Psychological Medicine

cambridge.org/psm

Original Article

Cite this article: Levis M, Dimambro M, Levy J, Dufort V, Fraade A, Winer M, Shiner B (2024). Characterizing Veteran suicide decedents that were not classified as high-suicide-risk. *Psychological Medicine* **54**, 3135–3144. https://doi.org/10.1017/S0033291724001296

Received: 11 December 2023
Revised: 18 March 2024
Accepted: 9 May 2024
First published online: 16 September 2024

Keywords:
machine learning bias; machine learning; suicide prevention; veterans and military

Corresponding author:
Maxwell Levis;
Email: maxwelle.levis@va.gov

Characterizing Veteran suicide decedents that were not classified as high-suicide-risk

Maxwell Levis[1,2], Monica Dimambro[1], Joshua Levy[3], Vincent Dufort[1], Abby Fraade[4], Max Winer[1] and Brian Shiner[1,2,5]

[1]White River Junction VA Medical Center, White River Junction, VT, USA; [2]Department of Psychiatry, Geisel School of Medicine at Dartmouth, Hanover, NH, USA; [3]Pathology and Computational Biomedicine, Cedars Sinai Medical Center, Los Angeles, CA, USA; [4]Long Island University, Brooklyn, NY, USA and [5]National Center for PTSD Executive Division, White River Junction, VTS, USA

Abstract

Background. Although the Department of Veterans Affairs (VA) has made important suicide prevention advances, efforts primarily target high-risk patients with documented suicide risk, such as suicidal ideation, prior suicide attempts, and recent psychiatric hospitalization. Approximately 90% of VA patients that go on to die by suicide do not meet these high-risk criteria and therefore do not receive targeted suicide prevention services. In this study, we used national VA data to focus on patients that were not classified as high-risk, but died by suicide.

Methods. Our sample included all VA patients who died by suicide in 2017 or 2018. We determined whether patients were classified as high-risk using the VA's machine learning risk prediction algorithm. After excluding these patients, we used principal component analysis to identify moderate-risk and low-risk patients and investigated demographics, service-usage, diagnoses, and social determinants of health differences across high-, moderate-, and low-risk subgroups.

Results. High-risk ($n = 452$) patients tended to be younger, White, unmarried, homeless, and have more mental health diagnoses compared to moderate- ($n = 2149$) and low-risk ($n = 2209$) patients. Moderate- and low-risk patients tended to be older, married, Black, and Native American or Pacific Islander, and have more physical health diagnoses compared to high-risk patients. Low-risk patients had more missing data than higher-risk patients.

Conclusions. Study expands epidemiological understanding about non-high-risk suicide decedents, historically understudied and underserved populations. Findings raise concerns about reliance on machine learning risk prediction models that may be biased by relative underrepresentation of racial/ethnic minorities within health system.

In 2019, over 6000 Veterans died from suicide in the US, a rate that is, when adjusting for age and sex, 57% higher than civilian adults' suicide rate (OMHSP, 2021). While the US Department of Veterans Affairs (VA) has made important suicide prevention contributions, these efforts have primarily targeted high-risk Veterans with documented suicide risk, including suicidal ideation, prior suicide attempts, and other flagged concerns (Matarazzo et al., 2023; McCarthy et al., 2015). More than 90% of VA patients that go on to die by suicide (Kessler et al., 2017), however, do not meet these criteria and therefore do not fall into this high-risk tier, nor receive targeted suicide prevention services (McCarthy et al., 2021).

Predicting suicide remains notoriously challenging (Nock, Ramirez, & Rankin, 2019), constrained by the relatively low rates of suicide and the diversity of symptom typologies. Given this difficulty, the VA's suicide prevention strategy pragmatically focuses on patients with the highest likelihood of dying by suicide (Kessler et al., 2017). This strategy has led to a range of risk classification innovations, including Recovery Engagement and Coordination for Health – Veterans Enhanced Treatment (REACH-VET: Cannizzaro, 2017; Kessler et al., 2017), a suicide risk prediction algorithm that automatically evaluates all VA patients for high-risk status. REACH-VET conceptualizes high-risk patients as the top 1% risk tier, which includes 10% of deaths by suicide.

REACH-VET, which uses a network of structured electronic health record (EHR) risk-variables ranging from health service usage to psychotropic medication usage, diagnoses, prior suicide prior attempts, and socio-demographics, is the VA's most sophisticated and far-reaching suicide prediction method (Jobes, Haddock, & Olivares, 2019; Matarazzo, Brenner, & Reger, 2019). That being said, REACH-VET's impact is constrained to the small subgroup of patients with known risk factors, like prior suicide attempts, opioid usage, and inpatient mental health treatment. Unfortunately, this leaves the majority of patients that go on to die by suicide without accurate risk classification or designated high-impact treatments (Mann, Michel, & Auerbach, 2021; Stanley et al., 2009). Recognizing this limitation, REACH-VET's authors acknowledge that their model inevitably fails to impact the majority of Veterans at risk for death by suicide whose risk falls outside of the high-risk tier:

> To achieve substantial reductions in the burden of suicide, it will be necessary to target larger strata of patients at lower – but still elevated – risk; for example, the 5.00% of patients who account for 24% of suicides in VA patients over 1 year. Because so many patients fall into this stratum, and because of the magnitude of the resources that would be required for a comprehensive approach for these patients, demonstration projects and research are needed to develop and validate an array of risk-stratified interventions that can be realistically delivered across a health care system (McCarthy et al., 2015).

Departing from the conventional focus on the high-risk population, this study analyzes VA patient suicide risk dimensions across risk tiers. REACH-VET, which uses a 0–100 percentile risk scoring system, with 100% being the lowest possible risk score, offers a potential method to identify not only the high-risk population, but also those with lower classified risk. In this study, we use these scores to evaluate classified risk, alongside demographic, social determinants of health, service usage, and diagnostic variables, to gain clinical and epidemiological understanding of all Veterans that died by suicide.

Following the REACH-VET authors' guidance, this study helps to lay the groundwork for future suicide prevention services for 'larger strata of patients at lower – but still elevated – risk'. While lower-risk populations represent more than 90% of VA suicide decedents, to date they have been understudied and underserved. As an initial step toward reaching this community, this study uses national VA data to characterize patients across suicide risk strata to aid epidemiological understanding about risk dimensions and to help scaffold future development of evidenced-based suicide prevention services for all Veterans.

Methods

Sample selection

To develop the study sample, we linked EHR data from the VA Corporate Data Warehouse (CDW) EHR with cause of death data from the VA-Department of Defense Mortality Data Repository (MDR; VA/DoD, 2017) to identify all patients who died by suicide that had at least one VA health care encounter in either 2017 or 2018. REACH-VET scores are recalculated monthly for all VA patients. With support from the VA Office of Mental Health and Suicide Prevention, we identified each case's REACH-VET risk score during the month before death. Using CDW data, we pulled demographics, social determinants of health, service usage, prescriptions, and diagnostic information. In addition to REACH-VET, the VA uses a suicide risk warning system that can be manually designated by patients' clinicians ('high-risk flag') (Hein, Peltzman, Hallows, Theriot, & McCarthy, 2021). We included high-risk flag indication from any time point within 3 months before death date as a descriptive variable within our analytic model to evaluate the overlap between algorithmic and clinical risk monitoring.

Analysis

There were three steps to our analysis. First, we evaluated suicide risk concentration as measured by REACH-VET. We identified the high-risk population as any patient with $a \leq 1$ REACH-VET score based on the high concentration of suicide deaths within this tier (McCarthy et al., 2015). Second, we used a data-driven method to assess patterns in demographics, diagnoses, and service usage among the remaining patients who died by suicide but were not identified as high-risk. We used principal component analysis (PCA; Jolliffe, 2002), an orthogonal linear transformation technique that maps data into coordinates based on projections of the greatest amounts of variance, to identify subgroups within the non-high suicide risk population. Variables with greater than 25% missing data were excluded from our PCA. Outliers were removed from the derived model based on standardized distribution (Serneels & Verdonck, 2008). Second, we investigate the link between patient characteristics and their REACH-VET scores in order to calculate patient risk tiers. We performed K-Means clustering on data transformed via PCA (Ding & He, 2004). This enabled us to categorize patients into clusters that we hypothesized would correspond to varying levels of REACH-VET risk scores. To differentiate patients into moderate and low-risk groups effectively, we sought an optimal cut-off point within the REACH-VET percentile scores. This cut-off was derived by taking the mean of the lower quartile of REACH-VET scores in the first cluster and the upper quartile in the second cluster, an approach we have designated as the Q1–Q3 method. The subsequent risk tiers, delineated by this REACH-VET cut-off, demonstrated a high degree of association with the original clusters. This was corroborated using the adjusted Rand Index (Steinley, 2004), which accounts for chance when measuring the similarity of two clustering solutions: the one obtained through our unsupervised K-Means approach based on patient characteristics, and the other defined by our established REACH-VET cut-off. A comprehensive assessment of the adjusted Rand Index across various potential REACH-VET percentile cut-offs confirmed the appropriateness of the Q1–Q3 method in mirroring the initial clustering. Third, we compared our identified non-high-risk subgroups to the REACH-VET identified high-risk subgroup. We used odds ratios (OR) to compare subgroups using variables included within REACH-VET's predictive model, as well as additional demographics, social determinants of health, disability, mental health services, medical services, and diagnoses variables.

Results

We identified a total of 4810 cases. Within this sample, 9.4% ($n = 452$) were identified as 'high-risk' as indicated by REACH-VET risk percentile, as presented in Fig. 1. After excluding the high-risk subgroup, we analyzed the remaining non-high-risk population using PCA, which led to two components, as presented in Fig. 2. We established cut points on REACH-VET by averaging the first component's lowest quartile with the second component's highest quartile. This REACH-VET cut point allowed us to differentiate 'moderate-risk' ($n = 2149$ or 44.7% of sample), and 'low-risk' ($n = 2209$ or 45.9% of sample) subgroups, as presented in Fig. 3. Subgroup REACH-VET scores were significantly different ($p < 0.001$).

When comparing the high, moderate, and low-risk subgroups, marked differences in demographics, social determinants of health, disability, mental health services, medical services, medications, and diagnoses were detected, as presented in Table 1. The high-risk subgroup included the plurality of patients who received clinical high-risk flags. Patients classified as high-risk tended to be younger, White, unmarried, experienced homelessness, and have elevated

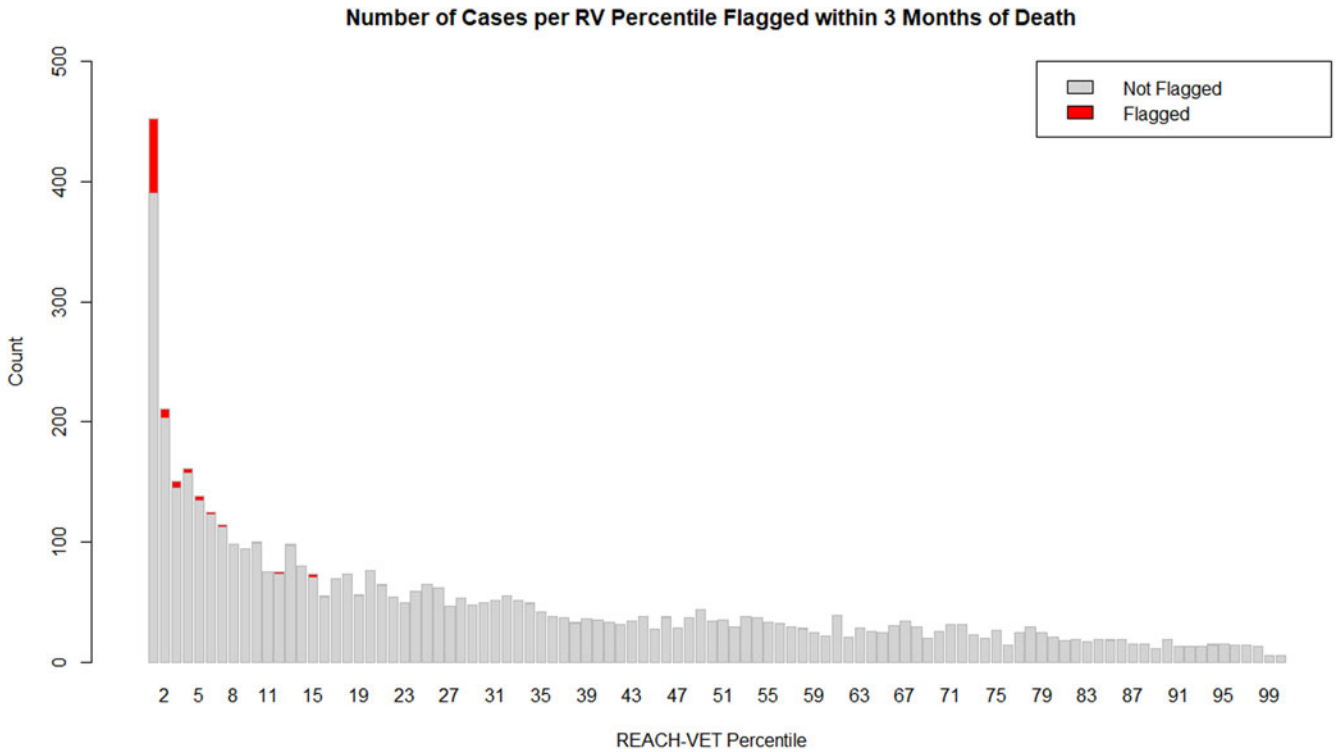

Figure 1. Plotting Veterans Affairs patient suicide deaths by Recovery Engagement and Coordination for Health – Veterans Enhanced Treatment (REACH-VET) (Cannizzaro, 2017; Kessler et al., 2017) score and clinical high-risk indication ('high-risk flag') (Hein et al., 2021) among patients that died by suicide in 2017 and 2018. The figure identifies the high-degree of overlap between REACH-VET, an algorothimic risk modeling technique, and high-risk flagging, a clinical risk modeling techinque.
Notes: RV is used in the figure title as an abbreviation for REACH-VET.

mental health burden, when compared to moderate- and low-risk patients. Patients that were classified as moderate- and low-risk tended to be older, married, Black, and Native American or Pacific Islander, and have elevated physical health burden. Lower-risk patients had substantially more missing data than high-risk patients.

When comparing high-risk and moderate-risk subgroups, we found that recent homelessness, elevated mental health burden,

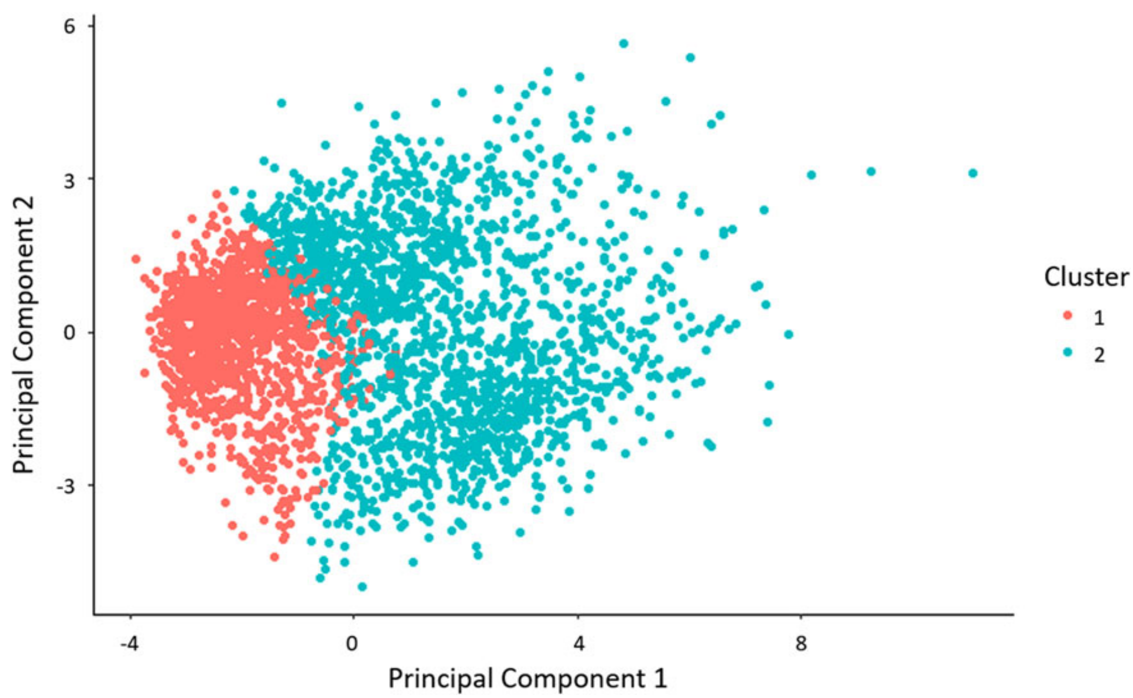

Figure 2. Principal components analysis (Jolliffe, 2002) of Veterans Affairs patient suicide deaths in 2017 and 2018, not-including those classified as high-risk by Recovery Engagement and Coordination for Health – Veterans Enhanced Treatment (REACH-VET; Cannizzaro, 2017; Kessler et al., 2017).

https://doi.org/10.1017/S0033291724001296 Published online by Cambridge University Press

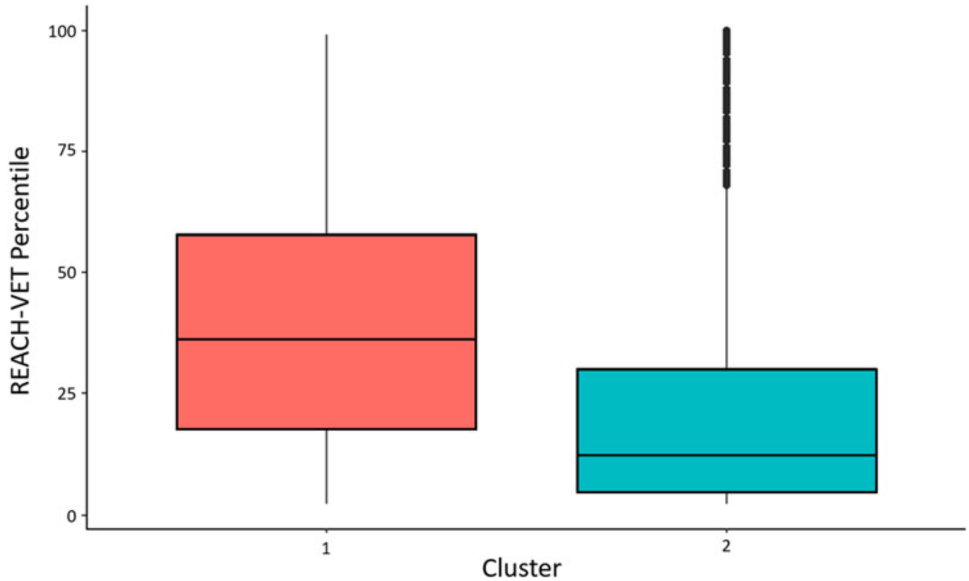

Figure 3. Boxplot presenting principal components analysis (Jolliffe, 2002) derived clusters. Boxplots were used to identify the cut point between low-risk and moderate-risk patients' Recovery Engagement and Coordination for Health – Veterans Enhanced Treatment (REACH-VET; Cannizzaro, 2017; Kessler et al., 2017) scores. The cut point was identified by averaging the first cluster's first REACH-VET quartile with the second cluster's third quartile. Clusters' REACH-VET scores were significantly different ($p < 0.001$).

depression, substance abuse, and clinical high-risk flag substantially increased the odds of being in the high-risk subgroup. Conversely, we found that being married, over the age of 75, and having served during the Vietnam era substantially reduced the odds of being in the high-risk subgroup.

When comparing moderate-risk and low-risk subgroups, we found that recent homelessness, elevated mental health burden, conduct disorder, clinical high-risk flag, depression, and receiving an antidepressant prescription substantially increased the odds of being in the moderate-risk subgroup. Conversely, we found that being married, over the age of 75, being Black, Native American or Pacific Islander, having any missing variables, and having some disability substantially reduced the odds of being in the moderate-risk subgroup.

When comparing high-risk and low-risk subgroups, we found that elevated mental health, anxiety, depression, personality and conduct disorders, substance abuse, clinical high-risk flag, and receiving psychopharmaceutical treatment dramatically increased the odds of being in the high-risk subgroup. Conversely, we found that being Black, married, over the age of 75, and having some physical disability substantially reduced the odds of being in the high-risk subgroup. Full OR results are presented in Table 2.

Discussion

Given the challenge of predicting suicide risk, REACH-VET was designed to identify patients with the highest suicide risk concentration (Kessler, Bossarte, Luedtke, Zaslavsky, & Zubizarreta, 2020), a measurement defined by the REACH-VET authors as 'the ratio of observed case patients to the case patients that would be expected if the distribution were uniform across strata' (McCarthy et al., 2015). Our high-risk subgroup confirms this distribution, accounting for a large swatch of patients in one risk stratum. As anticipated this ratio was much higher in the high-risk (9.4× greater than the risk rate in the overall sample) than the moderate-risk (1.9× greater than the risk rate in the overall sample) and low-risk (0.4× lower than the risk rate in the overall sample) subgroups. In addition to being identified as high-risk algorithmically, the high-risk subgroup tends to also be flagged by clinicians as being at high suicide risk nine times more than the moderate-risk and 58 times more than the low-risk subgroups. This high-risk subgroup tends to be younger, White, not married, and utilizes more mental health services and fewer medical services, when compared to patients with lower estimated risk. This relatively homogeneous subgroup stands in contrast to the much larger and more heterogeneous moderate-risk and low-risk subgroups, each of which is five times larger than the high-risk subgroup.

While there are obvious reasons to prioritize the high-risk subgroup's suicide prevention care, including their elevated risk concentration, mental health burden, relative demographic homogeneity, and prior utilization of VA mental health care, it is nonetheless imperative to prevent suicide across risk tiers, especially as the lower tiers represent the majority of suicide deaths. It is also necessary to recognize limitations associated with current prediction methods, shedding light on risk tier demographic disparities. To this end, we found that the high-risk subgroup included 10.06, 7.21, and 5.88% of the White, Native American or Pacific Islander, and Black patients. In contrast, the low-risk subgroup included 49.08, 63.10, and 64.3% of these respective populations. When looking at age distribution, we found that the high-risk subgroup included 13.11% of 18–34 and 1.8% of 75+ years-old patients, while the low-risk subgroup included 38.36 and 68.26% of these populations. These results indicate that select patient populations' risk burdens are undervalued within both algorithmic and clinical risk models, minimizing their ability to receive prevention care.

Representation

It is important to emphasize that REACH-VET's model was developed by identifying and weighting a network of machine

https://doi.org/10.1017/S0033291724001296 Published online by Cambridge University Press

Table 1. Veterans Affairs patients that died by suicide in 2017 and 2018 based on suicide risk subgroup

	1: High-risk (*N* = 452)	2: Moderate-risk (2–24) (*N* = 2149)	3: Low-risk (25–100) (*N* = 2209)
Demographics			
Female	23 (5.1%)	72 (3.4%)	84 (3.8%)
Non married	318 (70.4%)	1358 (63.2%)	899 (40.7%)
Married	119 (26.3%)	684 (31.8%)	1135 (51.4%)
Rural	79 (17.5%)	444 (20.7%)	396 (17.9%)
Homeless_prior24m	71 (15.7%)	140 (6.5%)	45 (2.0%)
Any missing variables	275 (60.8%)	1349 (62.8%)	1599 (72.4%)
Some disability	31 (6.9%)	144 (6.7%)	261 (11.8%)
Eligibility			
Not service connected	170 (37.6%)	970 (45.1%)	904 (40.9%)
Service connected <50%	82 (18.1%)	339 (15.8%)	467 (21.1%)
Service connected 50–100%	166 (36.7%)	691 (32.2%)	671 (30.4%)
Age group			
1: 18–34	80 (17.7%)	296 (13.8%)	234 (10.6%)
2: 35–54	156 (34.5%)	514 (23.9%)	393 (17.8%)
3: 55–74	194 (42.9%)	988 (46.0%)	780 (35.3%)
4: 75+	22 (4.9%)	351 (16.3%)	802 (36.3%)
Age			
Mean (s.d.)	51.8 (15.4)	58.3 (17.4)	65.1 (19.4)
Median [Min, Max]	53.5 [22.0, 92.0]	61.0 [19.0, 101]	69.0 [19.0, 101]
Race			
Black	15 (3.3%)	76 (3.5%)	164 (7.4%)
Hispanic	26 (5.8%)	79 (3.7%)	88 (4.0%)
Native American or Pacific Islander	8 (1.8%)	33 (1.5%)	70 (3.2%)
Unknown	10 (2.2%)	45 (2.1%)	292 (13.2%)
White	393 (86.9%)	1916 (89.2%)	1595 (72.2%)
Deployment			
Vietnam	127 (28.1%)	785 (36.5%)	796 (36.0%)
Afghanistan or Iraq	213 (47.1%)	792 (36.9%)	609 (27.6%)
Physical health burden			
0 (None)	190 (42.0%)	734 (34.2%)	594 (26.9%)
1	143 (31.6%)	614 (28.6%)	580 (26.3%)
2 (Elevated)	90 (19.9%)	395 (18.4%)	329 (14.9%)
Missing	29 (6.4%)	406 (18.9%)	706 (32.0%)
Mental health burden			
0 (None)	25 (5.5%)	453 (21.1%)	944 (42.7%)
1	104 (23.0%)	802 (37.3%)	477 (21.6%)
2 (Elevated)	294 (65.0%)	488 (22.7%)	82 (3.7%)
Missing	29 (6.4%)	406 (18.9%)	706 (32.0%)
Mental health diagnosis/risk flag			
Anxiety	358 (79.2%)	1118 (52.0%)	461 (20.9%)
Bipolar	188 (41.6%)	407 (18.9%)	160 (7.2%)

(*Continued*)

https://doi.org/10.1017/S0033291724001296 Published online by Cambridge University Press

Table 1. (*Continued.*)

	1: High-risk (*N* = 452)	2: Moderate-risk (2–24) (*N* = 2149)	3: Low-risk (25–100) (*N* = 2209)
Conduct	29 (6.4%)	39 (1.8%)	5 (0.2%)
Depression	440 (97.3%)	1607 (74.8%)	728 (33.0%)
Neurocognitive	87 (19.2%)	221 (10.3%)	145 (6.6%)
OCD	25 (5.5%)	52 (2.4%)	29 (1.3%)
PTSD	250 (55.3%)	796 (37.0%)	468 (21.2%)
Personality	154 (34.1%)	272 (12.7%)	103 (4.7%)
Sleeping	287 (63.5%)	1037 (48.3%)	629 (28.5%)
Substance	374 (82.7%)	1026 (47.7%)	473 (21.4%)
Trauma	338 (74.8%)	1106 (51.5%)	656 (29.7%)
Combat trauma	136 (30.1%)	567 (26.4%)	501 (22.7%)
Military sexual trauma	22 (4.9%)	92 (4.3%)	46 (2.1%)
High suicide risk flag	62 (13.7%)	36 (1.7%)	6 (0.3%)
Number of inpatient mental health visits within 1 year of death			
Mean (s.d.)	1.19 (1.63)	0.0731 (0.372)	0.00860 (0.110)
Median [Min, Max]	1.00 [0, 14.0]	0 [0, 9.00]	0 [0, 2.00]
Number of inpatient mental health days within 1 year of death			
Mean (s.d.)	76.2 (117)	7.33 (44.0)	1.34 (20.5)
Median [Min, Max]	16.0 [0, 366]	0 [0, 366]	0 [0, 366]
Number of emergency department visits within 1 year of death			
Mean (s.d.)	2.71 (2.93)	1.32 (1.38)	1.08 (0.863)
Median [Min, Max]	1.00 [0, 28.0]	1.00 [0, 20.0]	1.00 [0, 17.0]
Prescriptions			
Opioid Rx_prior12	156 (34.5%)	593 (27.6%)	324 (14.7%)
Opioid Rx_prior24	198 (43.8%)	767 (35.7%)	443 (20.1%)
Mood Stabilizer Rx_prior12	282 (62.4%)	706 (32.9%)	272 (12.3%)
Mood Stabilizer Rx_prior24	320 (70.8%)	846 (39.4%)	338 (15.3%)
Antipsychotic Rx_prior12	224 (49.6%)	415 (19.3%)	105 (4.8%)
Antipsychotic Rx_prior24	250 (55.3%)	502 (23.4%)	134 (6.1%)
Antidepressant Rx_prior12	387 (85.6%)	1211 (56.4%)	405 (18.3%)
Antidepressant Rx_prior24	410 (90.7%)	1406 (65.4%)	480 (21.7%)

Subgroups were derived from Recovery Engagement and Coordination for Health – Veterans Enhanced Treatment (REACH-VET) (Cannizzaro, 2017; Kessler et al., 2017) risk scores (high-risk = 1, moderate-risk = 2–24, and low-risk = 25–100). Analysis includes demographics, social determinants of health, disability, mental health services, medical services, and diagnoses variables.
Notes: Variables marked as '_prior12' include data from the 12 months before death.
Variables marked as '_prior24' include data from the 24 months prior to death.

learning-derived variables to best predict suicide for the subset of the VA patient population with the highest likelihood of dying by suicide (Kessler et al., 2017). Accordingly, REACH-VET offers less predictive relevance for some patients, especially those who come from demographic backgrounds that have lower suicide risk concentration or for patients who receive services that are not associated with suicide risk. Furthermore, REACH-VET works by leveraging patient EHR; accurate REACH-VET scoring therefore hinges on accurate EHR data. One of our consistent findings is that patients at lower-risk levels use fewer mental health services and have higher levels of missing data (Peltzman, Rice, Jones, Washington, & Shiner, 2022; Sullivan et al., 2023). This could be associated with these patients accessing less services (Meffert et al., 2019), preferring non-VA services (Mattocks et al., 2019), having greater barriers to care (Elnitsky et al., 2013), or not disclosing personal information within care (Botero et al., 2020). Providing these underrepresented patients, and subgroups of patients, with alternative suicide prediction and prevention mechanisms is a critical step in reducing patient suicide (Jobes et al., 2019; Miller-Matero et al., 2023).

The concern of over- and under-representation of select populations is a latent critique of machine learning and prediction methods (Huang, Galal, Etemadi, & Vaidyanathan, 2022). Like other machine learning classification metrics (Gianfrancesco,

https://doi.org/10.1017/S0033291724001296 Published online by Cambridge University Press

Table 2. Odds ratios for risk subgroup classification (high-risk, moderate-risk, and low-risk) for Veterans Affairs patients that died by suicide in 2017 and 2018

	High *v.* moderate ($N = 2601$)	*p* Value[a]	Moderate *v.* low ($N = 4358$)	*p* Value[a]	High *v.* low ($N = 2661$)	*p* Value[a]
Demographics						
Gender (male)	0.65 [0.40, 1.05]	0.0962	1.14 [0.83, 1.57]	0.4632	0.74 [0.46, 1.18]	0.2356
Marital status (married)	**0.74 [0.59, 0.93]**	**0.0112**	**0.40 [0.35, 0.45]**	**<0.0001**	**0.30 [0.24, 0.37]**	**<0.0001**
Rural	**0.70 [0.53, 0.91]**	**0.0077**	0.98 [0.84, 1.15]	0.8737	**0.69 [0.53, 0.90]**	**0.0058**
Homeless_prior24m	**2.67 [1.97, 3.63]**	**<0.0001**	**3.35 [2.38, 4.71]**	**<0.0001**	**8.96 [6.07, 13.22]**	**<0.0001**
Any missing variables	0.92 [0.75, 1.13]	0.4246	**0.640 [0.57, 0.73]**	**<0.0001**	**0.59 [0.48, 0.73]**	**<0.0001**
Some disability	1.03 [0.69, 1.53]	0.9177	**0.54 [0.43, 0.66]**	**<0.0001**	**0.55 [0.37, 0.81]**	**0.0016**
Eligibility						
Not service connected	*Ref*	–	*Ref*	–	*Ref*	–
Other	1.64 [1.02, 2.63]	0.0546	0.92 [0.68, 1.26]	0.6354	1.51 [0.94, 2.43]	0.1079
Service connected <50%	**1.38 [1.03, 1.85]**	**0.0361**	**0.68 [0.57, 0.80]**	**<0.0001**	0.93 [0.70, 1.24]	0.6643]
Service connected >50%	**1.37 [1.08, 1.73]**	**0.0093**	0.96 [0.83, 1.10]	0.5689	**1.32 [1.04, 1.67]**	**0.0250**
Age group						
1: 18–34	*Ref*	–	*Ref*	–	*Ref*	–
2: 35–54	1.12 [0.83, 1.52]	0.4882	1.03 [0.83, 1.28]	0.7829	1.16 [0.85, 1.59]	0.3828
3: 55–74	**0.73 [0.54, 0.97]**	**0.0356**	1.00 [0.82, 1.22]	1.0000	**0.73 [0.54, 0.98]**	**0.0393**
4: 75+	**0.23 [0.14, 0.38]**	**<0.0001**	**0.35 [0.28, 0.43]**	**<0.0001**	**0.08 [0.05, 0.13]**	**<0.0001**
Race						
White	*Ref*	–	*Ref*	–	*Ref*	–
Black	1.00 [0.57, 1.76]	1.0000	**0.38 [0.29, 0.51]**	**<0.0001**	**0.38 [0.22, 0.66]**	**0.0001**
Hispanic	1.56 [0.99, 2.46]	0.0666	0.81 [0.59, 1.10]	0.2034	1.26 [0.80, 1.99]	0.3266
Native American or Pacific Islander	1.23 [0.56, 2.69]	0.5307	**0.38 [0.25, 0.58]**	**<0.0001**	**0.46 [0.22, 0.97]**	**0.0403**
Deployment						
Vietnam	**0.68 [0.55, 0.86]**	**<0.0001**	1.00 [0.88, 1.13]	1	**0.68 [0.55, 0.85]**	**<0.0001**
Afghanistan or Iraq	**1.55 [1.26, 1.91]**	**<0.0001**	**1.51 [1.32, 1.71]**	**<0.0001**	**2.34 [1.90, 2.88]**	**<0.0001**
Physical health burden						
0 (None)	*Ref*	–	*Ref*	–	*Ref*	–
1	0.90 [0.71, 1.15]	0.4240	0.86 [0.73, 1.00]	0.0551	**0.77 [0.60, 0.99]**	**0.0403**
2 (Elevated)	0.88 [0.67, 1.16]	0.3993	0.97 [0.81, 1.17]	0.7806	0.86 [0.64, 1.14]	0.3161
Mental health burden						
0 (None)	*Ref*	–	*Ref*	–	*Ref*	–
1	**2.35 [1.50, 3.69]**	**<0.0001**	**3.50 [2.99, 4.11]**	**<0.0001**	**8.23 [5.25, 12.91]**	**<0.0001**
2 (Elevated)	**10.92 [7.12, 16.75]**	**<0.0001**	**12.40 [9.57, 16.07]**	**<0.0001**	**135.38 [84.91, 215.86]**	**<0.0001**
Mental health diagnosis/risk flag						
Anxiety	**3.51 [2.75, 4.47]**	**<0.0001**	**4.10 [3.59, 4.68]**	**<0.0001**	**14.36 [11.19, 18.42]**	**<0.0001**
Bipolar	**3.04 [2.45, 3.78]**	**<0.0001**	**2.98 [2.46, 3.62]**	**<0.0001**	**9.08 [7.09, 11.61]**	**<0.0001**
Conduct	**3.71 [2.27, 6.06]**	**<0.0001**	**8.12 [3.19, 20.64]**	**<0.0001**	**30.08 [11.58, 78.16]**	**<0.0001**
Depression	**12.32 [6.89, 22.04]**	**<0.0001**	**6.01 [5.27, 6.86]**	**<0.0001**	**74.09 [41.47, 132.36]**	**<0.0001**
Neurocognitive	**2.08 [1.58, 2.73]**	**<0.0001**	**1.63 [1.31, 2.02]**	**<0.0001**	**3.38 [2.53, 4.50]**	**<0.0001**
OCD	**2.36 [1.45, 3.84]**	**0.0011**	**1.86 [1.17, 2.94]**	**0.0095**	**4.38 [2.54, 7.55]**	**<0.0001**
PTSD	**2.10 [1.71, 2.58]**	**<0.0001**	**2.18 [1.90, 2.49]**	**<0.0001**	**4.58 [3.70, 5.66]**	**<0.0001**

(*Continued*)

https://doi.org/10.1017/S0033291724001296 Published online by Cambridge University Press

Table 2. (*Continued.*)

	High *v.* moderate ($N = 2601$)	p Value[a]	Moderate *v.* low ($N = 4358$)	p Value[a]	High *v.* low ($N = 2661$)	p Value[a]
Personality	**3.56 [2.82, 4.49]**	**<0.0001**	**2.95 [2.33, 3.73]**	**<0.0001**	**10.52 [7.97, 13.88]**	**<0.0001**
Psychotic	**3.16 [2.47, 4.03]**	**<0.0001**	**2.42 [1.92, 3.73]**	**<0.0001**	**7.66 [5.78, 10.14]**	**<0.0001**
Sleeping	**1.86 [1.51, 2.30]**	**<0.0001**	**2.33 [2.06, 2.64]**	**<0.0001**	**4.34 [3.51, 5.37]**	**<0.0001**
Substance	**5.24 [4.05, 6.78]**	**<0.0001**	**3.34 [2.93, 3.81]**	**<0.0001**	**17.50 [13.43, 22.7]**	**<0.0001**
Trauma	**2.79 [2.22, 3.51]**	**<0.0001**	**2.50 [2.21, 2.83]**	**<0.0001**	**6.97 [5.53, 8.79]**	**<0.0001**
Combat	1.20 [0.96, 1.50]	0.1156	**1.22 [1.06, 1.40]**	**0.0048**	**1.47 [1.17, 1.84]**	**0.0011**
Military sexual trauma	1.14 [0.71, 1.84]	0.6127	**2.10 [1.47, 3.01]**	**<0.0001**	**2.41 [1.43, 4.04]**	**0.0016**
High suicide risk flag	**9.33 [6.10, 14.27]**	**<0.0001**	**6.26 [2.63, 14.88]**	**<0.0001**	**58.37 [25.08, 135.87]**	**<0.0001**
Prescriptions						
Opioid Rx_prior12	**1.38 [1.11, 1.72]**	**0.0036**	**2.22 [1.91, 2.58]**	**<0.0001**	**3.07 [2.44, 3.85]**	**<0.0001**
Opioid Rx_prior24	**1.40 [1.14, 1.73]**	**0.0013**	**2.21 [1.93, 2.54]**	**<0.0001**	**3.11 [2.51, 3.85]**	**<0.0001**
Mood Stabilizer Rx_prior12	**3.39 [2.75, 4.18]**	**<0.0001**	**3.48 [2.98, 4.07]**	**<0.0001**	**11.81 [9.40, 14.85]**	**<0.0001**
Mood Stabilizer Rx_prior24	**3.73 [3.00, 4.65]**	**<0.0001**	**3.59 [3.11, 4.15]**	**<0.0001**	**13.42 [10.62, 16.95]**	**<0.0001**
Antipsychotic Rx_prior12	**4.11 [3.32, 5.08]**	**<0.0001**	**4.80 [3.84, 6.00]**	**<0.0001**	**19.69 [15.04, 25.77]**	**<0.0001**
Antipsychotic Rx_prior24	**4.06 [3.29, 5.01]**	**<0.0001**	**4.72 [3.86, 5.77]**	**<0.0001**	**19.16 [14.85, 24.73]**	**<0.0001**
Antidepressant Rx_prior12	**4.61 [3.50, 6.08]**	**<0.0001**	**5.75 [5.01, 6.60]**	**<0.0001**	**26.52 [19.96, 35.23]**	**<0.0001**
Antidepressant Rx_prior24	**5.16 [3.71, 7.17]**	**<0.0001**	**6.82 [5.96, 7.80]**	**<0.0001**	**35.16 [25.20, 49.07]**	**<0.0001**

Odd ratios with significant p values are marked in bold.
Notes: [a]p values calculated using Fisher's exact method.
Variables marked as '_prior12' include data from the 12 months before death.
Variables marked as '_prior24' include data from the 24 months prior to death.

Tamang, Yazdany, & Schmajuk, 2018), REACH-VET may be constrained by concerns about lack of model transparency, biased training data, inconsistent analytic methods, and absence of clinical validations (Huang et al., 2022). As machine learning methods typically utilize pre-existing data, this approach frequently prioritizes historic trends over contemporary realities, replicating potential biases and service access limitations and skewing prediction models toward select race, age, and gender populations (Gianfrancesco et al., 2018; Huang et al., 2022; Nong, Williamson, Anthony, Platt, & Kardia, 2022). Indeed, disparities in mental healthcare utilization or underdiagnoses of depression and other suicide-related diagnoses among older adults and racial and ethnic minorities have been shown to contribute to misdiagnosis and under-hospitalization (Bailey, Mokonogho, & Kumar, 2019; Lavingia, Jones, & Asghar-Ali, 2020), factors that may weigh heavily in algorithmic risk modeling. Although a variety of analytic methods have been developed (Afrose, Song, Nemeroff, Lu, & Yao, 2022) that could alleviate some of modeling concerns, the efficacy of these approaches in regard to suicide prediction remains unknown. This study helps address these concerns by focusing on underrepresented patient populations with the intention of developing new prediction and prevention mechanisms that provide more equitable services.

Limitations

We specifically utilized a retrospective sample as this format was most sensible for studying VA suicide deaths. Given our focus on Veteran suicide, we intentionally only used Veteran data, and restricted non-Veteran generalization. Unfortunately, we have limited access to VA patients' usage of medical providers outside of the VA. Our models are somewhat constrained by the higher prevalence of missing data among low-risk patients than other risk subgroups.

Implications

Our study addresses select shortcomings associated with current high-risk suicide modeling methods and develops a framework to cluster lower-risk populations. Implications include expanding epidemiological understanding about Veteran suicide risk distribution. Our analysis highlighted the significant differences between risk subgroups, allowing identification of distinct care utilization trends and risk factors, information that could be subsequently used to scaffold tailored prevention programs. As current VA practice guidelines encourage utilizations of patient-centered and representative care models (VA/DoD, 2019), it is all the more important to prioritize the development of tailored prediction and prevention models for each of these subgroups. We are optimistic that this work can aid the development of adjunctive prediction models for populations that are under-represented within the REACH-VET, including non-White, older patients, and non-mental health service users.

Given the VA's dedication to combating Veteran suicide, it is essential to reach as many patients as possible across all risk tiers. As a first step toward this goal, it will be important to gain clinical and epidemiological understanding about these patient populations and risk subgroups. Our prior work suggests the utility of leveraging additional EHR data formats, including unstructured EHR like provider notes, to improve clinical awareness and risk prediction accuracy (Levis et al., 2023a; Levis, Levy, Dufort, Russ, & Shiner, 2023b). This approach has been shown to offer specific benefits for patients that are missing data (Shiner et al., 2021) and could, accordingly, help mitigate limitations associated with non-high-risk patient data. Improved data, including unstructured EHR-derived variables, could aid risk clustering and expand epidemiological and clinical information. Future research could in turn utilize derived information to develop targeted prediction and prevention methods. Additionally, following World Health Organization guidance (Wasserman, Tadić, & Bec, 2023), future research could utilize these materials to aid the development of universal suicide prevention strategies for low-risk patients (Klimes-Dougan, Klingbeil, & Meller, 2013) and selective suicide prevention strategies that aim to reach moderate-risk patients (Ahmedani & Vannoy, 2014), alongside indicated strategies for high-risk patients (McCarthy et al., 2015). We are hopeful that identifying these trends can lead to more effective risk appraisal regardless of risk strata and, in turn, toward targeted suicide prevention care across all patient populations.

Data availability statement. Data access is restricted due to the clinical nature of dataset and VA privacy protections.

Funding statement. Dr Levis was supported by a VA New England Career Development Award (V1CDA-2020-60) and by a VA Clinical Science Research and Development Career Development Award (CX002630). Dr Levy was supported by a DoD award (HT9425-23-1-0267).

Competing interests. None.

Ethical standards. The authors assert that all procedures contributing to this work comply with the ethical standards of the relevant national and institutional committees on human experimentation and with the Helsinki Declaration of 1975, as revised in 2008.

References

Afrose, S., Song, W., Nemeroff, C. B., Lu, C., & Yao, D. (2022). Subpopulation-specific machine learning prognosis for underrepresented patients with double prioritized bias correction. *Communications Medicine*, *2*(1), 111. https://doi.org/10.1038/s43856-022-00165-w

Ahmedani, B. K., & Vannoy, S. (2014). National pathways for suicide prevention and health services research. *American Journal of Preventive Medicine*, *47*(3), S222–S228. https://doi.org/10.1016/j.amepre.2014.05.038

Bailey, R., Mokonogho, J., & Kumar, A. (2019). Racial and ethnic differences in depression: Current perspectives. *Neuropsychiatric Disease and Treatment*, *15*, 603–609. https://doi.org/10.2147/NDT.S128584

Botero, G., Rivera, N. I., Calloway, S. C., Ortiz, P. L., Edwards, E., Chae, J., & Geraci, J. C. (2020). A lifeline in the dark: Breaking through the stigma of Veteran mental health and treating America's combat Veterans. *Journal of Clinical Psychology*, *76*(5), 831–840. https://doi.org/10.1002/jclp.22918

Cannizzaro, K. (2017). *REACH VET and the possible impact on integrated healthcare*. VA Health Care. Rocky Mountain Regional VAMC, Aurora, CO. https://avapl.org/conference/pubs/2018%20Conference%20Presentations/Cannizzaro%20-%20slides%20-%20Open%20Plenary.pdf

Ding, C., & He, X. (2004). K-means clustering via principal component analysis. *Twenty-first international conference on machine learning – ICML '04*, 29. https://doi.org/10.1145/1015330.1015408

Elnitsky, C. A., Andresen, E. M., Clark, M. E., McGarity, S., Hall, C. G., & Kerns, R. D. (2013). Access to the US Department of Veterans Affairs health system: Self-reported barriers to care among returnees of operations enduring freedom and Iraqi freedom. *BMC Health Services Research*, *13*(1), 498. https://doi.org/10.1186/1472-6963-13-498

Gianfrancesco, M. A., Tamang, S., Yazdany, J., & Schmajuk, G. (2018). Potential biases in machine learning algorithms using electronic health record data. *JAMA Internal Medicine*, *178*(11), 1544. https://doi.org/10.1001/jamainternmed.2018.3763

Hein, T. C., Peltzman, T., Hallows, J., Theriot, N., & McCarthy, J. F. (2021). Suicide mortality among Veterans health administration care recipients with suicide risk record flags. *Psychiatric Services*, *73*(3), 259–264. https://doi.org/10.1176/appi.ps.202000771

Huang, J., Galal, G., Etemadi, M., & Vaidyanathan, M. (2022). Evaluation and mitigation of racial bias in clinical machine learning models: Scoping review. *JMIR Medical Informatics*, *10*(5), e36388. https://doi.org/10.2196/36388

Jobes, D. A., Haddock, L. A., & Olivares, M. R. (2019). Military and Veteran suicide prevention. In E. C. Ritchie, & M. D. Llorente (Eds.), *Veteran psychiatry in the US* (pp. 51–71). New York, NY: Springer International Publishing. https://doi.org/10.1007/978-3-030-05384-0_5

Jolliffe, I. T. (2002). *Principal component analysis*. New York, NY: Springer-Verlag. https://doi.org/10.1007/b98835

Kessler, R. C., Bossarte, R. M., Luedtke, A., Zaslavsky, A. M., & Zubizarreta, J. R. (2020). Suicide prediction models: A critical review of recent research with recommendations for the way forward. *Molecular Psychiatry*, *25*(1), 168–179. https://doi.org/10.1038/s41380-019-0531-0

Kessler, R. C., Hwang, I., Hoffmire, C. A., McCarthy, J. F., Petukhova, M. V., Rosellini, A. J., … Bossarte, R. M. (2017). Developing a practical suicide risk prediction model for targeting high-risk patients in the Veterans health administration. *International Journal of Methods in Psychiatric Research*, *26*(3), e1575. https://doi.org/10.1002/mpr.1575

Klimes-Dougan, B., Klingbeil, D. A., & Meller, S. J. (2013). The impact of universal suicide-prevention programs on the help-seeking attitudes and behaviors of youths. *Crisis*, *34*(2), 82–97. https://doi.org/10.1027/0227-5910/a000178

Lavingia, R., Jones, K., & Asghar-Ali, A. A. (2020). A systematic review of barriers faced by older adults in seeking and accessing mental health care. *Journal of Psychiatric Practice*, *26*(5), 367–382. https://doi.org/10.1097/PRA.0000000000000491

Levis, M., Levy, J., Dent, K. R., Dufort, V., Gobbel, G. T., Watts, B. V., & Shiner, B. (2023a). Leveraging natural language processing to improve electronic health record suicide risk prediction for Veterans health administration users. *The Journal of Clinical Psychiatry*, *84*(4), 22m14568. https://doi.org/10.4088/JCP.22m14568

Levis, M., Levy, J., Dufort, V., Russ, C. J., & Shiner, B. (2023b). Dynamic suicide topic modelling: Deriving population-specific, psychosocial and time-sensitive suicide risk variables from electronic health record psychotherapy notes. *Clinical Psychology & Psychotherapy*, *30*(4), 795810. https://doi.org/10.1002/cpp.2842

Mann, J. J., Michel, C. A., & Auerbach, R. P. (2021). Improving suicide prevention through evidence-based strategies: A systematic review. *American Journal of Psychiatry*, *178*(7), 611–624. https://doi.org/10.1176/appi.ajp.2020.20060864

Matarazzo, B. B., Brenner, L. A., & Reger, M. A. (2019). Positive predictive values and potential success of suicide prediction models. *JAMA Psychiatry*, *76*(8), 869. https://doi.org/10.1001/jamapsychiatry.2019.1519

Matarazzo, B. B., Eagan, A., Landes, S. J., Mina, L. K., Clark, K., Gerard, G. R., … Reger, M. A. (2023). The Veterans health administration REACH VET program: Suicide predictive modeling in practice. *Psychiatric Services*, *74*(2), 206–209. https://doi.org/10.1176/appi.ps.202100629

Mattocks, K. M., Cunningham, K., Elwy, A. R., Finley, E. P., Greenstone, C., Mengeling, M. A., Pizer, S. D., … Bastian, L. A. (2019). Recommendations for the evaluation of cross-system care coordination from the VA state-of-the-art working group on VA/non-VA care. *Journal of General Internal Medicine*, *34*(S1), 18–23. https://doi.org/10.1007/s11606-019-04972-1

McCarthy, J. F., Bossarte, R. M., Katz, I. R., Thompson, C., Kemp, J., Hannemann, C. M., … Schoenbaum, M. (2015). Predictive modeling and concentration of the risk of suicide: Implications for preventive interventions in the US Department of Veterans Affairs. *American Journal of Public Health*, *105*(9), 1935–1942. https://doi.org/10.2105/AJPH.2015.302737

McCarthy, J. F., Cooper, S. A., Dent, K. R., Eagan, A. E., Matarazzo, B. B., Hannemann, C. M., … Katz, I. R. (2021). Evaluation of the recovery engagement and coordination for health – Veterans enhanced treatment suicide risk modeling clinical program in the Veterans health administration.

JAMA Network Open, *4*(10), e2129900. https://doi.org/10.1001/jamanetworkopen.2021.29900
Meffert, B. N., Morabito, D. M., Sawicki, D. A., Hausman, C., Southwick, S. M., Pietrzak, R. H., & Heinz, A. J. (2019). US Veterans who do and do not utilize Veterans Affairs health care services: Demographic, military, medical, and psychosocial characteristics. *The Primary Care Companion for CNS Disorders*, *21*(1), 18m02350. https://doi.org/10.4088/PCC.18m02350
Miller-Matero, L. R., Yeh, H.-H., Maffett, A., Mooney, J. T., Sala-Hamrick, K., Frank, C. B., … Ahmedani, B. K. (2023). Racial-ethnic differences in receipt of past-year health care services among suicide decedents: A case-control study. *Psychiatric Services*, *75*(2), 124–130. https://doi.org/10.1176/appi.ps.20220578
Nock, M. K., Ramirez, F., & Rankin, O. (2019). Advancing our understanding of the who, when, and why of suicide risk. *JAMA Psychiatry*, *76*(1), 11. https://doi.org/10.1001/jamapsychiatry.2018.3164
Nong, P., Williamson, A., Anthony, D., Platt, J., & Kardia, S. (2022). Discrimination, trust, and withholding information from providers: Implications for missing data and inequity. *SSM – Population Health*, *18*, 101092. https://doi.org/10.1016/j.ssmph.2022.101092
OMHSP. (2021). *National Veteran Suicide Prevention annual report*. VA Office of Mental Health and Suicide Prevention, Washington DC. https://www.mentalhealth.va.gov/docs/data-sheets/2021/2021-National-Veteran-Suicide-Prevention-Annual-Report-FINAL-9-8-21.pdf
Peltzman, T., Rice, K., Jones, K. T., Washington, D. L., & Shiner, B. (2022). Optimizing data on race and ethnicity for Veterans Affairs patients. *Military Medicine*, *187*(7–8), e955–e962. https://doi.org/10.1093/milmed/usac066
Serneels, S., & Verdonck, T. (2008). Principal component analysis for data containing outliers and missing elements. *Computational Statistics & Data Analysis*, *52*(3), 1712–1727. https://doi.org/10.1016/j.csda.2007.05.024
Shiner, B., Levis, M., Dufort, V. M., Patterson, O. V., Watts, B. V., DuVall, S. L., … Maguen, S. (2021). Improvements to PTSD quality metrics with natural language processing. *Journal of Evaluation in Clinical Practice*, *28*(4), 520–530. https://doi.org/10.1111/jep.13587
Stanley, B., Brown, G., Brent, D. A., Wells, K., Poling, K., Curry, J., … Hughes, J. (2009). Cognitive-behavioral therapy for suicide prevention (CBT-SP): Treatment model, feasibility, and acceptability. *Journal of the American Academy of Child & Adolescent Psychiatry*, *48*(10), 1005–1013. https://doi.org/10.1097/CHI.0b013e3181b5dbfe
Steinley, D. (2004). Properties of the Hubert-Arable adjusted rand index. *Psychological Methods*, *9*(3), 386–396. https://doi.org/10.1037/1082-989X.9.3.386
Sullivan, J. L., Simons, K. V., Mills, W. L., Hilgeman, M. M., Freytes, I. M., Morin, R. T., … Byers, A. L. (2023). A paucity of data on Veterans 65 and older and risk of suicide: A systematic review. *The American Journal of Geriatric Psychiatry*, *31*(7), 525–539. https://doi.org/10.1016/j.jagp.2022.11.005
VA/DoD. (2017). *Joint Department of Veterans Affairs (VA) and Department of Defense (DoD) Mortality Data Repository – National Death Index (NDI)*. Center of Excellence for Suicide Prevention, Washington DC. https://www.mirecc.va.gov/suicideprevention/documents/VA_DoD-MDR_Flyer.pdf
VA/DoD. (2019). *VA/DoD clinical practice guideline for the assessment and management of patients at risk for suicide (2.0)*. The Assessment and Management of Suicide Risk Work Group, Washington DC. https://www.healthquality.va.gov/guidelines/mh/srb/
Wasserman, D., Tadić, I., & Bec, C. (2023). Vision zero in suicide prevention and suicide preventive methods. In K. E. Björnberg, S. O. Hansson, M.-Å. Belin, & C. Tingvall (Eds.), *The vision zero handbook* (pp. 1117–1142). New York, NY: Springer International Publishing. https://doi.org/10.1007/978-3-030-76505-7_43