THE JOURNAL OF
CLINICAL PSYCHIATRY
THE OFFICIAL JOURNAL OF THE AMERICAN SOCIETY OF CLINICAL PSYCHOPHARMACOLOGY

Supplementary Material

Article Title: Homicide-Suicide in the United States: Moving Toward an Empirically Derived Typology

Authors: Joshua T. Jordan, PhD, and Dale E. McNiel, PhD

DOI Number: 10.4088/JCP.20m13528

List of Supplementary Material for the article

Disclaimer

This Supplementary Material has been provided by the author(s) as an enhancement to the published article. It has been approved by peer review; however, it has undergone neither editing nor formatting by in-house editorial staff. The material is presented in the manner supplied by the author.

Supplemental Materials

Methods

Appendix 1. States contributing data to the NVDRS. Seven states began contributing data in 2003 (Alaska, Maryland, Massachusetts, New Jersey, Oregon, South Carolina, Virginia), five in 2004 (Colorado, Georgia, North Carolina, Oklahoma, Rhode Island), three began in 2005 (Kentucky, New Mexico, Utah), one each in 2006, 2011, and 2014 (Wisconsin, Ohio, and Michigan, respectively), and nine began contributing data in 2015 (Arizona, Connecticut, Hawaii, Kansas, Maine, Minnesota, New Hampshire, New York, and Vermont).

Data Analysis

Appendix 2. Determination of final latent class solution. For all latent class models, 100 random starts were used to avoid local solutions, and 50 were optimized. For each latent class solution *k*, statistical model fit was assessed by approximate fit indices and likelihood ratio tests. Approximate fit indices for consideration included the Bayesian Information Criterion (BIC), Sample-size Adjusted Bayesian Information Criterion (SABIC), and Akaike Information Criterion (AIC), where lower values indicate improved fit over the $k – 1$ class solution (e.g., a three- vs. two-class solution); likelihood ratio tests included the adjusted Lo-Mendell-Rubin (LMR) and bootstrap likelihood ratio test (BLRT), where significant *p*-values indicate model fit that is superior to the $k – 1$ class solution. As each model was estimated, the entropy index was used to assess the overall clarity of classification of homicide-suicide suspects into classes. Values $\geq$ .80 indicate “good” classification[1]. The Average Posterior Probability (AvePP) was also used to estimate how well a model classifies observations into their most likely class, where values > .70 indicate well-separated classes[2]. Model solutions were then evaluated based on theoretical and conceptual meaningfulness, by an examination of qualitative differences among the latent classes' conditional response probabilities – on all characteristics specified in the LCA[3].

Appendix 3. Predicting homicide-suicide group membership. To examine whether demographic and other characteristics distinguish specific homicide-suicide classes from all other classes, a series of least absolute shrinkage and selection operator (lasso) logistic regressions were used. This approach was chosen to identify the most salient features of each homicide-suicide class. In short, the lasso is a machine learning model that aims to identify relevant covariates of an outcome in a manner that reduces overfitting, thus enhancing generalizability to external samples. It is preferable to traditional stepwise regression models, as it addresses issues of multicollinearity and prevents overfitting through a penalization parameter (λ) that shrinks coefficients towards zero, thereby leaving the "best" predictors in a model[4]. The optimal λ for each model was obtained via 10-fold cross-validation. For models where there was significant class imbalance (e.g., a given class represented < 30% of the sample), the Synthetic Minority Oversampling TEchnique (SMOTE[5]) was used via the *DMwR* package[6] in R[7].[1] Models were then refitted using the optimal λ to evaluate variable importance, as well as model fit as measured by the Area Under the Curve (AUC). Due to relatively small sample sizes from the resulting latent classes, we elected to use a bias-corrected and

[1] Severe class imbalance can bias models to favor the majority class. Unlike traditional oversampling methods, which essentially duplicates observations in the minority class to achieve balanced classes, SMOTE creates new (non-duplicate) minority cases using the *k*-nearest neighbors of the pre-existing minority cases, and simultaneously under-samples from the majority class. SMOTE is considered the "de facto" standard for imbalanced datasets[9].

accelerated bootstrap with 5,000 replications of the data to characterize model uncertainty, rather than using a holdout sample[8].

References

1. Nylund-Gibson K, Choi AY. Ten frequently asked questions about latent class analysis. *Transl Issues Psychol Sci*. 2018;4(4):440-461.
2. Nagin DS. *Group-Based Modeling of Development*. Cambridge, MA: Harvard University Press; 2005.
3. Masyn KE. Latent class analysis and finite mixture modeling. In: Little TD, ed. *The Oxford Handbook of Quantitative Methods*. 2nd ed. New York, NY: Oxford University Press; 2013:551-611.
4. Tibshirani R. Regression shrinkage and selection via the lasso. *J R Stat Soc Ser B*. 1996;58:267-288.
5. Chawla N, Bowyer K, Hall L, Kegelmeyer W. SMOTE: Synthetic minority over-sampling technique. *J Artif Intell Res*. 2002;16:321-357. doi:https://doi.org/10.1613/jair.953
6. Torgo L. *Data Mining with R, Learning with Case Studies*. Boca-Raton, FL: Chapman & Hall/CRC; 2010.
7. R Core Team. A language and environment for statistical computing. 2018.
8. Smith D, Wang S, Carter M, Fox K, Hooley J. Longitudinal predictors of self-injurious thoughts and behaviors in sexual and gender minority adolescents. *J Abnorm Psychol*. 2019;Epub ahead. doi:http://dx.doi.org/10.1037/abn0000483
9. Fernández A, García S, Herrera F, Chawla N. SMOTE for learning from imbalanced data: Progress and challenges, marking the 15-year anniversary. *J Artif Intell Res*. 2018;61:863-905.

Supplementary Table 1

Lasso coefficients predicting group membership

Characteristic	Latent class groups[a]							
	Class 1 (*n* = 134)	Class 2 (*n* = 277)	Class 3 (*n* = 145)	Class 4 (*n* = 96)	Class 5 (*n* = 179)	Class 6 (*n* = 89)	Class 7 (*n* = 160)	Class 8 (*n* = 1,367)
Demographic								
Male Gender	-1.02	.16	.18	.15	.13	.17	--	.23
Race								
White	.18	.21	.24	--	--	--	.22	-.23
African-American	-.02	-.20	-.19	--	-.18	--	-.26	.16
Hispanic	--	-.16	-.14	--	--	--	-.21	--
Other	--	--	--	--	-.18	--	--	--
Age								
18-29	--	.16	-.14	--	--	--	--	--
30-39	.12	--	--	.41	--	--	-.08	-.01
40-49	--	-.04	--	.52	.31	--	.14	--
50-59	--	.05	.26	--	--	--	--	.03
60-69	-.18	.08	.79	--	.13	--	--	-.09
70 and Older	-.40	-.30	1.53	-.18	--	--	-.14	-.18
Marital Status								
Married	.22	-.49	.18	.37	--	-.15	--	-.15
Never Married/Single	-.18	--	--	-.31	-.05	--	.50	-.36
Widowed/Divorced/Separated	--	-.57	--	--	--	.26	-.23	.11
Served in U.S. Military	.05	.20	--	-.14	--	--	-.19	-.10
Other Characteristics								
Disclosed Suicide Intent	.05	-.16	.04	--	.24	--	--	-.17
Suicide Note	.26	-.28	.40	.09	--	-.39	-.12	-.14
Location of Death at Residence	.27	-.77	.58	.42	.19	-.38	--	.07
Alcohol use suspected	-.25	-.21	-.18	--	.13	--	.03	-.02
Used firearm	-.10	.14	--	-.09	--	--	.22	-.02
History of suicide attempt	--	.05	--	--	.01	--	--	-.08
History of mental health treatment	--	.34	--	--	.90	.20	.18	-1.07
Current mental health treatment	-.02	-.34	--	--	--	--	--	-.41
Area Under the Curve (95% CI)[b]	.81 (.79, .83)	.80 (.78, .82)	.91 (.90, .92)	.75 (.73, .77)	.78 (.76, .80)	.68 (.66, .70)	.73 (.71, .75)	.76 (.74, .78)

Note. Coefficients can be interpreted in the same manner as regular regression coefficients, where positive values indicate presence of the characteristic is associated with a greater likelihood of being in the latent class, and negative values indicate presence of the characteristic is associated with a lower likelihood of being in the latent class. [a]Class 1: Filicide Type, Class 2: Extrafamilial Type, Class 3: IP-Physical Health Type, Class 4: Familicide Type, Class 5: IP-Distress Type, Class 6: Indiscriminate/Rage Type, Class 7: Other Family Type, Class 8: IP-Relational Type. [b]Area Under the Curve estimates derived from 5,000 bias-corrected and accelerated bootstrapped replications.