

Early Risk Identification in Veteran Primary Care

Connecting fragmented clinical information to support PTSD risk assessment and clinician follow-up

Research collaboration involving MIT and the U.S. Department of Defense

Nucleus's role: Data pipelines, risk scoring, recommendations, and EMR integration

Overview

As part of a research collaboration involving MIT and the U.S. Department of Defense, Nucleus supported the development of a scalable approach to identifying Veterans who could benefit from further clinical assessment. The work focused on connecting fragmented patient information, reducing manual research, and returning patient-level risk scores and recommendations to the clinical environment.

Nucleus combines regulated data pipelines with electronic medical record (EMR) integration to bring existing health records together with additional data sources for patient-specific research. The resulting workflow supported earlier identification of potentially at-risk patients and gave providers a clearer basis for planning follow-up care.

The Challenge

The collaboration addressed concerns about self-harm among returning Veterans and the potential role of undiagnosed post-traumatic stress disorder (PTSD). Identifying patients who might need further assessment required reviewing signals distributed across encounter documentation, recordings, and existing health records.

Manual review was time-consuming and difficult to scale. Providers needed to locate relevant information, interpret it in the context of the patient's history, and determine appropriate follow-up. Fragmentation made it harder to establish a consistent view of each patient and contributed to delays between reviewing information and taking the next step.

The central challenge was to create a repeatable process for surfacing potential risk while fitting into established Veteran primary care workflows. The approach needed to make information more accessible and actionable without requiring providers to conduct extensive manual research for every patient.

What Nucleus Delivered

Nucleus developed an end-to-end nightly processing pipeline that transformed EMR exports into patient-level PTSD likelihood scores. Each score was paired with recommendations to support clinician review and follow-up.

The workflow brought together data ingestion, normalization, feature preparation, classification, and delivery of results. Rather than leaving outputs in a separate research environment, Nucleus returned them in an integration-friendly CSV format for ingestion into the client's EMR environment.

This connected analytical processing with an operational path for using the results. Providers could receive information relevant to follow-up within the client's existing workflow, helping limit disruption and reducing the effort required to assemble patient information manually.

The PTSD likelihood score served as a signal for further assessment. It should be distinguished from a confirmed diagnosis or a direct prediction of self-harm; clinical interpretation remained essential to determining the appropriate response.

How the Workflow Operated

Data ingestion and preparation. The nightly batch process ingested source data from an S3 or SFTP landing zone. Automated parsing and normalization prepared the incoming records for consistent downstream processing.

Patient-level analysis. A secure patient store held the prepared information. Curated features fed a classification service that generated a PTSD likelihood score for each patient processed.

Recommendations and delivery. The workflow paired scores with recommendations for clinician follow-up and returned the results as CSV files to the client environment. These outputs supported EMR ingestion and provider notification.

The illustrative processing sequence was:

EMR export, S3/SFTP landing zone, Parsing and normalization, Secure patient store and curated features, Classification service, Score and recommendations, CSV return for EMR ingestion

This batch architecture provided a repeatable way to process incoming information and surface results for review. Its nightly cadence supported ongoing follow-up preparation rather than real-time emergency monitoring.

Reported Impact

The structured processing approach supported more consistent identification of patients requiring review, compared with relying solely on manual examination of fragmented records. Returning outputs to the client's environment helped make the findings available through established operational processes.

Capabilities Demonstrated

The collaboration illustrates Nucleus's ability to connect regulated data processing with practical clinical workflows. Its contribution combined patient-level risk scoring, actionable recommendations, and operational integration through CSV and EMR interfaces.

The broader value was a clearer path from fragmented information to clinician review: reducing the effort required to find relevant signals and supporting timely, informed follow-up for Veteran patients.