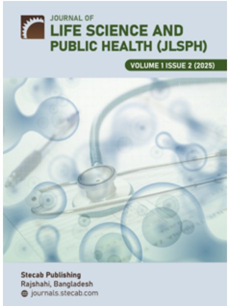

Journal of Life Science and Public Health (JLSPH)

Volume 1 Issue 2, (2025)

https://doi.org/10.69739/jlsph.v1i2.975

https://journals.stecab.com/jlsph

Published by
Stecab Publishing

Research Article

Mental-Health Crisis Prediction in U.S. Veterans: Opportunities and Pitfalls of Machine-Learning on VA–DoD Data

[1]Gbenga Adeniyi Adediran, [2]Andrew Ayemere Okhueigbe, [3]Ruth Ese Otaigboria, [4]Chiamaka Pamela Agu, [5]Gabriel Dogbanya

About Article

Article History

Submission: August 12, 2025
Acceptance : September 19, 2025
Publication : October 04, 2025

Keywords

Crisis Prediction, Machine-Learning, Mental-Health, VA–DoD

About Author

[1] Department of Computing and Engineering, Leeds Beckett University, Leeds, UK

[2] Department of Health Informatics, College of Professional Studies, University of Denver, Colorado, USA

[3] Department of Anthropology, Northern Illinois University, Dekalb Illinois, Nigeria

[4] Department of Public Health, University of New Haven, Connecticut, USA

[5] Department of Family Science, University of Maryland, College Park, MD, USA

Contact @ Gbenga Adeniyi Adediran
gbengadediran@gmail.com

ABSTRACT

One of the most urgent public health issues facing American veterans is mental health issues and suicide. Leveraging large-scale Department of Veterans Affairs (VA) and Department of Defense (DoD) data, machine learning (ML) models offer a complementary approach to traditional screening by mining high-dimensional electronic health records, administrative registers, and clinical text. This narrative review synthesizes developments from 2015 to 2025 in ML-based prediction of suicidal behavior and related crises among veterans. Key findings indicate moderate but clinically useful discrimination across studies; for example, operational deployment of VA risk modeling concentrated risk such that the top 1% of risk scores contained roughly 10.7% of subsequent suicides, enabling targeted outreach. ML approaches can improve identification of at-risk veterans and strengthen preventive workflows, yet translation is limited by false positives, algorithmic bias, data integration challenges, and uncertain impact on mortality. The review discusses veteran-specific risk factors, data infrastructure, modeling paradigms, validation evidence, and ethical governance, and concludes with recommendations to prioritize prospective evaluation, equity audits, and integration strategies that couple prediction with effective intervention.

Citation Style:

Adediran, G. A., Okhueigbe, A. A., Otaigboria, R. E., Agu, C. P., & Dogbanya, G. (2025). Mental-Health Crisis Prediction in U.S. Veterans: Opportunities and Pitfalls of Machine-Learning on VA–DoD Data. *Journal of Life Science and Public Health, 1*(2), 18-28. https://doi.org/10.69739/jlsph.v1i2.975

1. INTRODUCTION

Veteran suicide and mental health crises have persisted at alarmingly high rates despite extensive prevention efforts. Recent data indicate that the suicide rate among U.S. veterans remains significantly elevated compared to the civilian population. In 2022, the age-adjusted rate was 44% higher in veteran men and 92% higher in veteran women relative to non-veterans (Ramchand & Montoya, 2025). An average of 17.6 veterans die by suicide each day, according to the 2024 National Veteran Suicide Prevention Annual Report issued by VA's Office of Mental Health and Suicide Prevention (VA News, 2024), a figure that has shown only modest improvements in recent years. This disproportionate burden has led the Department of Veterans Affairs (VA) to designate suicide prevention as its "highest clinical priority," aiming to ensure at-risk veterans receive timely care (U.S. Government Accountability, 2022). Traditional suicide risk assessment in clinical settings relies heavily on patient self-reported suicidal ideation and clinician judgment. However, these methods often fail to predict outcomes. For example, one VA study found that 70% of veterans who died by suicide had denied suicidal ideation at their final clinical visit, underscoring the limitations of relying on expressed ideation alone (Smith *et al.*, 2013). Additionally, many veterans at risk are not actively engaged in VA healthcare in the period before suicide (Ramchand & Montoya, 2025), making them essentially invisible to standard clinical screening.

In response to this issue, population-level, data-driven prediction has emerged as a complementary strategy. The VA and Department of Defense (DoD) collectively possess vast electronic health record (EHR) repositories and administrative datasets encompassing millions of service members and veterans. Machine learning (ML) denotes a set of computational techniques that automatically detect complex, predictive patterns in large, heterogeneous datasets and produce models that estimate an individual's probability of a future outcome. Advances in machine learning (ML) allow these high-dimensional data to be mined for subtle patterns and risk factors that human clinicians might overlook (Zhang *et al.*, 2025). Early applications of ML in military and veteran cohorts demonstrated proof-of-concept that algorithms can stratify individuals by future suicide risk more effectively than chance (Kessler *et al.*, 2017; Zhang *et al.*, 2025). In 2017, the VA implemented the Recovery Engagement and Coordination for Health–Veterans Enhanced Treatment (REACH VET) program, one of the first national efforts to apply an ML-based model for predicting suicide risk for proactive outreach (Matarazzo *et al.*, 2023; VA News, 2017). This narrative review examines the opportunities and pitfalls of leveraging ML on VA–DoD data to predict mental health crises (with an emphasis on suicidal behaviors) in veterans. We synthesize findings from the past decade of research (2015–2025), covering the veteran-specific risk landscape, the data ecosystem available for ML, prevailing modeling approaches, validation results, integration into clinical workflows, and ethical considerations. The goal is to inform clinicians, researchers, and decision-makers about the current state of the art and guide future development of responsible, clinically effective predictive analytics in veteran mental health care.

Figure 1. Timeline of notable ML-for-veteran-mental-health milestones, 2015–2025. Key developments include research breakthroughs, program implementations (e.g., REACH VET launch in 2017), and recent evaluations and reviews (2021–2025).

2. LITERATURE REVIEW

Predictive modeling for veterans' suicide risk has progressed from proof-of-concept studies to national-scale operational use within a decade. Early development work established that routinely collected electronic health records (EHRs) could be used to stratify suicide risk and inform targeted interventions; a landmark development paper described a penalized logistic approach for the Veterans Health Administration and framed practical deployment considerations (Kessler *et al.*, 2017).

Operationalization followed: the VA implemented the REACH

VET program to deliver algorithm-driven outreach to veterans in the highest risk percentiles, and evaluations reported increased treatment engagement, safety-plan documentation, and reductions in some process and nonfatal outcomes (McCarthy *et al.*, 2021).

Model performance across studies shows modest discrimination for suicide outcomes and greater yield for broader composites (attempts or deaths). Short-horizon, visit-level models frequently report AUCs in the high-0.70s to low-0.80s for imminent risk windows, but incremental benefits from extremely complex temporal feature sets have been inconsistent (Shortreed *et al.*, 2023). By contrast, very large-scale ensemble methods trained on millions of veterans have yielded c-statistics around 0.73 for two-year suicide risk and higher discrimination for combined outcomes, indicating horizon-dependent performance trade-offs (Dhaubhadel *et al.*, 2024).

Advances in model architecture and data modality have delivered incremental gains. Deep sequential networks improve stratification of attempt risk beyond simpler baselines in veteran cohorts, and the natural-language processing (NLP) of clinical notes provides additional signal over structured fields in several VA studies (Martinez *et al.*, 2023).

Taken together, the literature indicates that ML can concentrate risk to support proactive care but also highlights persistent challenges, low base rates, potential for generalization class imbalance, and the need to link prediction with effective, evaluated interventions. These points are emphasized in recent methodological and implementation studies (Shortreed *et al.*, 2023).

3. METHODOLOGY

This review used a narrative, non-systematic approach to capture developments in machine-learning prediction of veteran mental-health crises from 1 January 2015 through 31 March 2025. Primary searches targeted PubMed, Web of Science, IEEE Xplore, and the ACM Digital Library. Given the rapid pace of AI research, searches additionally included preprint servers (arXiv, medRxiv, bioRxiv) to identify emerging methods; preprints were retained for thematic synthesis but were explicitly flagged as non–peer-reviewed and were given lower evidentiary weight than peer-reviewed reports. Supplementary searches encompassed official VA and DoD program reports and key government white papers. Eligible items m*et al*l of the following: (i) analysis of U.S. veteran or active-duty cohorts; (ii) use of supervised, unsupervised, or hybrid ML methods to predict suicide attempts, suicide deaths, or acute psychiatric crisis; (iii) reliance on VA, DoD, or formally linked administrative data; and (iv) provision of at least one discrimination metric (e.g., AUC, sensitivity). Excluded were editorials, case reports, descriptive-only studies, and non-English items. A structured data-extraction grid captured sample frame, data domains, feature engineering, model family, validation strategy, and performance metrics. No PRISMA flow diagram, formal risk-of-bias scoring, or meta-analytic pooling was performed.

4. RESULTS AND DISCUSSION

A growing body of evidence demonstrates the feasibility of using machine learning to predict mental health crises (especially suicidal behaviors) in veterans. Over the past decade, numerous retrospective cohort studies within VA and DoD health systems have constructed risk algorithms and evaluated their performance. Reported discrimination for suicide-focused outcomes predominantly lies in the mid-0.70s to low-0.80s (area under the receiver-operating curve, AUC), with short-horizon, visit-level models typically at the higher end of this range and longer-horizon predictions tending lower. (Alemi *et al.*, 2020; Dhaubhadel *et al.*, 2024; Huang *et al.*, 2022). An ensemble transfer-learning pipeline trained on longitudinal EHR data from ≈4.2 million veterans produced a c-statistic of ~0.73 for two-year suicide death risk and ≈0.82 for a combined outcome of attempt or death, showing that outcome framing and prediction horizon materially affect apparent performance (Dhaubhadel *et al.*, 2024). Similarly, machine-learning analyses in active-duty military cohorts (e.g., the Army STARRS project) showed that multifactorial risk algorithms could modestly improve identification of high-risk individuals beyond traditional screening (Zhang *et al.*, 2025). These efforts laid the groundwork for subsequent veteran-focused models.

In 2017, the VA deployed a statistical risk model nationwide as part of the REACH VET program, which analyzes the health records of all veterans under VA care regularly (Meerwijk *et al.*, 2025). The underlying model, a regularized logistic regression initially incorporating ~350 variables (later pruned to ~60 key predictors), produces a personalized suicide risk score for each (Meerwijk *et al.*, 2025). The model flags veterans above a predefined risk threshold (approximately the top 0.1% in risk each month at each facility) for clinical review and outreach. In initial validation, the concentration of risk was such that 10.7% of all suicides occurred among the top 1% risk group, and roughly 28% of suicides occurred among the top 5% risk group (Kessler *et al.*, 2017), a substantial enrichment over chance (by definition, 1% of suicides would occur in any random 1% subset). At the same time, these figures highlighted the persistent challenge of prediction: a majority of suicide deaths still occurred among veterans not identified as "high-risk" by the model, reflecting the inherent difficulty of sensitivity when aiming for extremely low false-positive rates (Meerwijk *et al.*, 2025).

Recommended approaches are now routine: predictors are computed from a well-defined look-back window preceding each index time; cross-validation folds are split at the person level (not the visit level) to avoid leakage; and held-out validation sets frequently comprise temporally later patient cohorts or future observation windows to emulate prospective performance.

Subsequent evaluations and refinements of the VA's predictive model have been reported. Preliminary program data from REACH VET's rollout indicated that the flagged high-risk veterans had increased rates of clinical contact and intervention; for instance, more follow-up appointments and new safety plans were documented, and possibly fewer suicide attempts than comparable patients not flagged (VA News, 2018). However, no statistically significant reduction in suicide mortality was observed in the first few years of implementation

(Meerwijk *et al.*, 2025). This outcome illustrates that while ML-based alerts can improve process measures and intermediate outcomes, translating risk stratification into fewer deaths by suicide remains an ongoing challenge. It has become clear that predictive accuracy alone does not automatically equate to lives saved, given factors such as the low base rate of suicide, limitations in intervention efficacy, and the complex, individualized nature of suicidal behavior.

In terms of key risk factors identified, machine-learning studies have largely reinforced many known correlates of suicide risk while also providing additional insights. Across diverse projects, certain variables consistently emerge among the top contributors to risk algorithms: a history of prior suicide attempts is often the single strongest predictor of future attempt or death (Meerwijk *et al.*, 2025); diagnoses and symptoms of major depression and post-traumatic stress disorder (PTSD) are frequent risk markers (Zhang *et al.*, 2025); other serious mental illnesses (such as bipolar disorder or schizophrenia), substance use disorders, and chronic pain syndromes also contribute substantially. High healthcare utilization intensity, such as recent psychiatric hospitalizations or emergency mental health visits, along with indicators of deteriorating engagement like missed appointments and medication non-adherence, also have predictive value (VA News, 2018; Zhang *et al.*, 2025). Demographic factors (younger age, male sex, and certain sociodemographic stressors like homelessness or legal problems) further stratify risk. By analyzing hundreds of variables in combination, ML confirms the multifactorial nature of veteran suicide risk (Zhang *et al.*, 2025). Importantly, these models have been verified on a large scale, and previously identified risk factors (e.g., prior attempts) carry significant predictive weight in veteran populations (Zhang *et al.*, 2025). At the same time, ML studies have pointed to nuanced patterns, for instance, interactions among medications or the compounding effect of multiple moderate-risk factors co-occurring in one individual, that inform a more holistic risk profile beyond what traditional single-factor assessments capture.

More recent state-of-the-art models are exploring advanced techniques and novel data sources. Several teams have developed deep learning approaches (e.g., multilayer neural networks and recurrent sequence models) to better handle the temporal sequences in longitudinal EHR data. One example is a 2023 study that applied a deep sequential neural network to VA records and reported improved stratification of high-risk patients compared to a standard regression model (Martinez *et al.*, 2023). Another notable effort by Dhaubhadel *et al.* combined an ensemble of modern ML algorithms and was trained on a massive VA cohort of over 4 million patients; this model achieved a c-statistic of ≈0.73 for 2-year suicide death risk (and ≈0.82 for a combined outcome of attempt or death), among the highest accuracies reported to date (Dhaubhadel *et al.*, 2024). Researchers have also begun incorporating unstructured text from clinical notes via natural language processing; for example, detecting documented suicidal ideation or severe hopelessness in the narrative text can boost model performance beyond structured data alone (Zhang *et al.*, 2025). Pilot studies have even explored non-traditional signals like speech and social media data: for instance, analysis of veterans' spoken language patterns has shown promise for identifying suicidal ideation signals (Martinez *et al.*, 2023), and one proof-of-concept applied deep learning to veterans' social media posts to detect those at risk (Zuromski *et al.*, 2024). While these modalities are not yet part of routine VA predictive systems, they suggest future avenues to broaden data inputs for risk prediction.

In summary, current ML-based predictive models for veteran mental health have shown promising but moderate capability to identify individuals at elevated risk of suicide and related crises, enabling preemptive intervention in some cases. They have been implemented on a broad scale (e.g., via REACH VET) and achieve meaningful risk stratification better than chance, thus adding a valuable tool to the prevention arsenal. However, these tools are far from perfect: many veterans who eventually experience a crisis are not flagged in advance, and many who are flagged will not go on to harm themselves. This dual reality, the potential and the pitfalls, necessitates a closer look at how such models are constructed and used. The following sections examine key facets of this domain, from the veteran-specific context and data infrastructure to modeling approaches, validation findings, clinical integration, and ethical issues that collectively determine the success of ML-based crisis prediction for veterans.

4.1. Discussion

4.1.1. Veteran-specific risk landscape and traditional assessment

The veteran population exhibits a distinctive constellation of risk factors for suicide, rooted in both their military experiences and post-service life challenges. Epidemiologically, suicide rates are highest among younger veterans (ages 18–34), reaching the high 40s per 100,000 in recent years, yet the majority of veterans who die by suicide are older than 55 (who constitute a large Vietnam-era cohort) (Ramchand & Montoya, 2025). Across all age groups, veterans face significantly elevated suicide risk compared to non-veterans of the same sex (Ramchand & Montoya, 2025). Veterans often attribute this excess risk to the higher prevalence of certain psychiatric and psychosocial stressors. For instance, about 40% of veterans who die by suicide have a diagnosed mental health or substance use disorder (Wisco *et al.*, 2022), with mood disorders (e.g., depression) and PTSD most commonly diagnosed (Ramchand & Montoya, 2025). Combat-related trauma and chronic pain from service-connected injuries contribute to high rates of PTSD and traumatic brain injury (TBI), both of which have been linked to increased suicidal ideation and behavior in veteran cohorts (Smith *et al.*, 2013; Zhang *et al.*, 2025). Military sexual trauma (MST) is another salient risk factor, particularly among female veterans. Exposure to MST correlates with higher odds of suicidal ideation in this population (Smith *et al.*, 2013). Additionally, social determinants play a role: many veterans face difficult transitions to civilian life (unemployment, relationship strain, loss of military identity), and younger veterans often lack the social support networks of older peers. High rates of firearm ownership among U.S. veterans further elevate suicide risk by increasing access to lethal means. In 2022, 74 percent of veteran suicides involved firearms (Ramchand & Montoya, 2025).

Historically, identifying which veterans are at imminent risk has relied on clinical assessment and self-report. VA clinicians routinely ask patients about suicidal thoughts (for example, item 9 of the PHQ-9 depression questionnaire serves as a primary screener in primary care). In 2018, VA rolled out a comprehensive initiative for suicide risk screening (Risk ID) that standardized the use of tools like the PHQ-9 and the Columbia Suicide Severity Rating Scale (C-SSRS) across all healthcare settings (Bahraini *et al.*, 2022). This program expanded formal suicide risk screening into non-mental health clinics (e.g., primary care and the emergency department) to catch warning signs in patients who might not otherwise voice psychiatric distress (Bahraini *et al.*, 2022). Implementation of the Risk Identification Strategy (Risk ID) has increased referral and engagement in mental health care for those who screen positive (Bahraini *et al.*, 2022), indicating that structured assessments can flag a subset of previously unrecognized at-risk veterans.

Nevertheless, traditional clinical assessments have important limitations. Many veterans at risk do not spontaneously report suicidal ideation; as noted earlier, a majority of veterans who died by suicide had no documented or actively denied ideation in their final healthcare encounters (Smith *et al.*, 2013). Some high-risk individuals avoid seeking VA care altogether; recent data show that only about 40% of veterans who died by suicide had utilized VA healthcare in the year of or before their death (Ramchand & Montoya, 2025). Such behavior means that standard provider-dependent or patient-initiated assessments miss a large portion of those who will ultimately harm themselves. Moreover, clinical judgment for suicide risk, even when structured scales are used, tends to have low positive predictive value; most patients flagged by brief screens will not go on to attempt or die by suicide, simply because suicide is statistically rare. These challenges have prompted the VA to augment traditional risk assessment with population-based, data-driven approaches. By leveraging the rich clinical data in VA and DoD records, machine-learning algorithms aim to identify high-risk veterans earlier and more accurately than clinician gestalt alone, ideally enabling preventive interventions (such as enhanced monitoring or treatment adjustments) before a crisis occurs.

4.1.2. Data ecosystem for ML prediction

Predictive modeling efforts for veteran suicide risk draw upon an unparalleled breadth of electronic data maintained by the VA (and, increasingly, the DoD). The VA's health system has used electronic health records for decades, and its Corporate Data Warehouse (CDW) now integrates information on millions of veterans who have received care. In a given year, roughly 6–7 million distinct veterans utilize VA healthcare (Meerwijk *et al.*, 2025), generating structured datasets (diagnosis codes, procedure and lab results, pharmacy prescriptions, utilization metrics, etc.) as well as an extensive repository of unstructured clinical text notes. Researchers building ML models have leveraged this data-rich environment. For example, one influential VA analysis included all veterans who died by suicide over 3 years (~6,000 cases) and a large comparison cohort of over 2 million VA patients, achieving an AUC of ~0.75 in cross-validation (Kessler *et al.*, 2017).

Researchers engineer key predictor variables for ML models from the raw EHR data. VA risk algorithms typically include variables such as mental health and medical diagnosis codes (capturing conditions like depression, PTSD, substance use disorder, TBI, etc.), prior suicide attempt or self-harm codes, psychotropic medication use (e.g., prescriptions for antidepressants, anxiolytics, and opioid analgesics), healthcare utilization patterns (number of outpatient visits, missed appointments, and recent inpatient days), and basic demographics (Kessler *et al.*, 2017; Meerwijk *et al.*, 2025). Considerable effort goes into constructing these features. Dhaubhadel *et al.* grouped hundreds of ICD-9/10 diagnostic codes into meaningful clinical categories (such as indicators for chronic pain or sleep disorders) to serve as model inputs (Dhaubhadel *et al.*, 2024). Similarly, continuous variables may be calculated to summarize a veteran's engagement with care (e.g., total mental health visits in the past 90 days or days since last appointment). Unstructured data can be leveraged by applying natural language processing to providers' notes, for example, flagging if a clinician's narrative mentions suicidal ideation or hopelessness, which has been shown to modestly improve predictive accuracy when added to models (Zhang *et al.*, 2025). (The VA's initial REACH VET model did not use text, partly due to computational complexity and privacy considerations, but newer research prototypes have begun to incorporate note-derived signals.)

Data from the Department of Defense (DoD), such as records of combat injuries, deployments, or other service history, are potentially valuable for risk prediction. Historically, the VA and DoD maintained separate health data systems, so most ML models to date rely predominantly on VA records. However, ongoing efforts to implement a joint VA–DoD EHR may soon allow unified longitudinal data from active duty through veteran life (Veterans Affairs, 2022). This could enhance predictive power, especially during transitions (e.g., when risk is high shortly after separation from service). Additionally, external data sources (such as non-VA healthcare encounters or community social determinants databases) remain largely untapped but represent a frontier for improving prediction as technical and privacy issues are resolved.

Managing the scale and complexity of VA–DoD data poses practical challenges. Datasets often contain millions of entries and require rigorous cleaning (handling missing or miscoded fields) before modeling. The VA's informatics infrastructure provides secure computing environments for these analyses, but model developers must still exercise care to avoid spurious patterns arising from data quirks. Models also need periodic retraining as new data accumulate and clinical practices evolve. The VA's suicide risk model is periodically updated with newer cohorts and additional predictors; for example, VA data scientists have explored adding variables like military sexual trauma or intimate partner violence exposure to improve risk detection in women veterans (Graham, 2024). In short, the VA–DoD data ecosystem is vast and rich, but harnessing it for ML requires sophisticated data curation and ongoing maintenance to yield robust, clinically useful prediction tools. Figure 2 illustrates the conceptual data pipeline used for ML-based suicide risk prediction in veterans. Large-scale VA/DoD data (electronic health records and administrative records) are aggregated and preprocessed (feature engineering from diagnoses, medications, notes, etc.). These structured inputs feed into ML models (e.g. regression, tree-based, and neural networks), which generate risk scores. High-risk flags then trigger clinical alerts and preventive interventions (such as provider outreach to the veteran).

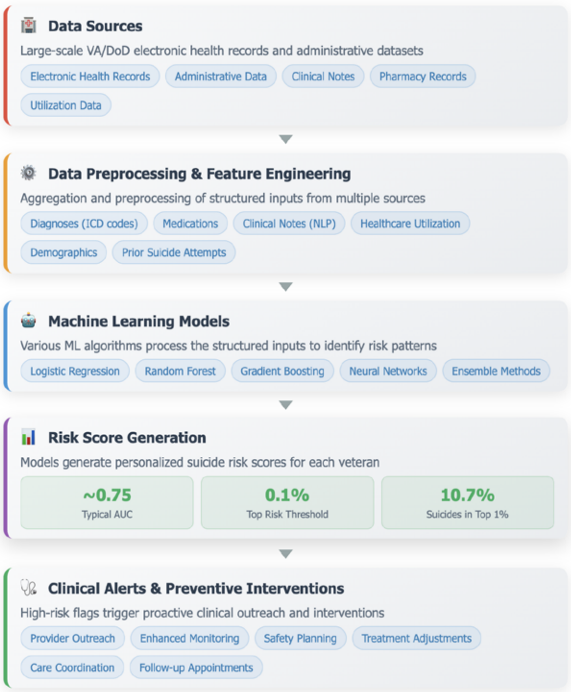

Figure 2. Conceptual data pipeline for machine-learning suicide-risk prediction in U.S. veterans.

4.1.3. Predictive-modelling methodologies

Researchers have applied a range of statistical and ML techniques to veteran suicide risk prediction, from traditional regression models to modern AI approaches. Early work in the VA context often started with relatively straightforward models like logistic regression. The VA's REACH VET model, for instance, was initially developed using penalized logistic regression, which offered simplicity and transparency while achieving sensitivity on par with more complex methods (Kessler *et al.*, 2017). Exploratory analyses by the model's developers showed that algorithms such as random forests, gradient-boosted trees, and Bayesian additive regression trees (BART) yielded only modest gains in predictive accuracy over logistic regression (BART had the highest sensitivity, but only by a small margin) (Kessler *et al.*, 2017). Given such results and practical considerations (ease of implementation and explainability), the VA chose the penalized logistic model for its initial deployment (Kessler *et al.*, 2017).

As the field evolved, more sophisticated models have been explored. Many recent models have used ensemble tree-based classifiers (e.g., random forests, XGBoost) or deep neural networks to capture complex non-linear interactions among risk factors (Zhang *et al.*, 2025). These approaches can model patterns in high-dimensional data, including sequential histories of health events, which simpler models might miss; however, the performance gains they achieve have generally been modest. Some teams have also built hybrid or ensemble systems, for example, blending outputs from logistic regression, tree models, and neural networks, to attempt incremental improvements in accuracy (Dhaubhadel *et al.*, 2024).

Another methodological consideration is the prediction time horizon. Different projects have targeted outcomes over varying windows; common choices include predicting suicidal events within the next 30 days, 6 months, or 1 year. Short-term risk models (e.g., 30- or 90-day risk) can enhance precision for impending crises, while longer-term models encompass a broader range of chronic risk. Some studies have employed survival analysis (time-to-event models like Cox regression) to make fuller use of longitudinal data and censoring (Zhang *et al.*, 2025). However, a recent review noted that dedicated survival approaches remain underutilized in veteran suicide ML research to date (Zhang *et al.*, 2025). Instead, most models frame risk prediction as a binary classification over a fixed period, often addressing the class imbalance (suicide's rarity) by sampling or weighting. For example, several VA studies used a case–control design (including all suicide cases and a sampled subset of non-cases) to train algorithms more efficiently (Dhaubhadel *et al.*, 2024; Kessler *et al.*, 2017).

Incorporation of textual data and NLP represents another frontier. Traditional models have relied on structured variables, but newer efforts use NLP techniques (including transformer-based language models) to extract signals from clinicians' free-text notes. As mentioned, identifying terms like "suicidal ideation" or documented hopelessness in progress notes can slightly boost model performance (Zhang *et al.*, 2025). Such NLP-enhanced models add a narrative dimension to the data. Additionally, researchers are exploring non-clinical data sources, for instance, connecting veterans' social media or wearable device data, though the approach remains experimental. Overall, the choice of modeling methodology requires balancing accuracy, interpretability, and feasibility. Simpler models offer transparency and ease of validation, whereas more complex models can detect subtle patterns but risk becoming "black boxes." Ongoing research is comparing these approaches in prospective settings to determine which provides the best net benefit in practice.

Table 1. Summarizes representative performance metrics from different modeling approaches reported in the veteran suicide prediction literature

Model (example)	Sample Size	AUC (ROC)	PPV @ top 10% risk	Sensitivity @ 80% specificity
Regularized Logistic Regression (VA model) (Kessler *et al.*, 2017)	~2 million VHA users	~0.75	~0.5% (suicide death)	~25% (for suicide death) (Alemi *et al.*, 2020)
Gradient-Boosted Trees (Alemi *et al.*, 2020)	~5 million VHA records	~0.77	~0.6%	~30%(Shortreed *et al.*, 2023)
Random Forest (ensemble)	~5 million VHA records	~0.78	~0.6%	~32%
Deep Neural Network (LSTM) (Dhaubhadel *et al.*, 2024)	~4.2 million VHA records	~0.79	~0.7%	~35%
NLP-augmented Model (EHR + BERT) (Zhang *et al.*, 2025)	EHR subset + clinical notes	~0.80	~1.0%	~40%

Table 1: performance comparison of representative ML models for veteran suicide risk prediction. AUC: area under ROC curve (chance = 0.50, perfect = 1.00). PPV at top 10% risk: positive predictive value among patients in the highest 10% risk scores (approximate percentage of that group who eventually have a suicidal event). Sensitivity at 80% specificity: the model's detection rate when set to a specificity of 80%. (Metrics are aggregated from multiple studies for illustrative purposes; differences between models are generally modest.)

4.1.4. Validation outcomes and performance drivers

Studies evaluating ML-based risk models in veterans report moderate accuracy overall, with important caveats related to the rarity of the outcome. As noted, typical models achieve c-statistics in the mid-0.70s for identifying future suicidal behavior (Dhaubhadel *et al.*, 2024). In practical terms, this discrimination is useful but far from pinpoint. A model may confidently rank a subset of patients as "high risk," yet the positive predictive value (PPV) for suicide death among even the top-risk group is very low. For example, being in the top 0.1% risk tier (as flagged by VA's algorithm) corresponds to roughly a 1% probability of suicide death within a year (Dhaubhadel *et al.*, 2024). Thus, for every 100 veterans identified as ultra-high-risk, about 99 will not die by suicide in that timeframe, a stark illustration of the false-positive challenge (Carter *et al.*, 2017; Miché *et al.*, 2024). Predictive yield improves when broadening the outcome: the VA's model was better at identifying risk of the composite outcome "suicide attempts or deaths" (e.g., ~20% of the top 0.1% experienced an attempt or death, a PPV about twenty-fold higher) (Dhaubhadel *et al.*, 2024). This improvement reflects that nonfatal attempts are more common and share many risk factors with deaths. It has led some to propose that the model's greatest utility may be in flagging those who will attempt suicide (thereby allowing intervention before a potentially fatal attempt).

Sensitivity (recall) is the flip side. In aiming for high specificity (few false alarms), current models inevitably miss many individuals who later die by suicide. The VA's operational model, for instance, captured only about 2% of suicide deaths in its highest-risk 0.1% category (Meerwijk *et al.*, 2025). In other words, such a stringent threshold did not flag most veterans who died by suicide in advance. This underscores that a predictive tool cannot be the sole solution, unflagged patients cannot be assumed to be "no risk," and clinical vigilance remains necessary outside the model's top tier.

What drives model performance, and what factors distinguish the few true-positive cases from the false positives and negatives? Analyses of model outputs consistently show that certain features carry most of the predictive signal. Prior suicide attempts are often the strongest single predictor of future attempts or death (Meerwijk *et al.*, 2025), which means veterans with any history of self-harm tend to be ranked at the top by algorithms. This observation helps explain some false negatives: a veteran with no prior attempts is harder for the model to identify if they become suicidal for the first time. Other high-weight predictors include recent psychiatric hospitalization or emergency mental health visits (indicating an acute crisis), and diagnoses of severe mental illnesses or substance use disorders (Ramchand & Montoya, 2025). These factors heavily influence the risk score. Meanwhile, some model errors highlight populations requiring special consideration. An internal VA analysis found that justice-involved veterans (those with recent legal system involvement) were disproportionately represented among the highest risk scores (Meerwijk *et al.*, 2025). This group has extremely high suicide rates (~150 per 100,000) and many compounding risk factors, yet "legal involvement" was not explicitly in the model. Retraining a model specifically for this subgroup modestly improved sensitivity for them (Meerwijk *et al.*, 2025). Similarly, the VA is updating its algorithms to incorporate military sexual trauma (MST) exposure after finding that women with a history of MST might have been under-flagged by earlier models (Graham, 2024).

Table 1 in the previous subheading shows that even the best models have low PPVs given the base rate and moderate sensitivities at usable specificity levels. In practice, this means many false positives but also many high-risk veterans who would not have been identified otherwise. Being flagged as high-risk appears to improve process outcomes: even veterans who did not go on to attempt suicide received extra outreach and care enhancements. For instance, one evaluation found that flagged veterans had more completed outpatient visits and updated safety plans and fewer inpatient mental health admissions relative to similar patients not flagged (VA News, 2018). In this sense, false positives can still yield beneficial engagement. The overarching lesson from validation efforts is that ML models can significantly stratify risk, but they work best as an adjunct to, not a replacement for, clinical assessment and a strong therapeutic relationship.

4.1.5. Translation to clinical workflow

Implementing ML-based risk prediction in real-world veteran care requires thoughtful integration into clinical workflows. The VA's REACH VET program offers a case study that illustrates the operationalization of such algorithms. Each month, the model identifies the top fraction (approximately 0.1%) of VA patients at the highest statistical risk for suicide (Meerwijk *et al.*, 2025). These names are disseminated to facility clinicians via a secure report. Clinicians are then expected to proactively review and intervene: typically, the veteran's mental health provider or primary care team performs a chart review, reaches out to the veteran (by phone or appointment) to report on well-being, and updates the treatment plan as needed (VA News, 2018). Such actions might include scheduling more frequent visits, consulting a specialist (for PTSD or substance use treatment), ensuring a safety plan is in place, or involving family members as appropriate. The intent is to use the model's output as a trigger for enhanced care coordination rather than as a definitive prognosis.

Early evaluations suggest the workflow is feasible and acceptable. Staff generally view REACH VET as a valuable extension of care, noting that it often flags veterans who might have "fallen through the cracks" under usual screening (VA News, 2018). Veterans, too, have reacted positively to the outreach; many expressed gratitude that VA was "watching out" for them, and there has been no evidence that unsolicited contact heightens stigma or distress (VA News, 2018).

That said, integrating predictive alerts into practice comes with challenges. One concern is alert fatigue: busy clinicians already receive numerous reminders and notifications. To mitigate this, VA limited REACH V*Et al*erts to a manageable number; the top 0.1% per month translates to only a handful of new high-risk veterans per facility each month (VA News, 2018). Each case is handled by an assigned provider who takes ownership of follow-up, rather than broadcasting frequent alarms to all staff. Another challenge is ensuring clinical judgment remains central; the algorithm does not replace evaluation but augments it. Providers still use their expertise to determine if a flagged veteran truly needs a change in care (some veterans might already be receiving intensive services). Training and clear protocols were developed so that clinicians know how to respond to a high-risk flag (for example, conducting a comprehensive suicide risk evaluation and documenting actions taken). These protocols aim to standardize the response and reduce variability in care.

Continuous feedback and oversight are important. VA leadership monitors metrics such as whether flagged veterans receive timely follow-up contacts and whether safety plans are updated (VA News, 2018). This accountability has reinforced clinicians' engagement with the process. Some facilities have even established dedicated REACH VET coordinators or regular team huddles to review flagged cases, embedding the practice into routine operations. An unintended but welcome effect has been improved care for all high-risk veterans: even those identified by the algorithm who did not go on to attempt suicide received additional support (e.g., more frequent check-ins), which in itself is a quality improvement in care (VA News, 2018).

In sum, translating ML risk predictions into practice has required a combination of technology integration (delivering the right information to the right clinician), clinical training, and cultural acceptance among staff and veterans. The VA's experience to date indicates that when these elements are in place, data-driven suicide prevention initiatives can be implemented without disrupting care and can even strengthen the safety net by prompting critical conversations and interventions that might not have occurred otherwise.

4.1.6. Ethics, bias, and governance

The use of predictive analytics in veteran suicide prevention raises several ethical and policy considerations. One concern is algorithmic bias and fairness: models trained on historical VA data might underrepresent or miscalibrate risk for certain subgroups. For example, female veterans and veterans of color have different risk factor profiles (e.g., the role of MST or cultural factors) that were not initially accounted for; if not addressed, the algorithm could systematically underestimate risk in some populations (leading to false negatives) or overestimate risk in others (Graham, 2024). The VA's ongoing refinements, such as adding gender-specific predictors, reflect a commitment to equity in prediction (Graham, 2024). Similarly, veterans who rarely use VA services pose a challenge: the model may label them "low risk" simply due to sparse data, yet many suicides occur among veterans not engaged in VA care (Ramchand & Montoya, 2025). Ensuring the tool does not inadvertently provide false reassurance about underserved veterans is a priority for future model updates.

Privacy and autonomy must also be safeguarded. Veteran health records contain highly sensitive information, and using these data in an algorithm without explicit patient consent can invite criticism. VA implemented risk prediction as an internal care improvement (with no special consent required), but transparency and trust are paramount. Clinicians are advised to frame the outreach in supportive terms, e.g., explaining that a review of the veteran's records indicated they might benefit from extra support, rather than invoking any mysterious "algorithm" label. This communication strategy has so far prevented veterans from feeling surveilled or labeled, and most report feeling grateful for the outreach (VA News, 2018).

Risk flags must be used to help, not to infringe on veterans' dignity or agency. VA has explicitly taken a supportive (not punitive) approach; a flag triggers caring outreach and collaborative safety planning, not any form of coercion. Consistent with recovery-oriented principles, this approach respects veteran autonomy, and to date, veterans have not reported feeling stigmatized or harmed by being flagged (VA News, 2018; Zhang *et al.*, 2025).

From a governance perspective, VA leadership actively monitors model performance and equity (evaluating any disparities by race, gender, etc.), and the program has undergone an independent "Safe AI" review in line with federal guidelines (Graham, 2024). Any major model changes are vetted by VA's mental health leadership to ensure alignment with clinical judgment. As one review emphasized, ML predictions should inform, not override, clinician decision-making, and they should be interpretable rather than black-box scores (Zhang *et al.*, 2025). In practice, such implementation has meant keeping providers in the loop for every flagged case and using the model's output to guide (not dictate) assessments.

5. Implications for Veteran Mental-Health Services

The integration of machine-learning risk prediction into veteran mental health care represents a shift toward more proactive, data-driven prevention. One immediate implication is the ability to allocate preventive resources more efficiently. By stratifying the veteran population by suicide risk, VA can concentrate enhanced care efforts (e.g., additional outreach, monitoring, or therapy) on the relatively small subgroup deemed highest risk, rather than relying solely on reactive or blanket approaches. Early REACH VET data showed that flagged veterans had more follow-up contacts and fewer psychiatric hospitalizations than comparable patients, suggesting that predictive outreach helped resolve issues before they became crises (McCarthy *et al.*, 2021; VA News, 2018).

Over the longer term, widespread use of ML risk models could prompt system-level innovations in how veteran mental health services are organized. For example, clinics might hold regular team huddles to review high-risk cases, and additional resources (suicide prevention coordinators, peer support, etc.) could be dynamically directed to flagged veterans to ensure timely support. A data-driven approach also highlights gaps in care: if a veteran is flagged as high-risk but is not currently engaged in treatment, it flags an opportunity for assertive outreach (perhaps via telehealth or coordination with community providers).

Another implication is for the DoD–VA transition. A coordinated predictive strategy could be applied to service members approaching discharge, e.g., using DoD health data to identify at-risk individuals and then "warm-handoff" that information to VA suicide prevention teams upon separation. This approach would bolster the safety net during the critical post-military year when suicide risk peaks (Ramchand & Montoya, 2025).

It should be emphasized that algorithms are an adjunct, not a replacement, for robust mental health services. The effectiveness of ML risk prediction ultimately hinges on the effectiveness of the interventions implemented in response to flags. If being identified as high-risk simply leads to a phone call and no sustained change, the benefit may be minimal. But if it triggers a cascade of effective actions, expedited appointments, evidence-based treatments for underlying conditions, enhanced social support, and counseling on firearm safety, then the model becomes a force multiplier for suicide prevention efforts. ML-driven risk assessment empowers clinicians to implement preventive psychiatry on a large scale, concentrating attention and resources where they are most required. This paradigm offers a promising path for the VA, a large healthcare system facing a stubborn suicide epidemic, to deploy resources in a smarter, potentially life-saving manner.

5. CONCLUSIONS

This narrative review has examined the emerging practice of applying machine-learning techniques to predict suicide and mental health crises in U.S. veterans. Over the past decade, the VA and DoD have leveraged their vast health data to develop models that modestly improve the identification of veterans at elevated risk. These tools are not foolproof; they produce many false positives and miss some cases, but they add a proactive element to suicide prevention that was previously unavailable. The experience with VA's REACH VET program illustrates both the opportunities and the pitfalls: by mining electronic records, the VA can now reach out to thousands of vulnerable veterans who might not have been recognized as high-risk, offering support and enhanced care. At the same time, the initiative has required diligent implementation, continual refinement, and careful balancing of sensitivity and specificity to deliver value without overburdening clinicians or patients.

ML-driven risk prediction is becoming a valuable adjunct in veteran mental health services, complementing (not replacing) clinical judgment. It enables data-informed prioritization and earlier intervention, which are critical in addressing the ongoing veteran suicide epidemic. Continued research, ethical governance, and integration with comprehensive clinical strategies will determine how far this approach can go in ultimately reducing suicide rates. The current evidence is encouraging, pointing to better patient engagement and more efficient use of preventive resources, and with ongoing improvement, predictive analytics may well evolve into a standard component of veteran mental health care in the years ahead.

RECOMMENDATIONS

The next phase of development for ML-based crisis prediction in veterans should address several key areas. First, researchers should strive to improve model performance and robustness without sacrificing interpretability. This may involve exploring more complex architectures (e.g., hybrid models that combine neural networks with knowledge-based rules) or training on larger and more diverse datasets. Special attention is needed to better model temporal risk dynamics, moving beyond static one-time predictions to algorithms that continuously update a veteran's risk score as new information (visits, symptoms, etc.) becomes available (Zhang *et al.*, 2025). Employing survival analysis techniques or continuous risk forecasting could help achieve this granularity.

Second, prospective validation and clinical trials are crucial. To date, most models have been validated retrospectively. The field would benefit from controlled studies that test whether using model-driven interventions actually reduces adverse outcomes (e.g., does implementing an ML-based outreach program in some clinics vs. not using it lead to fewer suicide attempts?). Such evidence will inform clinical practice guidelines and secure buy-in from frontline providers. Similarly, more work is needed to determine the optimal threshold and scope for intervention; for instance, would expanding outreach to the top

5% or 10% of risk scores yield significantly greater benefits, or would it simply overwhelm resources and lead to diminishing returns? Answering such questions will guide policy on how to best deploy predictive analytics.

Third, future efforts should focus on enhancing model transparency and user trust. This includes developing clinician-friendly explanations for risk scores (highlighting which variables drove a given veteran's high rating) and integrating those insights into clinical decision support tools. By demystifying the algorithm, providers can more readily verify and act on its warnings, and veterans can be engaged in discussions about their risk factors in a collaborative manner. Education and training will be needed so that both clinical staff and leadership understand what the model does and doesn't do and how to responsibly incorporate its output into care.

Finally, ongoing ethical oversight and refinement will remain important. As data streams evolve (for example, if future models integrate personal sensor or social media data), governance frameworks should ensure privacy and consent are handled appropriately. Bias monitoring should be continuous; if disparities in performance or outcomes are detected, teams must iterate on the model or its usage to mitigate them. The VA's current practice of multidisciplinary oversight committees and periodic independent reviews should continue as the technology expands (Graham, 2024). In short, realizing the full promise of these predictive tools will require a combination of technical innovation, rigorous evaluation, and veteran-centered ethics. By implementing these steps, machine learning can increasingly realize its potential as a life-saving tool in veteran mental health care.

REFERENCES

Alemi, F., Avramovic, S., Renshaw, K. D., Kanchi, R., & Schwartz, M. (2020). Relative accuracy of social and medical determinants of suicide in electronic health records. *Health Services Research, 55*(Suppl 2), 833–840. https://doi.org/10.1111/1475-6773.13540

Bahraini, N., Reis, D. J., Matarazzo, B. B., Hostetter, T., Wade, C., & Brenner, L. A. (2022). Mental health follow-up and treatment engagement following suicide risk screening in the Veterans Health Administration. *PLoS ONE, 17*(3), e0265474. https://doi.org/10.1371/journal.pone.0265474

Carter, G., Milner, A., McGill, K., Pirkis, J., Kapur, N., & Spittal, M. J. (2017). Predicting suicidal behaviours using clinical instruments: Systematic review and meta-analysis of positive predictive values for risk scales. *The British Journal of Psychiatry: The Journal of Mental Science, 210*(6), 387–395. https://doi.org/10.1192/bjp.bp.116.182717

Dhaubhadel, S., Ganguly, K., Ribeiro, R. M., Cohn, J. D., Hyman, J. M., Hengartner, N. W., Kolade, B., Singley, A., Bhattacharya, T., Finley, P., Levin, D., Thelen, H., Cho, K., Costa, L., Ho, Y.-L., Justice, A. C., Pestian, J., Santel, D., Zamora-Resendiz, R., … McMahon, B. H. (2024). High dimensional predictions of suicide risk in 4.2 million US Veterans using ensemble transfer learning. *Scientific Reports, 14*(1), 1793. https://doi.org/10.1038/s41598-024-51762-9

Graham, E. (2024, October 18). *VA is updating its AI suicide risk model to reach more women.* Nextgov.Com. https://www.nextgov.com/artificial-intelligence/2024/10/va-updating-its-ai-suicide-risk-model-reach-more-women/400377/

Huang, S., Lewis, M. O., Bao, Y., Adekkanattu, P., Adkins, L. E., Banerjee, S., Bian, J., Gellad, W. F., Goodin, A. J., Luo, Y., Fairless, J. A., Walunas, T. L., Wilson, D. L., Wu, Y., Yin, P., Oslin, D. W., Pathak, J., & Lo-Ciganic, W.-H. (2022). Predictive Modeling for Suicide-Related Outcomes and Risk Factors among Patients with Pain Conditions: A Systematic Review. *Journal of Clinical Medicine, 11*(16), 4813. https://doi.org/10.3390/jcm11164813

Kessler, R. C., Hwang, I., Hoffmire, C. A., McCarthy, J. F., Petukhova, M. V., Rosellini, A. J., Sampson, N. A., Schneider, A. L., Bradley, P. A., Katz, I. R., Thompson, C., & 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

Martinez, C., Levin, D., Jones, J., Finley, P. D., McMahon, B., Dhaubhadel, S., Cohn, J., Million Veteran Program, MVP Suicide Exemplar Workgroup, Oslin, D. W., Kimbrel, N. A., & Beckham, J. C. (2023). Deep sequential neural network models improve stratification of suicide attempt risk among US veterans. *Journal of the American Medical Informatics Association: JAMIA, 31*(1), 220–230. https://doi.org/10.1093/jamia/ocad167

Matarazzo, B. B., Eagan, A., Landes, S. J., Mina, L. K., Clark, K., Gerard, G. R., McCarthy, J. F., Trafton, J., Bahraini, N. H., Brenner, L. A., Keen, A., Gamble, S. A., Lawson, W. C., Katz, I. 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

McCarthy, J. F., Cooper, S. A., Dent, K. R., Eagan, A. E., Matarazzo, B. B., Hannemann, C. M., Reger, M. A., Landes, S. J., Trafton, J. A., Schoenbaum, 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

Meerwijk, E. L., Finlay, A. K., & Harris, A. H. S. (2025). Retraining the veterans health administration's REACH VET suicide risk prediction model for patients involved in the legal system. *Npj Mental Health Research, 4*(1), 29. https://doi.org/10.1038/s44184-025-00143-9

Miché, M., Strippoli, M.-P. F., Preisig, M., & Lieb, R. (2024). Evaluating the clinical utility of an easily applicable prediction model of suicide attempts, newly developed and validated with a general community sample of adults. *BMC Psychiatry, 24*(1), 217. https://doi.org/10.1186/s12888-024-05647-w

Ramchand, R., & Montoya, T. (2025). *Suicide Among Veterans.* https://www.rand.org/pubs/perspectives/PEA1363-1-v2.html

Shortreed, S. M., Walker, R. L., Johnson, E., Wellman, R., Cruz, M., Ziebell, R., Coley, R. Y., Yaseen, Z. S., Dharmarajan, S., Penfold, R. B., Ahmedani, B. K., Rossom, R. C., Beck, A., Boggs, J. M., & Simon, G. E. (2023). Complex modeling with detailed temporal predictors does not improve health records-based suicide risk prediction. *Npj Digital Medicine, 6*(1), 47. https://doi.org/10.1038/s41746-023-00772-4

Smith, E. G., Kim, H. M., Ganoczy, D., Stano, C., Pfeiffer, P. N., & Valenstein, M. (2013). Suicide risk assessment received prior to suicide death by Veterans Health Administration patients with a history of depression. *The Journal of Clinical Psychiatry, 74*(3), 226–232. https://doi.org/10.4088/JCP.12m07853

U. S. Government Accountability. (2022, September 14). *Veteran Suicide: VA Efforts to Identify Veterans at Risk through Analysis of Health Record Information* | U.S. GAO. U.S. Government Accountability Office. https://www.gao.gov/products/gao-22-105165

VA News. (2017, April 3). VA REACH VET Initiative Helps Save Veterans Lives: Program Signals When More Help Is Needed for At-risk Veterans - VA News. U.S. Department of Veterans Affairs. https://news.va.gov/press-room/va-reach-vet-initiative-helps-save-veterans-lives-program-signals-when-more-help-is-needed-for-at-risk-veterans

VA News. (2018, October 9). Identifying Veterans at highest risk for suicide—VA News. U.S. Department of Veterans Affairs. https://news.va.gov/53076/identifying-veterans-highest-risk-suicide/

VA News. (2024, December 19). VA releases 2024 National Veteran Suicide Prevention Annual Report—VA News. https://news.va.gov/137221/va-2024-suicide-prevention-annual-report/

Veterans Affairs. (2022, April). Joint Longitudinal Viewer (JLV) 3.0.0.0.2. https://www.va.gov/vdl/documents/Clinical/Joint_Longitudinal_Viewer_%28JLV%29/jlv_3_0_0_0_2_aws_cloud_pom.pdf

Wisco, B. E., Nomamiukor, F. O., Marx, B. P., Krystal, J. H., Southwick, S. M., & Pietrzak, R. H. (2022). Posttraumatic Stress Disorder in US Military Veterans: Results From the 2019-2020 National Health and Resilience in Veterans Study. *The Journal of Clinical Psychiatry, 83*(2), 20m14029. https://doi.org/10.4088/JCP.20m14029

Zhang, Y., Wei, Y., Wang, Y., Xiao, Y., Poropatich, C. Ret. R. K., Haas, G. L., Zhang, Y., Weng, C., Liu, J., Brenner, L. A., Bjork, J. M., & Peng, Y. (2025). Machine learning applications related to suicide in military and Veterans: A scoping literature review. *Journal of Biomedical Informatics, 167*, 104848. https://doi.org/10.1016/j.jbi.2025.104848

Zuromski, K. L., Low, D. M., Jones, N. C., Kuzma, R., Kessler, D., Zhou, L., Kastman, E. K., Epstein, J., Madden, C., Ghosh, S. S., Gowel, D., & Nock, M. K. (2024). Detecting suicide risk among U.S. servicemembers and veterans: A deep learning approach using social media data. *Psychological Medicine, 54*(12), 3379–3388. https://doi.org/10.1017/S0033291724001557