Sepsis Prevention System

A Clinical Decision Support System to Reduce the Rate of Sepsis Among Hospital Inpatients

Ardak Baizhaxynova, Pallavi Rajan, Santiago Enriquez
April 28, 2025

Abstract

Sepsis—the body's extreme response to infection—afflicts hundreds of thousands of hospital inpatients each year and drives significant mortality. Rapid identification and intervention are essential, yet many clinical workflows lack accessible, data-driven decision support. Our prototype DSS stores patient vitals and laboratory results in a PostgreSQL relational database and mimics electronic health record integration by allowing data entry through a web form. Upon entry, an XGBClassifier model, trained on hourly-granular data, calculates a sepsis risk score using features such as heart rate, respiratory rate, blood pressure, temperature, age, white blood cell count, creatinine, bilirubin, platelets, and lactate. The system ranks all admitted patients by risk score and visualizes the overall distribution via a histogram. Built with Streamlit, the front end enables editing of patient records and on-the-fly risk recalculation. By delivering clear risk rankings and visual summaries, our DSS supports care teams in prioritizing high-risk patients, improving ICU resource allocation, and ultimately reducing sepsis-related morbidity and mortality.

Contents

Background and Rationale

Sepsis Overview & Stakeholders

Around 850,000 cases of sepsis are encountered in emergency departments every year in the United States. Sepsis is the body's extreme response to an infection, which requires rapid treatment to avoid undesired consequences for patients, such as tissue damage, organ failure, or even death. People are vulnerable to getting infections, and almost any infection can lead to sepsis. A rapid identification and initiation of the required measures is necessary to optimize the outcomes of patients with sepsis. For that reason, developing decision support systems (DSS) to assist clinicians in the early diagnosis of sepsis is highly desirable, leading to appropriate and timely treatment of patients. There are some reports of automated sepsis alert systems that have been developed and implemented to improve compliance with sepsis guidelines and protocols, showing a decreased length of hospital stay and even in-hospital mortality (Uffen et al., 2021).

Sepsis affects approximately 1.7 million adults in the U.S. each year and is associated with over 250,000 deaths (Rhee et al., 2017). Studies estimate that it contributes to 30% to 50% of hospital deaths (Liu et al., 2014). Sepsis advances quickly, making timely intervention crucial for reducing mortality. Studies indicate that many sepsis-related deaths are preventable with better care, highlighting the importance of predictive analytics in improving outcomes. Machine learning shows promise in aiding early detection, complementing clinical diagnosis. Key stakeholders impacted by this solution include ICU physicians, bedside nurses, hospital administrators, and patients, whose outcomes depend on early recognition and treatment of sepsis. The objective of this report is to introduce a DSS that predicts sepsis through a machine learning model, enabling frequent risk assessment for ICU patients, assisting clinicians to prioritize high-risk patients, and optimizing resource allocation to reduce the rate of sepsis in hospitals.

Existing DSS Approaches

Sepsis prediction and management have become key focus areas in clinical decision support system (DSS) development, given the high morbidity, mortality, and cost associated with sepsis. In recent years, several DSS solutions have emerged with the goal of enabling earlier recognition and intervention. However, while some notable successes exist, significant technological and clinical gaps remain.

One of the most prominent examples of a sepsis-focused DSS is the Targeted Real-time Early Warning System (TREWS)(Henry et al., 2022), developed by Bayesian Health in collaboration with Johns Hopkins University. TREWS uses machine learning ensemble models trained on large-scale electronic health record (EHR) data to identify patients at risk of developing sepsis. TREWS represents a significant advancement over traditional rules-based sepsis alert systems by capturing non-linear relationships between clinical variables and sepsis onset. In clinical deployments, TREWS demonstrated improved early detection rates compared to manual screening methods, contributing to faster treatment initiation and improved patient outcomes. Despite these strengths, studies revealed that TREWS alerts were more likely to be dismissed during night shifts, when staffing levels are lower and the risks associated with delayed recognition are higher. Additionally, TREWS's effectiveness is diminished when patients present with atypical symptoms, leading providers to rely on clinical heuristics rather than trust the system’s recommendations. Furthermore, TREWS requires site-specific recalibration prior to deployment in each hospital, limiting its scalability and ease of adoption.

Another notable example is Prenosis' Sepsis Identification Algorithm, an FDA-approved solution utilizing random forest machine learning techniques. Prenosis’ model leverages high-impact physiological features such as heart rate, respiratory rate, and temperature to predict sepsis risk. The algorithm has demonstrated strong performance in clinical validation studies and is designed to integrate seamlessly into existing hospital IT infrastructure. Despite its approval and success, Prenosis also faces challenges common to machine learning-based healthcare applications: difficulties in model explainability, variability in performance across different patient populations, and dependence on high-quality, real-time data that is not always uniformly available across healthcare systems.

Overall, while current sepsis DSS models have significantly advanced early detection capabilities compared to historical methods, important limitations persist. Many systems still struggle with user trust and adoption, particularly during periods of atypical clinical presentation or staff shortages. Workflow integration also remains a critical barrier; alerts that are not seamlessly incorporated into clinician routines tend to be overlooked or ignored. Finally, most systems require time-intensive customization for each deployment site, reducing their ability to scale across diverse healthcare environments. These gaps highlight the pressing need for a next-generation DSS for sepsis prediction that is more generalizable, more transparent in its reasoning, and better integrated into the dynamic realities of hospital workflows. Our proposed DSS aims to address these limitations by offering a flexible, interpretable, and adaptive approach to sepsis risk prediction.

Policy & Management Consideration

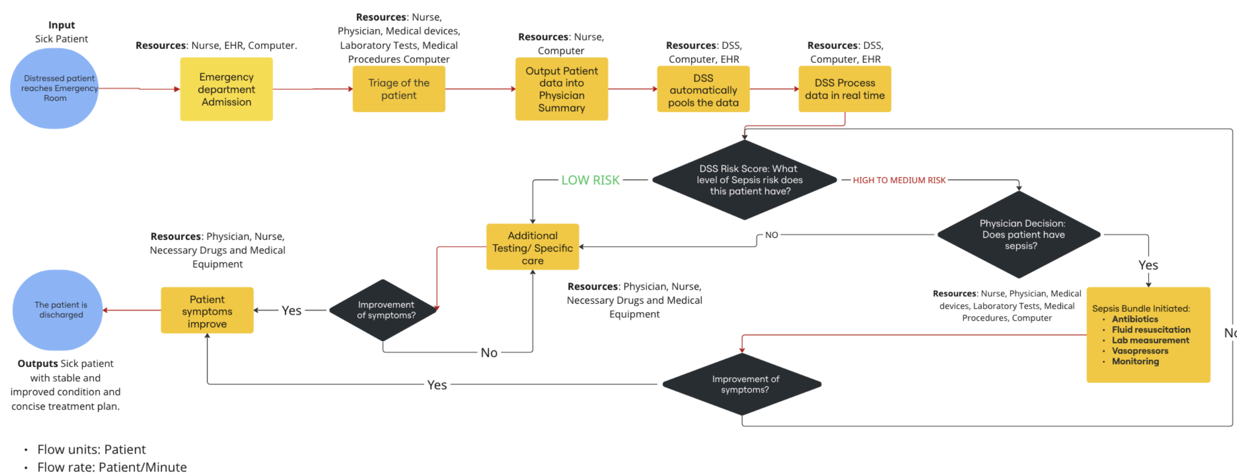
The adoption of sepsis decision support systems introduces important regulatory, legal, and operational challenges. Under FDA guidance, predictive tools that impact clinical decision-making are classified as medical devices and must meet established safety and efficacy standards (*Clinical Decision Support Software - Guidance for Industry and Food and Drug Administration Staff | FDA*, 2022). Hospitals must also manage liability risks, as clinicians could face legal exposure if they disregard accurate DSS recommendations or act on faulty ones (Price et al., 2019). Building clinician trust is essential for successful deployment; studies show that DSS tools poorly integrated into workflows are often ignored due to alert fatigue(Sutton et al., 2020). Finally, DSS applications must comply with HIPAA regulations to ensure patient data privacy and secure handling of clinical information (Rights (OCR), 2025). Addressing these concerns proactively is crucial for ensuring both regulatory compliance and meaningful clinical adoption.

DSS Model: Sepsis Prevention System (SPS)

Introduction & High-Level Process Model

Our DSS project, Sepsis Prevention System, uses evidence and data to identify inpatients at elevated risk for sepsis rather than relying solely on clinicians' judgment. While the overall clinical workflow remains largely unchanged, this evidence-based step enhances decision making by providing quantitative risk assessment at the bedside. To this end, our system has been modeled on hourly-granular clinical records from publicly available sepsis datasets, encompassing hundreds of patient encounters with and without sepsis outcomes (detailed dataset description appears in later sections). Clinicians enter patient vitals and laboratory values via a web form to compute real-time risk scores and prioritize high-risk patients accordingly.

Process Flow Diagram



Conceptual IT Architecture

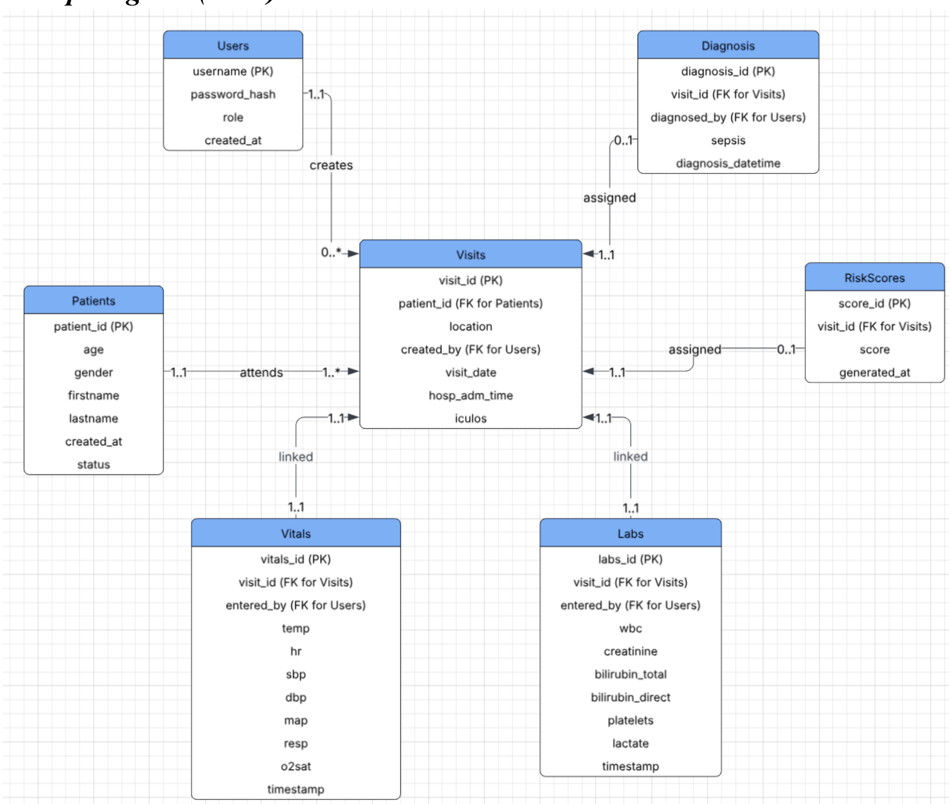
Relational Database Design

The Sepsis Prevention System (SPS) relies on a PostgreSQL relational database to serve as its clinical data repository. The schema is centered around the concept of a unique visit, with each visit linked to a specific patient and allowing for multiple visits per patient to support readmission scenarios. This design choice mirrors real-world clinical workflows, where patients may present to the hospital on different days with evolving clinical statuses. The database consists of seven primary tables:

- `Patients`: Stores demographic data such as name, age, gender, and identifiers.
- `Visits`: Captures each hospital encounter, including location and ICU length of stay.
- `Vitals`: Logs user-entered vital signs for each visit.
- `Labs`: Logs laboratory results for each visit.
- `RiskScores`: Stores computed sepsis risk scores per visit.
- `Diagnosis`: Records physician-confirmed sepsis diagnoses.
- `Users`: Tracks users entering or modifying patient data.

Key clinical attributes captured in the database correspond directly to features used by the machine learning model for risk prediction. The relational structure ensures referential integrity across patients, visits, and clinical measurements, providing a robust framework for real-time risk scoring and longitudinal tracking of patient outcomes.

Entity-relationship diagram (ERD).



HIPAA & Standards Impact

Our Sepsis Prevention Decision Support System (DSS) leverages established healthcare data standards to ensure interoperability, accuracy, and scalability:

- **FHIR (Fast Healthcare Interoperability Resources):** Enables real-time data exchange between systems, allowing dynamic updates of critical information like heart rate, blood pressure, and lab results.
- **HL7 Messaging Standards:** Provide the structure for securely and consistently transmitting clinical data across different healthcare systems.
- **LOINC (Logical Observation Identifiers Names and Codes):** Standardizes the reading and interpretation of lab and physiological data, such as white blood cell counts, ensuring model reliability across hospital systems.
- **SNOMED CT:** Offers precise and consistent labeling of clinical events, diagnoses, and symptoms, helping align DSS outputs with real-world medical documentation.
- **ICD-10 Codes:** Classify comorbid conditions and infection types that are essential inputs for model training and patient outcome analysis.
- **USCDI (United States Core Data for Interoperability):** Defines the essential patient data elements for exchange, ensuring compatibility with federal interoperability initiatives.

By implementing these standards, our DSS ensures accurate, scalable, and interoperable sepsis risk prediction to support timely clinical decision-making.

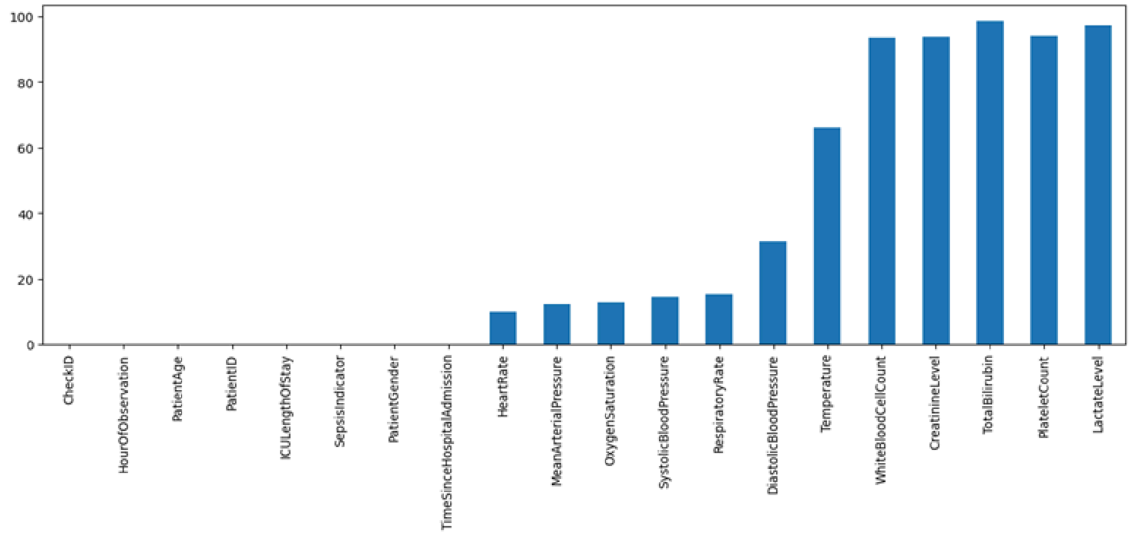
Data and Modeling

Data Sources and Feature Selection

We utilized the publicly available Kaggle Sepsis Prediction Dataset, containing de-identified time-series clinical data for thousands of patients. The dataset initially consisted of over 1.5 million observations across 37 clinical variables, each representing a patient's condition at a given hour. The original target variable, SepsisIndicator, was a binary classification of sepsis presence. For model development, we refined the target to predict a continuous Sepsis Risk Score between zero and one.

Feature selection was guided by a literature review and clinical relevance. Variables with excessive missingness or limited predictive value were excluded. For example, Direct Bilirubin was removed due to collinearity with Total Bilirubin and low availability. Features prioritized for inclusion included heart rate, respiratory rate, blood pressure, oxygen saturation, lactate level, creatinine level, platelet count, and total bilirubin, reflecting established sepsis physiology.

Data completeness per column.



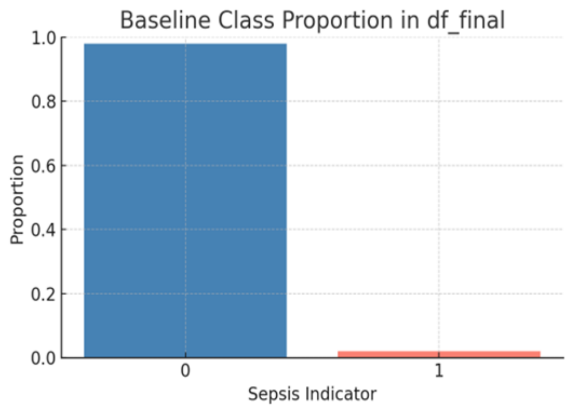
Exploratory Data Analysis

Each row in the dataset corresponds to one hour of patient observation. Column names were standardized to align with clinical terminology. Summary statistics and exploratory visualizations revealed substantial physiological variability, expected in clinical populations. Outliers were observed in key features such as ICULengthOfStay, HeartRate, LactateLevel, and WhiteBloodCellCount, consistent with sepsis pathology.

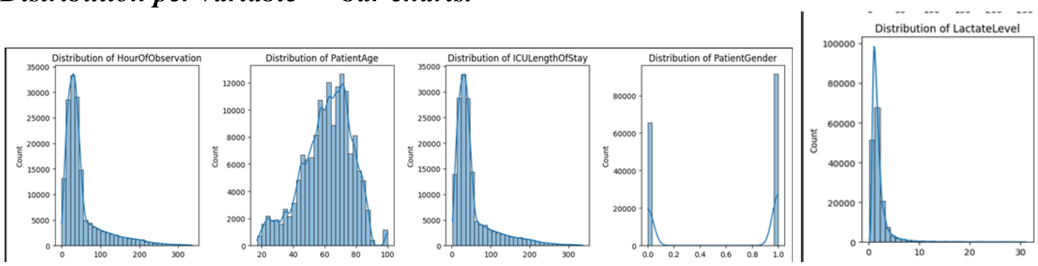
Missingness was common due to irregular testing intervals. Features like LactateLevel and PlateletCount had relatively high completeness, while time-based identifiers and demographic fields exhibited more missingness. To address this, we applied a forward fill method within each patient ID timeline to preserve real-time continuity, mimicking clinical practice. Remaining incomplete rows were dropped, yielding a cleaned dataset of 157,030 records.

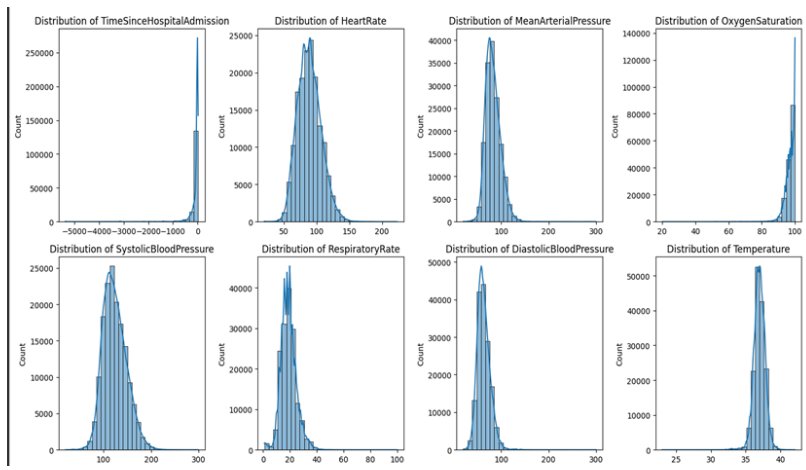
Class imbalance posed a major challenge, with only 5% of observations labeled sepsis-positive. We balanced the dataset by oversampling the minority class, creating a 50/50 split for model training with 15,984 total rows.

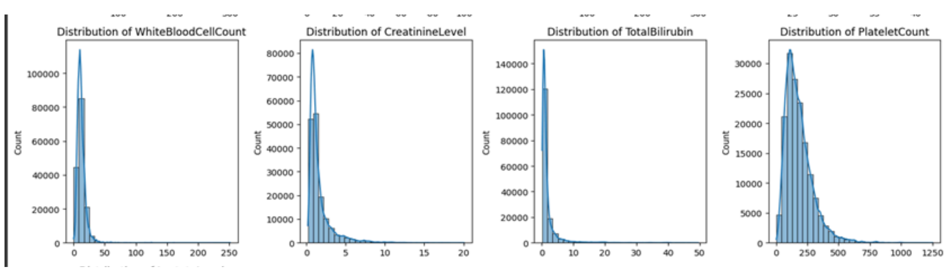
Baseline class proportion before oversampling.



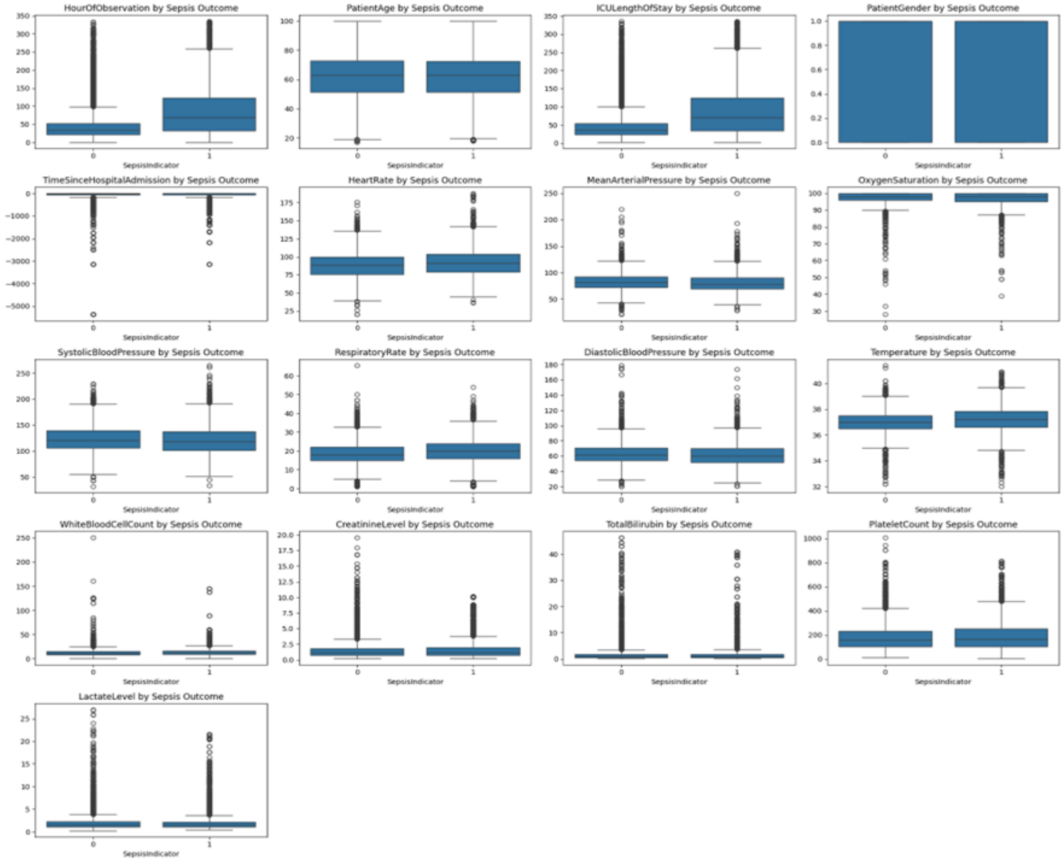
Distribution per variable — bar charts.



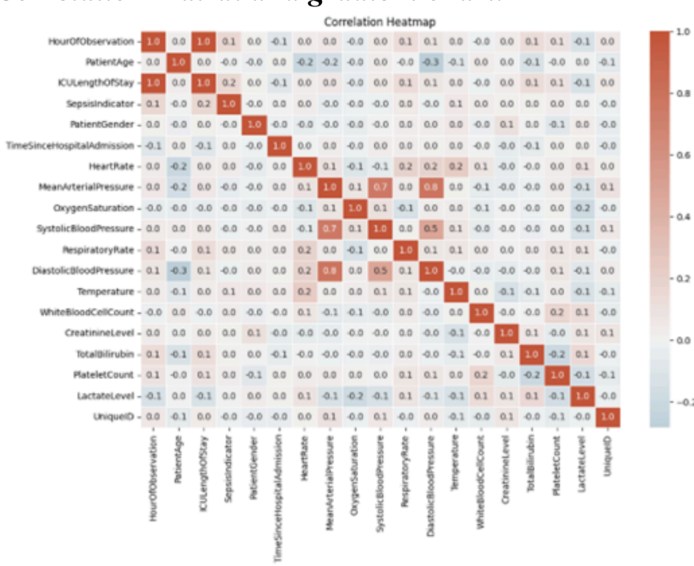


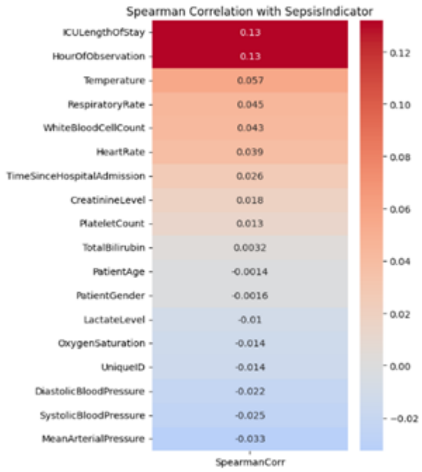


Distribution per variable — box plots.



Correlation matrix and gradient chart.



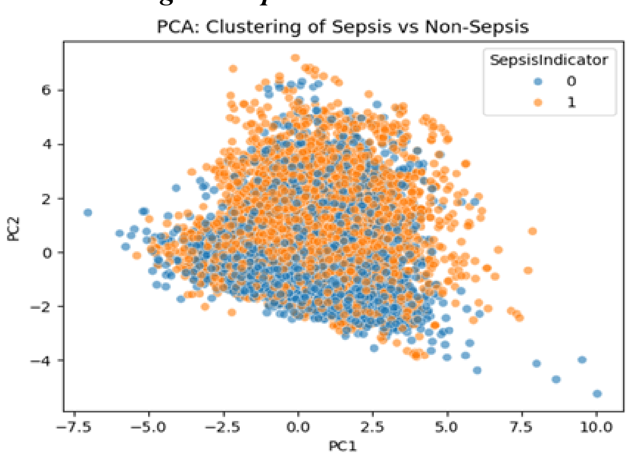


Model Choice and Evaluation

Distribution analysis through skewness plots and boxplots confirmed that most clinical features were skewed and contained outliers. *Z*-score standardization was applied to normalize feature ranges.

Correlation analysis showed that no single variable had a strong monotonic relationship with sepsis (Spearman's correlation ~0.13 for top predictors), indicating the need for a model capable of capturing complex, non-linear interactions. Principal Component Analysis (PCA) revealed substantial overlap between sepsis and non-sepsis cases, supporting the choice of non-linear modeling.

PCA clustering scatterplot.



We selected XGBoost for its robustness to outliers, ability to model non-linear relationships, and native handling of imbalanced datasets. The model achieved:

- Accuracy: 89%
- Recall (Sepsis class): 92%
- Precision (Sepsis class): 86%
- F1 Macro Score: 0.89 (via 5-fold Stratified Cross-Validation)

Risk scores ranged from probabilities like 72.08% to 98.55%, demonstrating the model's ability to prioritize high-risk patients.

```
              precision    recall  f1-score   support

           0       0.92      0.85      0.88      1998
           1       0.86      0.92      0.89      1998

    accuracy                           0.89      3996
   macro avg       0.89      0.89      0.88      3996
weighted avg       0.89      0.89      0.88      3996

Example risk scores: [98.55 33.83 60.49  7.75 92.62  9.66 41.8  60.13 67.08 72.08]
```

Logistic Regression and Support Vector Machine (SVM) models were also tested due to their interpretability but achieved lower performance (66.9% and 72.1% accuracy, respectively).

```
Logistic Regression Results:
              precision    recall  f1-score   support

           0       0.64      0.76      0.70      1998
           1       0.71      0.58      0.64      1998

    accuracy                           0.67      3996
   macro avg       0.68      0.67      0.67      3996
weighted avg       0.68      0.67      0.67      3996

Accuracy: 0.669
Precision: 0.707
Recall: 0.579
F1 Score: 0.636
# (Optional) also view a few sample risk scores
Example risk scores: [0.724  0.4278 0.5178 0.3625 0.6991 0.359  0.5231 0.4474 0.7062 0.3351]
```

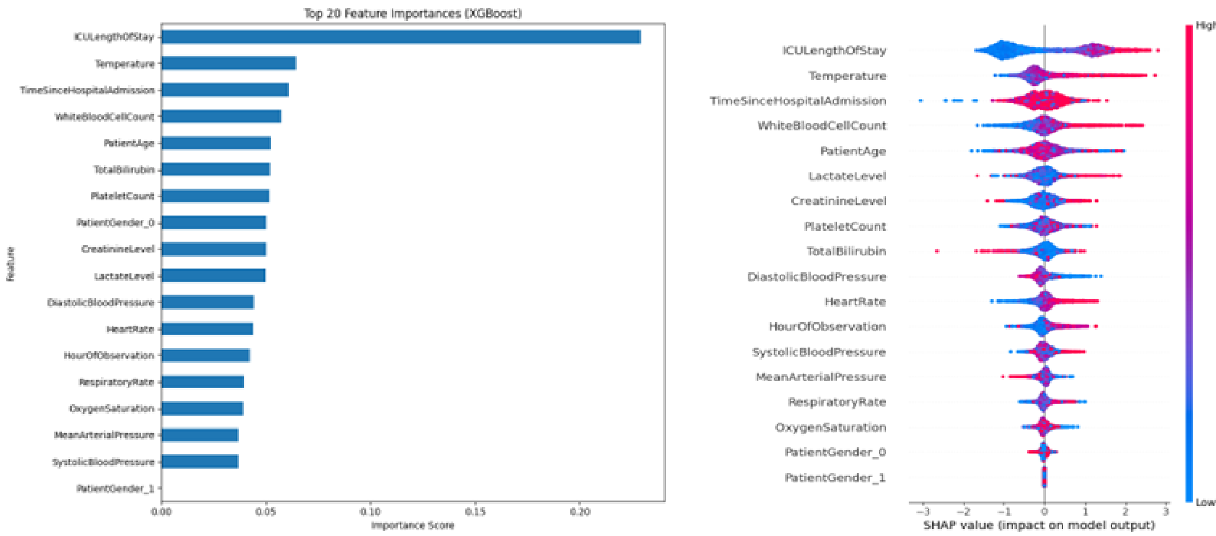
```
SVM Results:
              precision    recall  f1-score   support

           0       0.71      0.76      0.73      1998
           1       0.74      0.68      0.71      1998

    accuracy                           0.72      3996
   macro avg       0.72      0.72      0.72      3996
weighted avg       0.72      0.72      0.72      3996

Accuracy: 0.721
Precision: 0.74
Recall: 0.684
F1 Score: 0.711
# (Optional) also view a few sample risk scores
Example risk scores: [0.8473 0.6381 0.5137 0.2483 0.8677 0.1553 0.554  0.3356 0.8531 0.6269]
```

XGBoost outperformed both, offering better recall and F1 scores critical for high-risk sepsis detection. Beyond predictive superiority, XGBoost demonstrated operational advantages, including resilience to missing data, scalability, and generation of meaningful feature importance outputs.

Feature Importance and Interpretability (SHAP)

Feature importance plots identified ICULengthOfStay, Temperature, and WhiteBloodCellCount as top contributors to sepsis risk. SHAP (SHapley Additive exPlanations) analysis was conducted to quantify feature impacts at both global and individual levels.

Feature importance and SHAP value charts.



SHAP summary plots revealed that higher ICU stays, elevated temperatures, and increased WBC counts positively shifted predictions toward sepsis. Other features showed smaller or context-dependent impacts, while variables like gender and blood pressure readings contributed minimally. This multi-level interpretability ensures transparency, a critical requirement for integration into clinical workflows.

Sensitivity and Robustness Analysis

We tested model robustness by simulating the absence of top predictive features. Dropping ICULengthOfStay, Temperature, or WhiteBloodCellCount one at a time resulted in minimal drops in performance:

- Dropping ICULengthOfStay: Recall = 94%, F1 = 91%
- Dropping Temperature: Recall = 93%, F1 = 91%
- Dropping WhiteBloodCellCount: Recall = 93%, F1 = 91%

The model's ability to maintain performance despite missing high-impact features suggests strong real-time resilience, crucial in clinical environments where data availability may fluctuate.

Limitations

Despite strong performance, limitations exist. Forward filling mitigated but did not eliminate missing data artifacts. The robustness analysis only addressed single-feature loss; real-world data shifts could introduce broader challenges. External validation on independent datasets is necessary to assess generalizability. Future work will involve testing model performance across different patient populations and EHR systems.

Implementation and Deployment (PEIT Framework)

Process Factors

- Volume: The DSS will initially impact a moderate number of ICU processes, primarily targeting sepsis risk assessment during patient monitoring. Since the prediction uses existing EHR data already collected for clinical purposes, additional workload volume is minimal.
- Novelty: The system introduces moderate novelty. Clinicians are accustomed to assessing sepsis risk manually based on clinical judgment. The DSS introduces a structured, real-time risk score derived from machine learning models, complementing traditional assessment with data-driven insights. Moreover, real-time alerts are displayed to facilitate interpretability and ensure that physicians and nurses use the tool.
- Complexity: The complexity of adoption is low. The DSS integrates into existing EHR workflows with minimal additional data entry required. Sepsis risk predictions are automatically updated and displayed, minimizing disruption to clinician routines.

Technology Factors

- Tool: The DSS consists of an XGBoost-based predictive model integrated with a user-friendly dashboard that displays real-time sepsis risk scores and contributing factors to support decision-making.
- Template: The tool builds on standard clinical data collected in EHRs (e.g., vital signs, lab results) and presented in a color-coded, intuitive format. Data inputs conform to FHIR, LOINC, and SNOMED CT standards for compatibility and interoperability.

People

- Discretion: Clinicians maintain full discretion over how they use the DSS outputs. The system provides risk scores and explanations but leaves final clinical decisions to human providers, maintaining professional autonomy.

- Sophistication: No advanced technical skills are required to use the DSS. Clinicians simply review risk scores and key contributing factors presented visually. Basic orientation sessions may be provided to familiarize users with the dashboard layout and interpretation.

Strategies

- Big-bang Approach: Given the low complexity and moderate novelty, a big-bang rollout is feasible. Immediate full deployment across ICU units is recommended to prevent delays and streamline adoption, rather than a slow phased rollout.
- Rational-Empirical Approach: Initial training and communication efforts will focus on demonstrating the accuracy, reliability, and workflow benefits of the DSS. Sharing real-world case examples where early risk identification prevented adverse outcomes will help foster clinician trust and adoption.

Pitfalls

- Inertia: Resistance to change is expected to be low to moderate. Since the system complements rather than replaces clinical judgment, and since it minimizes workflow disruption, inertia is unlikely to significantly impede adoption.
- Nonuse: There is a risk that clinicians might initially ignore DSS outputs, particularly during periods of high workload or when the risk prediction conflicts with clinical intuition. Clear training on the system's strengths and continuous feedback collection will be necessary to encourage consistent use.

Conclusion, Policy Implications, & Recommendations

Conclusion

The early identification and management of sepsis remain among the most critical challenges in hospital care. Our Sepsis Prevention Decision Support System (DSS) leverages machine learning models integrated with existing electronic health record (EHR) data to predict sepsis risk in real-time. By offering clinicians a tool that highlights high-risk patients dynamically, the DSS enhances clinical decisions regarding septic patients without disrupting workflow. Machine learning algorithms can achieve high accuracy, precision, and recall, making them well-suited for sensitive, high-stakes applications like sepsis prevention. With rigorous adherence to interoperability standards and HIPAA compliance, the system is a first step for getting a DSS ready for real-world deployment and integration into ICU environments.

Policy Implications

The deployment of machine learning-based DSS tools like ours has important implications for healthcare policy. Policymakers should support initiatives that facilitate seamless, real-time data sharing and promote the development of interoperable predictive tools across hospital systems to improve health care outcomes, not just in early sepsis detection, but in different medical conditions. Additionally, as the use of AI in healthcare grows, regulatory bodies must develop clear guidelines for the ethical deployment of predictive analytics, focusing on patient safety, model transparency, clinician oversight, and accountability. Institutions must ensure that DSS outputs are interpretable and that clinicians remain the ultimate decision-makers in patient care, providing the patient with clear information about the reasons of the clinical decisions.

Furthermore, funding incentives or policy frameworks that encourage hospitals to adopt predictive DSS solutions for early sepsis detection could reduce preventable morbidity and mortality rates nationally, even mitigating the increase in health care costs for delaying the antibiotic treatment initiation. Finally, healthcare accreditation bodies could consider including clinical DSS integration and usage metrics as part of hospital quality assessments, so the hospitals and clinics are encouraged to invest in technologies for improving healthcare attention for patients.

Recommendations

To maximize the clinical impact of our Sepsis Prevention DSS, several actions are recommended:

- Conduct pilot deployments in ICU settings to gather user feedback, refine the dashboard interface, and ensure seamless workflow integration.
- Develop comprehensive training modules to improve clinician familiarity with interpreting machine learning-based risk scores and establish monitoring systems to continuously assess the model's performance and recalibrate thresholds as clinical populations or treatment patterns evolve.
- Collaborate with hospital IT departments to ensure that DSS alerts are properly delivered across the necessary devices in the hospital to increase accessibility of all physicians and nurses, especially during night shifts when sepsis detection delays are more common.
- Advocate for broader adoption of EHR interoperability standards at the institutional level to enable smoother scaling and cross-institutional learning.
- Comprehensive training ensures that healthcare providers understand how the DSS generates sepsis risk scores, interpret the outputs correctly, and confidently integrate the system's recommendations into clinical decision-making, thereby maximizing the tool's effectiveness and trust in real-world settings.

By following these recommendations, the Sepsis Prevention DSS can serve as a reliable, scalable, and transformative tool for improving sepsis care outcomes.

References

Clinical Decision Support Software—Guidance for Industry and Food and Drug Administration Staff | FDA. (2022, September 28). https://www.fda.gov/media/109618/download

Henry, K., Adams, R., Parent, C., Soleimani, H., Sridharan, A., Johnson, L., Hager, D., Cosgrove, S., Markowski, A., Klein, E., Chen, E., & Saheed, M. (2022). *Factors driving provider adoption of the TREWS machine learning-based early warning system and its effects on sepsis treatment timing*. *28*, 1447–1454. https://doi.org/10.1038/s41591-022-01895-z

Liu, V., Escobar, G., & Greene, J. (2014). *Hospital deaths in patients with sepsis from 2 independent cohorts*. *312*(1), 90–92. https://doi.org/10.1001/jama.2014.5804

Price, W., Gerke, S., & Cohen, I. (2019). *Potential Liability for Physicians Using Artificial Intelligence*. *322*(18), 1765–1766. https://doi.org/10.1001/jama.2019.15064

Rhee, C., Dantes, R., & Epstein, L. (2017). *CDC Prevention Epicenter Program. Incidence and trends of sepsis in US hospitals using clinical vs claims data*. *318*(13), 1241–1249. https://doi.org/10.1001/jama.2017.13836.

Rights (OCR), O. for C. (2025, March 14). *Summary of the HIPAA Privacy Rule* [Page]. https://www.hhs.gov/hipaa/for-professionals/privacy/laws-regulations/index.html

Sutton, R., Pincock, D., Baumgart, D., Sadowski, D., Fedorak, R., & Kroeker, K. (2020). *An overview of clinical decision support systems: Benefits, risks, and strategies for success*. *3*. https://doi.org/10.1038/s41746-020-0221-y

Uffen, J., Oosterheert, J. J., Schweitzer, V. A., Thursky, K., Kaasjager, H. A. H., & Ekkelenkamp, M. B. (2021). *Interventions for rapid recognition and treatment of sepsis in the emergency department: A narrative review*. *27*, 192–203.