% to 80% indicating acceptable; 80% to 90% indicating excellent; >90% indicating outstanding performance (Hosmer et al., 2013). Additionally, Bootstrapped Bias-Corrected 95% Confidence Intervals (BC 95% CI) for the AUC were calculated.

The agreement between the observed rates of suicidal ideation and the model predictions was assessed using a GiViTI calibration belt plot (Nattino et al., 2017). This allowed for the visual identification of deviations or miscalibrations in observed frequencies compared to expected probabilities at certain confidence levels. This included a calibration test to assess whether any deviations from the bisector were significant (Fenlon et al., 2018; Finazzi et al., 2011). Finally, the overall performance of the model was assessed by estimating the Brier score, which is a measure of the discrepancy between the predicted and observed outcomes using the standard benchmark of 0 for no disagreement, 0.25 for predictions no better than chance (50/50), and 1 for complete disagreement (Brier, 1950).

Sensitivity Analysis

Two sensitivity analyses were performed to evaluate the clustered nature of the data, as well as to assess the impact of incomplete data on the model results. First, the AUC was re-estimated using bootstrap-resampling to account for clustering by schools. Second, the expectation-maximization (EM) algorithm was employed to impute missing values and the predictive model was refitted to the imputed data using the same lasso method (Dempster et al.,

12

1977; Little & Rubin, 2020). The performance of the model was then re-evaluated using AUC and calibration statistics as described above for the main model.

Results

Descriptive Statistics

The majority of the respondents reported being in 9th and 11th grade with 42.93% reported being in 9th grade, 6.32% in 10th grade, 45.00% in 11th grade, and 5.75% in 12th grade. For gender and sexuality, the majority of students reported being cisgender (97.05%) with 50.63% of respondents self-identified as female, and 16.98% self-identifying as non-heterosexual. For the additional categorical predictors, 0.34% of respondents reported being in foster care with 2.53% reported being homeless. For breakfast consumption, 45.04% of the respondents reported not eating breakfast the day they completed their respective survey. Additionally, 38.98% of respondents disclosed experiencing a depressive symptom in the past 12 months.

Prevalence of Suicidal Ideation and Predictors Retained

In the sample group, 16.79% of Native American high school students reported experiencing suicidal ideation in the past 12 months. Using lasso regression with a mean lambda of 0.0078412, the model identified that ten out of the 17 factors were significant in predicting which Native American high schoolers would report experiencing suicidal ideation (see Table 2). These predictors included: depressive symptoms; school-based victimization; sexual and gender minority status; lifetime use of alcohol, vapes, and cannabis; breakfast consumption; access to alcohol and other drugs; and parent education level.

Predictors		Model derived using lasso regression
Individual (Demographic)		
1	Gender	
	Male	x
	Female	x
2	Gender Minority Status	
	Cisgender	-0.01
	Transgender	x
3	Sexual Minority Status	
	Heterosexual	-0.34
	Non-Heterosexual	x
Individual (Behavioral)		
4	Academic Performance	x
5	Lifetime use of Alcohol	0.07
6	Lifetime use of Cigarettes	x
7	Lifetime use of Vapes	0.10
8	Lifetime use of Cannabis	0.02
9	Lifetime use of other Drugs	x
Emotional-Social		
10	Foster Care Placement	
	No	x
	Yes	x
11	Parent Education Level	0.00
12	Positive Substance Use Education	x
13	School-Based Victimization	0.35
14	Access to Drugs and Alcohol	0.08
Mental-Political		
15	Depressive Symptoms	
	No	-1.12
	Yes	x
Physical-Economic		
16	Homeless	
	No	x
	Yes	x
17	Breakfast Consumption	
	No	0.09
	Yes	x
Intercept		-2.28
Total number of predictors retained from the 17 in the models		10
AUC (BC 95% CI) derived using 10-fold cross-validation		0.87 (0.84, 0.88)
AUC (BC 95% CI) derived using 10-fold cross-validation which adjusted for clustering by school		0.87 (0.84, 0.88)
Brier score		0.10

Discrimination and Distribution of Risk Scores

After conducting a 10-fold cross-validation to evaluate internal validation, the final

model (as shown in Figure 3) achieved a cross-validated mean AUC of 87.36 (BC 95% CI: 84.48

to 88.34). This indicates that if a Native American high school student were chosen at random,

there would be an 87% chance that the model would accurately assign a higher risk score to a

student experiencing suicidal ideation compared to one not experiencing ideation. The model's

predicted risk scores ranged from 1.92% to 93.63%, with the median predicted risk score among

the sample being 4.83% (IQR: 2.64 to 26.84). The Brier score was calculated to be 0.10

signifying no disagreement between the predicted and observed outcomes using the standard
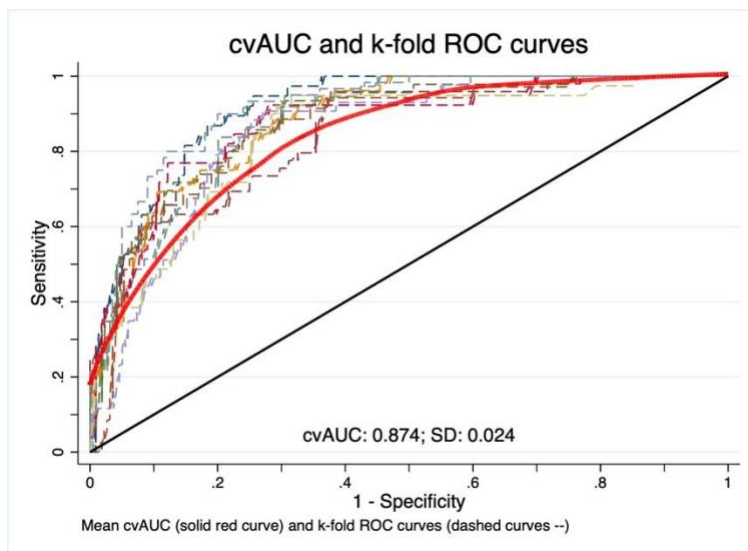
benchmark.

Figure 3 - Model discrimination: Receiver Operator Characteristic (ROC) Curve Showing Mean Cross-Validated Area Under the Curve (AUC)

Calibration

The calibration belt plot indicated that there were some ranges of miscalibration at the 95% and 99% confidence levels (Figure 4). In logistic regression, miscalibration refers to a situation where the predicted probabilities of a model do not correspond well with the observed outcomes. In other words, the model may predict a higher probability of an event occurring than the actual probability, or vice versa. This is a form of bias in the model and can lead to inaccurate predictions. For the 95% confidence level, miscalibration occurred under the bisector for predicted values ranging from 0.02 to 0.06 as seen with the inner belt (light gray area). For the 99% confidence level, miscalibration also occurred under the bisector for predicted values ranging from 0.02 to 0.04 as seen with the outer belt (dark gray area). The corresponding p-value was significant (test statistic = 9.49; p = 0.002) further indicating the evidence of miscalibration. For low-risk students, both belts are under the bisector where the observed presence of suicidal ideation is lower than expected indicating a potential for type 1 errors or false positives. For the purposes of preventing suicide, false positives are preferred to the alternative as students who are experiencing any degree of risk, but may not experiencing suicidal ideation, can still be identified and connected to the appropriate services.
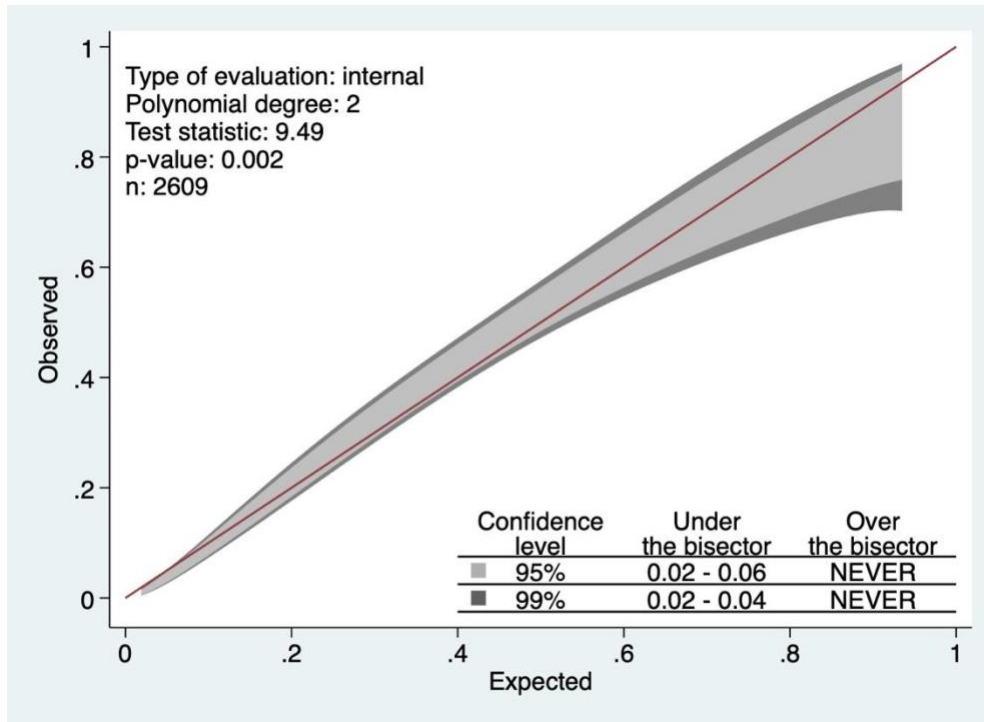
Figure 4 - Calibration Belt Showing Deviations from the Bisector (45% line of Perfect Fit) at the 95% (Inner Belt: Light Grey Area) and 99% (Outer Belt: Dark Grey Area) Confidence Levels

Sensitivity Analysis

Overall, the model discriminations did not show much difference when AUCs were derived using bootstrap-resampling to account for clustering by schools (See Table 2). This indicates that the clustered nature of the data did not have a significant impact on the predictive accuracy of the model (AUC = 87.08.4; BC 95% CI [84.22, 88.14]). In the model that imputed missing values, the AUC was consistent with the previous models (AUC = 88.02; BC 95% CI [87.83, 88.34]), however there was more evidence of miscalibration (See Appendix B). In the sensitivity analysis accounting for missing data, the ten previously identified predictors (depressive symptoms, school-based victimization, sexual and gender minority status, lifetime use of alcohol, vapes, and cannabis, breakfast consumption, access to alcohol and other drugs,

and parent education level) were retained as important for predicting suicidal ideation among Native American high schoolers. However, in the sensitivity analysis that accounted for clustering by schools, only nine of the ten predictors were retained— in this case, the predictor of lifetime cannabis was omitted. This suggests that when controlling for the variance within individual schools, the importance of lifetime cannabis use as a predictor for suicidal ideation can be challenged. Nevertheless, for the final model, lifetime use of cannabis was retained as a predictor due to the strong predictive ability in the original model as designated by the AUC.

Discussion

Suicidal ideation among Native American youth is a complex and multifaceted issue that requires a comprehensive understanding of the risk factors that contribute to suicide-related behavior. Unfortunately, traditional psycho-centric models of suicide have been inadequate in explaining the nuanced experiences of Native American youth, who often face unique cultural, historical, and social challenges that can exacerbate their risk for suicide. Therefore, this study represents an important preliminary step in identifying factors that are relevant in (re)conceptualizing suicide-related behavior among Native American youth and moving from a psycho-centric model to a multi-level wholistic model. By identifying potentially relevant predictors and their relative predictive value, this study can help inform future research directions that seek to better understand the individualized risk of suicidal ideation across theoretical, conceptual, and practical areas. This research can eventually lead to better assessment of risk and more effective prevention.

The study found that a combination of factors, including: individual (such as sexual and gender minority status, lifetime use of alcohol, vapes, and cannabis); emotional-social

(school-based victimization, access to alcohol and other drugs, and parent education level); mental-political (depressive symptoms); and physical-economic (breakfast consumption) domains, could be used to predict the individualized risk of experiencing suicidal ideation among Native American high schools in California. The multiplicity of these domains disrupts the domination of a psycho-centric conceptualization of suicide by substantiating a complex network of factors relevant for predicting suicidal ideation beyond a singular psychological conception. Despite some evidence of miscalibration, the final model had a good overall performance with the 95% CI (84.48 to 88.34) indicating that the true discriminative ability, or AUC, of the model is acceptable. Therefore, the analysis offers a promising approach as a modeling tool to determine the individualized risk of suicidal ideation among Native American high schoolers in California.

While population-level analyses are important for understanding trends and differences among sub-groups, gaining a culturally relevant understanding of suicidal ideation among Native American high school students requires a (re)examination of factors that are relevant to their lived experiences. By identifying the adolescents who are at risk for suicide before its fatal outcome can occur, practitioners in the mental health, education, and social work fields can intervene early and prevent suicide. The resulting individualized risk score from the model can help practitioners identify those with the highest predicted risk and connect them with appropriate suicide prevention efforts. These efforts can be culturally relevant and may include programs such as Honoring the Children-Mending the Circle trauma-focused cognitive behavioral therapy (BigFoot & Schmidt, 2010), American Indian Life Skills (LaFromboise & Howard-Pitney, 1994; 1995), and Gathering of Native Americans (Kraus et al., 2017).

19

This study aimed to examine a set of predictors for suicide-related behavior among Native American youth using available data. However, the findings of this study should be interpreted with caution given its limitations. One major limitation of this study is that it relied on non-validated self-report instruments, which may have limited construct validity and generalizability. Non-validated self-report instruments may not accurately measure the construct of interest, leading to inaccurate conclusions about the variables being studied. Moreover, results from studies that use non-validated self-report instruments may not be generalizable to other populations since the instruments have not been validated across different groups of people. To address these limitations, future studies should incorporate validated, evidence-based instruments to measure suicide risk, such as the Suicide Probability Scale (Cull & Gill, 1982; Hill et al., 2020) or the Suicide Inventory Questionnaire (Keane et al., 1996; LaFromboise & Howard-Pitney, 1994; 1995; Reynolds, 1988), which have been cross-culturally validated for use in Native American contexts (Matsumoto, 2008). Validated, evidence-based instruments to measure the selected predictors should also be incorporated.

Another limitation of this study is the need for further expansion and validation work to optimize its accuracy before deploying it in a clinical setting to improve clinical decisions. For example, the model's transportability to Native American youth in different settings outside of California should be evaluated via external validation. The model's development sample comprised participants from a non-randomized study, which may not be generalizable to Native American young people not in a public high school setting. Future studies may also benefit from incorporating a longitudinal design of the data to substantiate temporal precedence beyond the single academic year evaluated in this study.

Furthermore, a limitation of the study is that it relied on existing data from a survey whose questions may not fully capture the relevant factors. Further research should investigate whether the model's predictive value could be improved by adding other factors such as predictors in the spiritual-historical domain, which is absent from CHKS. One example includes the Historical Loss Scale (Whitbeck et al., 2004), which is a validated instrument designed to measure historical trauma.

Before using the risk scores in practice, the clinical utility of the model and its feasibility and acceptability need to be measured. Although a factor may be statistically significantly associated with suicidal ideation, as seen in previous research, it may not necessarily have predictive power (Shmueli, 2010) and not contribute to predicting suicidal ideation in new cases as indicated by the subset of predictors excluded after lasso regression. As such, even though foster care placement was not retained as a predictor in the final model, future studies should still consider this risk factor in (re)conceptualizing suicide-related behavior.

Despite these limitations, the study's main analyses comprised 438 participants with the outcome (people who reported suicidal ideation) and 10 potential predictors, which conforms to the minimum of 10 events per predictor variable required for reliable prediction modeling (Harrell, 2001; Steyerberg et al., 2001). Moreover, a robust 10-fold cross-validation was utilized for internal validation, which produces less biased estimates than split-sample validation, resulting in more confident predictions (Steyerberg & Vergouwe, 2014).

Conclusion

In conclusion, the study represents an important preliminary step in identifying a broader set of factors that are relevant in (re)conceptualizing suicide-related behavior among

Native American youth beyond a psycho-centric model. The multiplicity of significant factors found in this study disrupts the domination of a psycho-centric conceptualization of suicide and highlights the need for a culturally relevant understanding of suicidal ideation among Native American high school students. While this study has some limitations, it offers a promising approach as a modeling tool to determine the individualized risk of suicidal ideation among Native American high schoolers in California. By identifying those at risk, practitioners can intervene early and prevent suicide through culture-based programs. Further validation work and clinical utility testing are necessary before deploying the model in a clinical setting. Ultimately, the study highlights the need for a more comprehensive understanding of suicide-related behavior among Native American youth and the importance of developing evidence-based suicide prevention strategies that incorporate culture-based theoretical perspectives.

References

Absolon, K. (2010). Indigenous Wholistic Theory: A Knowledge Set for Practice. *First Peoples Child and Family Review*, *5*(2), 14.

Ansloos, J. (2018). Rethinking Indigenous Suicide. *International Journal of Indigenous Health*, *13*(2), 8–28. https://doi.org/10.32799/ijih.v13i2.32061

Balis, T., & Postolache, T. T. (2010). *Ethnic Differences in Adolescent Suicide in the United States*.

BigFoot, D. S., & Schmidt, S. R. (2010). Honoring Children, Mending the Circle: Cultural Adaptation of Trauma-Focused Cognitive-Behavioral Therapy for American Indian and Alaska Native Children. *Journal of Clinical Psychology*, *66*(8), 847–856. https://doi.org/10.1002/jclp.20707

Bolton, S.-L., Elias, B., Enns, M. W., Sareen, J., Beals, J., Novins, D. K., The Swampy Cree Suicide Prevention Team, & The AI-SUPERPFP Team. (2014). A Comparison of the Prevalence and Risk Factors of Suicidal Ideation and Suicide Attempts in Two American Indian Population Samples and in a General Population Sample. *Transcultural Psychiatry*, *51*(1), 3–22. https://doi.org/10.1177/1363461513502574

Borowsky, I. W., Resnick, M. D., Ireland, M., & Blum, R. W. (1999). Suicide Attempts Among American Indian and Alaska Native Youth: Risk and Protective Factors. *Archives of Pediatrics & Adolescent Medicine*, *153*(6). https://doi.org/10.1001/archpedi.153.6.573

Brathwaite, R., Ssewamala, F. M., Neilands, T. B., Okumu, M., Mutumba, M., Damulira, C., Nabunya, P., Kizito, S., Sensoy Bahar, O., Mellins, C. A., & McKay, M. M. (2021). Predicting the Individualized Risk of Poor Adherence to Art Medication Among Adolescents Living with Hiv in Uganda: The Suubi+adherence Study. *Journal of the International AIDS Society*, *24*(6). https://doi.org/10.1002/jia2.25756

Breiman, L. (2001). Statistical Modeling: The Two Cultures. *Statistical Science*, *16*(3), 199–231.

Brier, G. W. (1950). Verification of Forecasts Expressed in Terms of Probability. *Monthly Weather Review*, *78*(1), 1–3.

CA Dept. of Education. (2022). *California Healthy Kids Survey—Alcohol, Tobacco & Other Drug Prevention*. California Healthy Kids Survey. https://www.cde.ca.gov/ls/he/at/chks.asp

CalSCHLS: California School Climate, Health, and Learning Survey System. (n.d.). *The Surveys*. About. Retrieved April 2, 2023, from https://calschls.org/about/the-surveys/

CDC: US Centers for Disease Control and Prevention. (2023). *10 Leading Causes of Death, United States 2018—2020, Am Indian/AK Native, Both Sexes* [Leading Causes of Death Report]. US Department of Health and Human Services.

Cederbaum, J. A., Gilreath, T. D., Benbenishty, R., Astor, R. A., Pineda, D., DePedro, K. T., Esqueda, M. C., & Atuel, H. (2014). Well-Being and Suicidal Ideation of Secondary School Students from Military Families. *Journal of Adolescent Health*, *54*(6), 672–677. https://doi.org/10.1016/j.jadohealth.2013.09.006

Cull, J. G., & Gill, W. S. (1982). *Suicide Probability Scale*. American Psychological Association.

Curtin, S. C., & Hedegaard, H. (2019). Suicide Rates for Females and Males by Race and Ethnicity: United States, 1999 and 2017. *National Center for Health Statistics (NCHS) Health E-Stat*, 6.

Cwik, M., Barlow, A., Tingey, L., Goklish, N., Larzelere-Hinton, F., Craig, M., & Walkup, J. T. (2015). Exploring Risk and Protective Factors with a Community Sample of American Indian Adolescents Who Attempted Suicide. *Archives of Suicide Research*, *19*(2), 172–189. https://doi.org/10.1080/13811118.2015.1004472

Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum Likelihood from Incomplete

    Data via the EM Algorithm. *Journal of the Royal Statistical Society. Series B*

    (Methodological), *39*(1), 1–38.

Elliott-Groves, E. (2018). Insights from Cowichan: A Hybrid Approach to Understanding

    Suicide in One First Nations' Collective. *Suicide and Life-Threatening Behavior*, *48*(3),

    328–339. https://doi.org/10.1111/sltb.12364

Espel, E. (2015). *American Indian/Alaska Native Adolescent Suicide: Risk Factors, Protective*

    Factors, and Prevention Opportunities in Education Settings: An Annotated Bibliography

    (p. 39) [A report from the West Comprehensive Center]. WestEd.

Evers, L., & Molina, F. S. (1987). *Yaqui Deer Songs/Maso Bwikam: A Native American*

    Poetry. University of Arizona Press.

Fenlon, C., O'Grady, L., Doherty, M. L., & Dunnion, J. (2018). A Discussion of Calibration

    Techniques for Evaluating Binary and Categorical Predictive Models. *Preventive*

    Veterinary Medicine, *149*, 107–114. https://doi.org/10.1016/j.prevetmed.2017.11.018

Finazzi, S., Poole, D., Luciani, D., Cogo, P. E., & Bertolini, G. (2011). Calibration Belt for

    Quality-of-Care Assessment Based on Dichotomous Outcomes. *PLOS ONE*, *6*(2),

    e16110. https://doi.org/10.1371/journal.pone.0016110

Harrell, F. E. (2001). *Regression Modeling Strategies: With Applications to Linear Models,*

    Logistic Regression, and Survival Analysis (pp. xxii–xxii). Springer.

    https://doi.org/10.1007/978-1-4757-3462-1

Hastie, T., Tibshirani, R., & Friedman, J. (2016). *The Elements of Statistical Learning: Data*

    Mining, Inference, and Prediction, Second Edition (2nd edition). Springer.

Hill, K., Van Eck, K., Goklish, N., Larzelere-Hinton, F., & Cwik, M. (2020). Factor Structure and Validity of the SIQ-Jr in a Southwest American Indian Tribe. *Psychological Services*, *17*(2), 207–216. https://doi.org/10.1037/ser0000298

Holleran, L., & Poon, G. (2018). Dying in the Shadows: Suicide Among the Homeless. *Harvard Public Health Review (Cambridge, Mass.)*, *20*, 1–5.

Hosmer, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied Logistic Regression* (Third edition). Wiley.

Kaplan, M. S., McFarland, B. H., Huguet, N., Conner, K., Caetano, R., Giesbrecht, N., & Nolte, K. B. (2013). Acute Alcohol Intoxication and Suicide: A Gender-Stratified Analysis of the National Violent Death Reporting System. *Injury Prevention*, *19*(1), 38–43. https://doi.org/10.1136/injuryprev-2012-040317

Keane, E. M., Dick, R. W., Bechtold, D. W., & Manson, S. M. (1996). Predictive and Concurrent Validity of the Suicidal Ideation Questionnaire Among American Indian Adolescents. *Journal of Abnormal Child Psychology*, *24*(6), 735. Periodicals Archive Online; Periodicals Index Online.

Kraus, C., Bartgis, J., Lahiff, M., & Auerswald, C. L. (2017). The Gathering of Native Americans Intervention: Cultivating Hope and Meaningful Relationships for Urban American Indian Adolescents in California. *Journal of Adolescent Health*, *60*(2), S1. https://doi.org/10.1016/j.jadohealth.2016.10.024

LaFromboise, T. D., & Howard-Pitney, B. (1994). The Zuni Life Skills Development Curriculum: A Collaborative Approach to Curriculum Development. *American Indian and Alaska Native Mental Health Research*, *Monograph* (4), 98–121. https://doi.org/10.5820/aian.mono04.1994.98

LaFromboise, T. D., & Howard-Pitney, B. (1995). The Zuni Life Skills Development

    Curriculum: Description and Evaluation of a Suicide Prevention Program. *Journal of*

    Counseling Psychology, *42*(4), 479–486.

Little, R. J. A., & Rubin, D. B. (2020). *Statistical Analysis with Missing Data* (Third edition).

    Wiley.

Manzo, K., Tiesman, H., Stewart, J., Hobbs, G. R., & Knox, S. S. (2015). A Comparison of

    Risk Factors Associated with Suicide Ideation/Attempts in American Indian and White

    Youth in Montana. *Archives of Suicide Research*, *19*(1), 89–102.

    https://doi.org/10.1080/13811118.2013.840254

Matsumoto, D. (2008). Cross-Cultural Research. In S. Davis (Ed.), *Handbook of Research*

    Methods in Experimental Psychology. John Wiley & Sons.

Moons, K. G. M., Royston, P., Vergouwe, Y., Grobbee, D. E., & Altman, D. G. (2009).

    Prognosis and Prognostic Research: What, Why, and How? *BMJ*, *338*.

    https://doi.org/10.1136/bmj.b375

Nattino, G., Lemeshow, S., Phillips, G., Finazzi, S., & Bertolini, G. (2017). Assessing the

    Calibration of Dichotomous Outcome Models with the Calibration Belt. *The Stata*

    Journal, *17*(4), 1003–1014. https://doi.org/10.1177/1536867X1801700414

Office of Minority Health. (2021). *Mental and Behavioral Health—American Indians/Alaska*

    Natives (p. 1) [Minority Population Profiles]. US Department of Health and Human

    Services. https://minorityhealth.hhs.gov/omh/browse.aspx?lvl=4&lvlid=39

Pavkov, T. W., Travis, L., Fox, K. A., King, C. B., & Cross, T. L. (2010). Tribal Youth

    Victimization and Delinquency: Analysis of Youth Risk Behavior Surveillance Survey

Data. *Cultural Diversity and Ethnic Minority Psychology*, *16*(2), 123–134.

https://doi.org/10.1037/a0018664

Perez-Brumer, A., Day, J. K., Russell, S. T., & Hatzenbuehler, M. L. (2017). Prevalence and

Correlates of Suicidal Ideation Among Transgender Youth in California: Findings from a

Representative, Population-Based Sample of High School Students. *Journal of the*

American Academy of Child & Adolescent Psychiatry, *56*(9), 739–746.

https://doi.org/10.1016/j.jaac.2017.06.010

Reynolds, M. W. (1988). *Suicidal Ideation Questionnaire: Professional Manual*.

Psychological Assessment Resources.

SAMHSA: Substance Abuse and Mental Health Services Administration. (2023). *Substance*

Abuse and Mental Health Services Administration Fiscal Year 2023 Justification of

Estimates for Appropriations Committee. US Department of Health and Human Services.

https://www.samhsa.gov/sites/default/files/samhsa-fy-2023-cj.pdf

Sarche, M., & Spicer, P. (2008). Poverty and Health Disparities for American Indian and

Alaska Native Children: Current Knowledge and Future Prospects. *Annals of the New*

York Academy of Sciences, *1136*, 126–136. https://doi.org/10.1196/annals.1425.017

Shmueli, G. (2010). To Explain or to Predict? *Statistical Science*, *25*(3).

https://doi.org/10.1214/10-STS330

Sierra, V. Q. (2023a). *A Guide for Crafting Indigenous Statistical Spaces for Computational*

Research

Sierra, V. Q. (2023b). Indigenous Wholistic Factors Project: Predicting Suicidal Ideation

among Native American High Schoolers in California. Retrieved from

https://osf.io/4dpwt/

Steyerberg, E. W., Eijkemans, M. J. C., Harrell, F. E., & Habbema, J. D. F. (2001). Prognostic

Modeling with Logistic Regression Analysis: In Search of a Sensible Strategy in Small

Data Sets. *Medical Decision Making*, *21*(1), 45–56.

https://doi.org/10.1177/0272989X0102100106

Steyerberg, E. W., & Vergouwe, Y. (2014). Towards Better Clinical Prediction Models: Seven

Steps for Development and an ABCD for Validation. *European Heart Journal*, *35*(29),

1925–1931. https://doi.org/10.1093/eurheartj/ehu207

Taussig, H. N., Harpin, S. B., & Maguire, S. A. (2014). Suicidality Among Preadolescent

Maltreated Children in Foster Care. *Child Maltreatment*, *19*(1), 17–26.

https://doi.org/10.1177/1077559514525503

The Trevor Project. (2020). *The Trevor Project Research Brief: American Indian/Alaskan

Native Youth Suicide Risk* [Research Brief]. The Trevor Project.

https://www.thetrevorproject.org/research-briefs/american-indian-alaskan-native-youth-

suicide-risk/

Walter, M., & Andersen, C. (2013). *Indigenous Statistics: A Quantitative Research

Methodology* (1st edition). Routledge.

Warne, D., & Lajimodiere, D. (2015). American Indian Health Disparities: Psychosocial

Influences. *Social and Personality Psychology Compass*, *9*(10), 567–579.

https://doi.org/10.1111/spc3.12198

WestEd. (2020). California Healthy Kids Survey (CHKS) & Biennial State CHKS. California

Dept. of Education. As cited on www.kidsdata.org, a program of Population Reference

Bureau. Retrieved on [February 1, 2023].

Wexler, L. M., & Gone, J. P. (2012). Culturally Responsive Suicide Prevention in Indigenous

    Communities: Unexamined Assumptions and New Possibilities. *American Journal of*

    Public Health, *102*(5), 800–806. https://doi.org/10.2105/AJPH.2011.300432

Whitbeck, L. B., Adams, G. W., Hoyt, D. R., & Chen, X. (2004). Conceptualizing and

    Measuring Historical Trauma Among American Indian People. *American Journal of*

    Community Psychology, *33*(3–4), 119–130.

    https://doi.org/10.1023/B:AJCP.0000027000.77357.31

White, J. (2017). What Can Critical Suicidology Do? *Death Studies*, *41*(8), 472–480.

    https://doi.org/10.1080/07481187.2017.1332901

Yoder, K. A., Whitbeck, L. B., Hoyt, D. R., & LaFromboise, T. (2006). Suicidal Ideation

    Among American Indian Youths. *Archives of Suicide Research*, *10*(2), 177–190.

    https://doi.org/10.1080/13811110600558240

Appendix A: Construction of Predictor Variables

- Gender (Binary): This variable represents the gender of each respondent. The variable was constructed from question **a4** (What is your gender?) of the Core Module. Raw data was dichotomously coded (0 = Male; 1 = Female).
- Gender Minority Status (Binary): This variable represents the gender minority status of each respondent. Gender minority status describes individuals whose gender identity is different from the sex they were assigned at birth. The variable was constructed from question **a5** (Some people describe themselves as transgender when how they think or feel about their gender is different from the sex they were assigned at birth. Are you transgender?) of the Core Module. Responses were categorically coded (1 - No, I am not transgender; 2 - Yes, I am transgender; 3 - I am not sure if I am transgender; 4 - Decline to respond). Raw data was recoded to create a binary variable, with raw data value 1 being recoded as 0 = cisgender (No), and the raw data values 2 - 4 were recoded as 1 = transgender (Yes).
- Sexual Minority Status (Binary): This variable represents the sexual minority status of each respondent. Sexual minority status described individuals who identify as lesbian, gay, bisexual, queer, or additional non-heterosexual identity. The variable was constructed from question **a18** of the Core Module. Responses were categorically coded (1 - Straight [not gay]; 2 - Gay or Lesbian; 3 – Bisexual; 4 - I am not sure yet; 5 - Something else; 6 - Decline to respond). Raw data was recoded to create a binary variable, with raw data value of 1 being recoded as 0 = heterosexual (No), and the raw data values 2 - 6 were recoded as 1= non-heterosexual (Yes).
- Academic Performance (Continuous): This variable represents the degree of academic performance among each respondent. The variable was constructed from the **a20** question (During the past 12 months, how would you describe the grades you mostly received in school?) from the Core Module. This variable was then standardized so the variable has a mean of zero and a standard deviation of one. Scores ranged from -2.54 to 1.30 with a high score signifying a higher degree of academic performance.
- Lifetime Use of Alcohol (Continuous): This variable represents the frequency a respondent consumes alcohol. The variable was constructed from the **a51_a52** question from the substance use component of the Core Module designed to assess the extent to which participants consumes mind altering substances. The raw data was then standardized so the variable had a mean of zero and a standard deviation of one. Scores ranged from -0.52 to 2.46 with a higher score signifying a higher frequency of substance use.
- Lifetime Use of Cigarettes (Continuous): This variable represents the frequency a respondent engages in smoking cigarettes. The variable was constructed from the **a48_a49** question from the substance use component of the Core Module. The raw data was then standardized so the variable had a mean of zero and a standard deviation of one. Scores ranged from -0.23 to 5.31with a higher score signifying a higher frequency of substance use.
- Lifetime Use of Vapes (Continuous): This variable represents the frequency a respondent engages in smoking via a vaporizer or vape product. The variable was constructed from the **a50_a51** question from the substance use component of the Core Module. The raw

data was then standardized so the variable had a mean of zero and a standard deviation of one. Scores ranged from -0.46 to 2.70 with a higher score signifying a higher frequency of substance use.

- Lifetime Use of Cannabis (Continuous): This variable represents the frequency a respondent consumes cannabis. The variable was constructed from the **a52_a53** question from the substance use component of the Core Module. The raw data was then standardized so the variable had a mean of zero and a standard deviation of one. Scores ranged from -0.49 to 2.42 with a higher score signifying a higher frequency of substance use.

- Lifetime Use of other Drugs (Continuous): This variable represents the frequency a respondent consumes substances not previously mentioned. The variable was constructed by merging responses from 11 questions from the substance use component of the Core Module. A continuous variable was constructed by summing the responses to the 11 questions. The raw data was then divided by the number of responses and standardized so the variable had a mean of zero and a standard deviation of one. Scores ranged from -0.26 to 6.47 with a higher score signifying a higher frequency of substance use.

- Foster Care Placement (Binary): This variable represents if a respondent is currently in a foster care placement. The variable was constructed from question **a9** (What best describes where you live?) of the Core Module. Responses were categorically coded (1- A home with one or more parents/guardian; 2 - Other relative's home; 3- A home with more than one family; 4 - Friend's home; 5 - Foster home, group care, or waiting place; 6 - Hotel or motel; 7 - Shelter, car, campground, or other temporary housing; 8 - Other living arrangement). Raw data was recoded to create a binary variable, with raw data values 1 – 3, and 5 – 8 being recoded as 0 – No, and the raw data value of 4 recoded as 1- Yes.

- Parent Education Level (Continuous): This variable represents the degree of education for each respondents' parents or guardians. The variable was constructed from the **a10** question (What is the highest level of |education your parents or guardians completed?) from the Core Module. The raw data was then standardized so the variable had a mean of zero and a standard deviation of one. Scores ranged from -2.54 to 1.30 with a higher score signifying a higher level of parental education.

- Positive Substance Use Education (Continuous): This variable represents the degree a respondent has received and internalized positive substance use education. The variable was constructed by merging responses from eight questions from the Core Module. A continuous variable was constructed by summing the responses to the eight questions. The raw data was then divided by the number of responses and standardized so the variable had a mean of zero and a standard deviation of one. Scores ranged from -1.81 to 1.03 with a higher score signifying a higher level of internalized education.

- School-Based Victimization (Continuous): This variable represents the frequency a respondent has experienced instances of victimization including bullying, physical or sexual violence, or other processes where an individual is harmed. The variable was constructed by merging responses from 24 questions from the Core Module. A continuous variable was constructed by summing the responses to the 24 questions. The raw data was then divided by the number of responses and standardized so the variable had a mean of zero and a standard deviation of one. Scores ranged from -0.64 to 4.48 with a higher score signifying a higher frequency of experiences of victimization.

- Access to Drugs and Alcohol (Continuous): This variable represents the degree of ease of access a respondent has to obtain drugs and/or alcohol. The variable was constructed by merging responses from four questions from the Core Module. A continuous variable was constructed by summing the responses to the four questions. The raw data was then divided by the number of responses and standardized so the variable had a mean of zero and a standard deviation of one. Scores ranged from -2.24 to 1.00 with a higher score signifying an easier degree of access to obtaining drugs and/or alcohol.
- Depressive Symptoms (Binary): This variable represents the presence of depressive systems for each respondent. The variable was constructed from question **a130_a118** (During the past 12 months, did you ever feel so sad or hopeless almost every day for two weeks or more that you stopped doing some usual activities?) of the Core Module. Raw data was dichotomously coded (0 = No; 1 = Yes).
- Homelessness (Binary): This variable represents if a respondent reported experiencing homelessness or housing insecurity. The variable was constructed from question **a9** (What best describes where you live?) of the Core Module. Responses were categorically coded (1- A home with one or more parents/guardian; 2 - Other relative's home; 3- A home with more than one family; 4 - Friend's home; 5 - Foster home, group care, or waiting place; 6 - Hotel or motel; 7 - Shelter, car, campground, or other temporary housing; 8 - Other living arrangement). Raw data was recoded to create a binary variable, with raw data values 1 – 5 recoded as 0 = No, and the raw data values of 6 - 8 recoded as 1 = Yes.
- Breakfast Consumption (Binary): This variable represents if a respondent consumed breakfast the day the survey was administered. The variable was constructed from question **a132_a120** (Did you eat breakfast today?) of the Core Module. Raw data was dichotomously coded (0 = No; 1 = Yes).
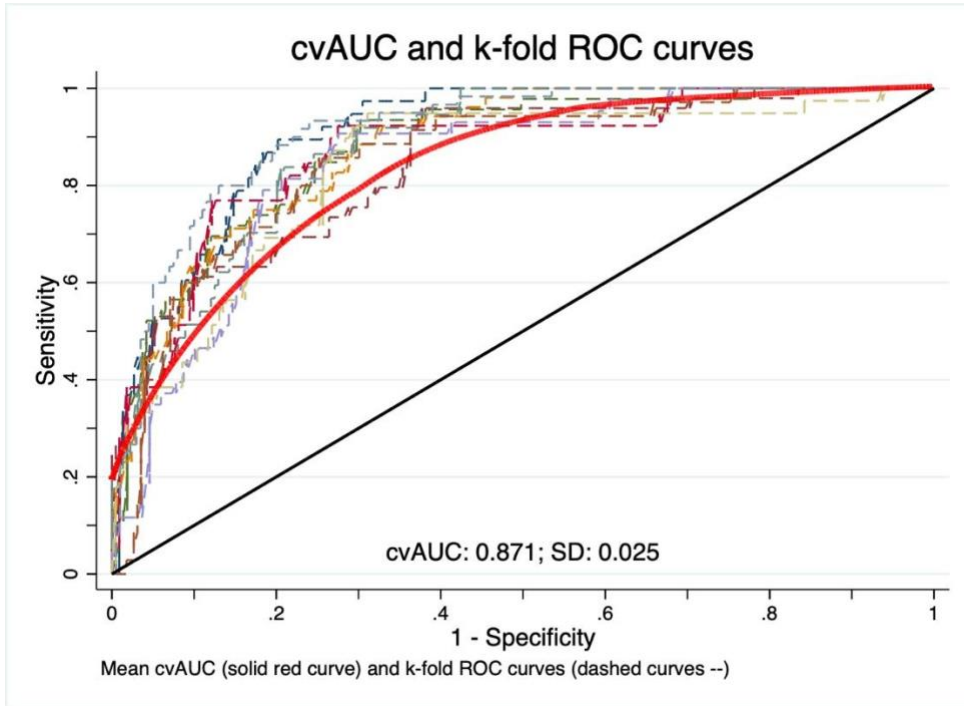
Appendix B: Sensitivity Analysis

Figure 5 - ROC curve showing mean cross-validated area under the curve (AUC) when accounting for clustering by schools
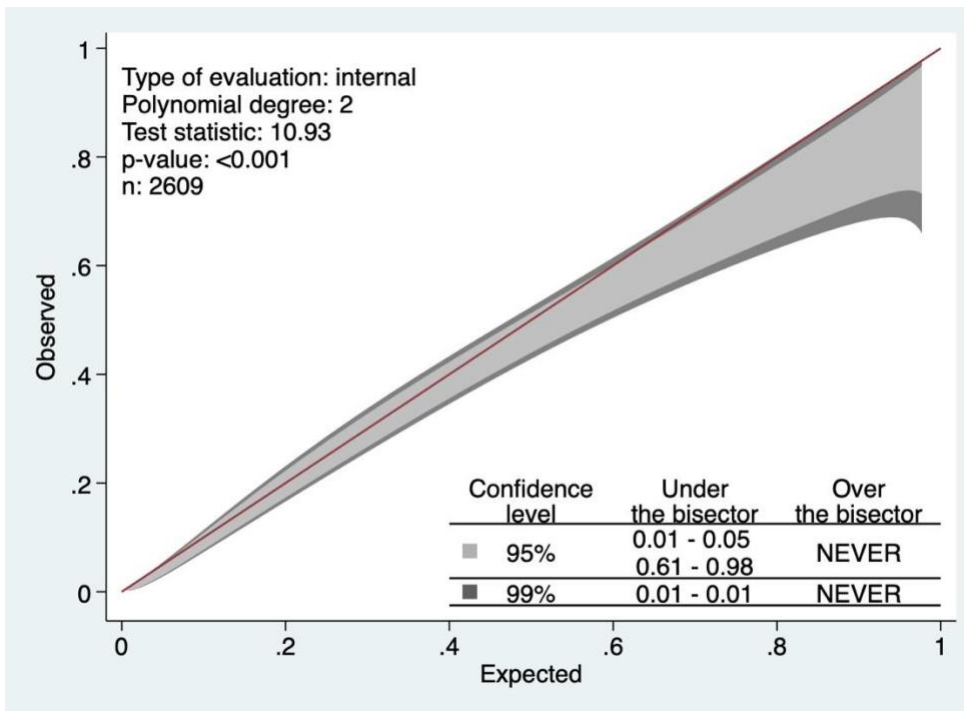

Figure 6 - Calibration belt showing deviations from the bisector (45% line of perfect fit) at the 95% (inner belt: light grey area) and 99% (outer belt: dark grey area) confidence levels when accounting for clustering by schools