

Predictive Sepsis Detection and Post-Op Health Report Automation

PBL Review Report – I

Submitted To:

Dr. Pradnya Sulas Borkar

Mentor

Submitted By:  
Tanmay Kalinkar, 23070521189

Vth, C

Siddhesh Mawale, 23070521229

Vth, C

Eavam Margamwar, 23070521209

Vth, C

August 2025

**Table of Contents**

|  |  |  |
| --- | --- | --- |
| Sr. No. | Title | Page No. |
| 1 | **Introduction** | **1** |
| 2 | **Problem Statement** | **3** |
| *2.1* | *Primary Problems* |  |
| *2.2* | *Clinical Impact* |  |
| 3 | **Objectives** | **3** |
| *3.1* | *Primary Objective* |  |
| *3.2* | *Specific Objectives* |  |
| 4 | **Engineering Systems, Tools, and Key Parameters** | **4** |
| *4.1* | *Technology Stack `(Revised)* |  |
| *4.2* | *Key Clinical Parameters* |  |
| 5 | **Literature Review and Existing Methods** | **5** |
| *5.1* | *Performance Comparison from Literature* |  |
| *5.2* | *Proposed Methodology* |  |
| 6 | **Future Trends** | **7** |
| *6.1* | *Emerging Technologies* |  |
| *6.2* | *Clinical Integration* |  |
| 7 | **References** | **7** |

**Introduction**

Sepsis continues to represent one of the most challenging and deadly conditions faced by modern healthcare systems worldwide. Each year, millions of patients experience sepsis, which can rapidly lead to organ failure, shock, and increased risk of mortality. The seriousness of this condition demands early diagnosis and timely clinical intervention, yet traditional hospital scoring systems and manual monitoring routines often fall short. The complexity and variability of sepsis presentation highlight the pressing need for innovation, particularly in environments such as intensive care units (ICUs), emergency departments, and postoperative wards.

Detecting sepsis early is especially problematic outside the ICU, where patients may not benefit from continuous monitoring. In many hospitals, vital signs are only checked every few hours, and patient data can be fragmented across multiple records or visits. For surgical patients, complications such as sepsis often develop postoperatively when they are most vulnerable and before obvious symptoms appear. The limits of current approaches create a dangerous delay in recognizing this rapidly progressing disease.

Beyond the immediate concerns in the ICU, sepsis is a major risk factor in the postoperative care process. Surgery can raise the likelihood of infections, and the onset of sepsis in the days after discharge is a key driver of hospital readmissions and adverse outcomes. Manual observation and reliance on standardized scoring—such as the qSOFA and SOFA indices—are widely used but generally lack the predictive power and speed necessary to flag deteriorating patients before they reach critical status [4]. These scores depend on a handful of clinical features and leave much of the patient journey unmonitored, especially once the patient leaves continuous care environments.

Recent advances in machine learning (ML) and data-driven healthcare signal a new era in predictive analytics and patient monitoring. Instead of waiting for obvious symptoms or depending solely on periodic checks, ML models can process large amounts of patient data—such as vital signs, lab results, and medical history—in real time. By identifying subtle changes and complex patterns, these models can predict the onset of sepsis hours or even days in advance. This capability not only increases survival rates but can also reduce the burden of unnecessary ICU admissions, improve triage routines, and help clinicians prioritize high-risk patients.

A key innovation has been the development of specialized ML algorithms, such as XGBoost, Light Gradient Boosting Machine (LGBM), and Random Forest, which have already demonstrated high accuracy in predicting sepsis with minimal data [1,2,3]. Instead of requiring dozens of complicated test results, modern models can deliver reliable predictions using just a few routine blood tests or baseline vital signs [1]. This makes the solution scalable and accessible, whether deployed in large tertiary hospitals or smaller community clinics.

The landscape of postoperative health monitoring has also begun to shift. With advances in wearable technology, remote patient monitoring, and automated reporting systems, clinicians can eventually keep track of patients’ vital signs, wound healing, and clinical status even after discharge. Automated health report systems display a promising way to catch early signs of sepsis and other complications, minimizing the risk of missed deteriorations.

The goal of this project is to bridge the gap between traditional manual monitoring and the promise offered by intelligent, automated prediction tools. By leveraging a scalable machine learning-based system, the project aims to provide not just early sepsis detection, but an integrated solution for postoperative health report automation. This will empower clinicians with real-time alerts, interpretable predictions, and actionable dashboards, enhancing both patient safety and workflow efficiency.

Overall, the convergence of accurate ML models and smart monitoring systems represents a turning point in managing acute health threats such as sepsis. It enables healthcare providers to move from reactive care to proactive, data-guided intervention—ultimately saving lives and optimizing health system resources. This report will explore the current state of sepsis prediction, the core technologies required, and the best approaches to building a scalable, user-friendly solution for today’s hospitals and clinics.

**Problem Statement**

**Primary Problems:**

• **Late sepsis detection** - Traditional scoring systems achieve AUCs of only 0.17-0.73 compared to ML models achieving 0.87-0.99 [3, 4].  
• **Manual monitoring limitations** - Vital signs monitored only every 4-8 hours postoperatively with no continuous surveillance after discharge  
• **High false alarm rates** - Existing sepsis alert systems suffer from alarm fatigue among clinicians  
• **Limited interpretability** - Current ML models lack clinical explainability for practical adoption  
• **Fragmented care** - No integrated system connecting sepsis prediction with postoperative health monitoring

**Clinical Impact:**

• Every hour of delayed sepsis treatment increases mortality risk by 4-9%  
• Approximately 350,000 sepsis-related deaths annually in the US alone  
• Substantial proportion of postoperative deaths occur after discharge

**Objectives**

**Primary Objective:**

To Develop a scalable machine learning system for early sepsis prediction integrated with automated postoperative health monitoring

**Specific Objectives:**

• **Early Detection**: To Predict sepsis onset 6-48 hours before clinical manifestation with >85% sensitivity  
• **Real-time Monitoring**: To Implement continuous postoperative vital sign tracking and automated reporting  
• **Clinical Integration**: To Create interpretable models using SHAP/LIME for clinical decision support  
• **Scalability**: To Design web-based platform capable of handling multiple healthcare facilities  
• **Risk Stratification**: To Identify high-risk surgical patients for targeted monitoring

**Engineering Systems, Tools and Key Parameters**

**Technology Stack (Revised):**

**Frontend**

* **Framework:** React + Tailwind CSS

Stable, widely used, excellent for scalable web apps

* **Data Visualization:** Recharts, Axios, Lucide

Powerful and flexible for clinical charts and dashboards

**Backend**

* **Framework:** FastAPI + Python 3.11  
  Lightweight, fast, async-ready, ideal for ML model serving

**Model Training & Serving:**

* **XGBoost + Scikit-learn** for model building
* **SHAP** Model interpretability and explanations

**Development Tools**

* **Version Control:** Git and GitHub

**Key Clinical Parameters:**

**Sepsis Prediction Features:**  
• Vital signs: Heart rate, blood pressure, temperature, respiratory rate, oxygen saturation  
• Laboratory values: Lactate, white blood cell count, prothrombin time, creatinine  
• Clinical scores: SOFA, qSOFA components  
• Demographics: Age, gender, comorbidities

**Postoperative Monitoring:**  
• Wound healing indicators  
• Pain assessment scores  
• Functional status measures  
• Medication adherence tracking  
• Early warning signs of complications

**Literature Review and Existing Methods**

**Traditional Sepsis Detection:**

• **qSOFA Score**: AUC 0.76, limited sensitivity for early detection [4].  
• **SOFA Score**: AUC 0.17-0.73, requires ICU-level monitoring [4].  
• **SIRS Criteria**: High false positive rates, poor specificity [4].

**Current ML Approaches:**

• **XGBoost Models**: Achieving AUCs 0.82-0.91 across multiple studies [3].  
• **Deep Learning**: LSTM networks with AUCs 0.94-0.99 but limited interpretability [3].  
• **Random Forest**: Consistent performance with AUC 0.87, good feature importance [2].  
• **Gradient Boosting**: Strong performance in MIMIC-IV studies with AUCs 0.83-0.91 [3,5].

**Postoperative Monitoring Systems:**

• **Manual Vital Signs**: Every 4-8 hours, prone to human error  
• **Remote Patient Monitoring**: Limited adoption, mixed results in reducing readmissions  
• **Wearable Devices**: Emerging technology with promising results for continuous monitoring

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Study | Year | Data Source | Model/Algorithm | Features | Sample Size | AUROC | Citation |
| Lin et al. | 2025 | ICU, Taiwan | LGBM, RFC, XGBoost | CBC+DIFF (Blood Counts) | 1,492 | 0.9 | [1] |
| Zhou et al. | 2024 | China + MIMIC-III | Random Forest, XGBoost | 18 Labs/Vitals/Scores | 1,968+ | 0.87 | [2] |
| Yadgarov et al. | 2024 | Meta-Analysis | NN, Decision Tree, Others | Various (Vitals, Labs) | 457,932 | 0.83\* | [5] |
| Shanmugam et al. | 2025 | Review (13 studies) | XGBoost, GBDT, DL | Labs, Vitals, EHR data | Varies | 0.83–0.99 | [3] |
| Liu et al. | 2025 | MIMIC-III/IV, AP | GBDT, SVM, KNN | 13 (Vitals, Labs, GCS, Age) | 1,672 | 0.985 | [6] |
| Liu et al. | 2025 | MIMIC-IV, ED | Gradient Boosting | Triage Vitals + Demographics + Hx | 189,617 | 0.83 | [6] |

|  |  |  |  |
| --- | --- | --- | --- |
| **Approach** | **Advantages** | **Disadvantages** | **AUC Performance** |
| **Traditional Scoring** | Established clinical use, Simple calculation | Poor sensitivity, Late detection | 0.17-0.76 |
| **Machine Learning** | High accuracy, Early prediction | Can be 'black box' models, requiring interpretability tools (e.g., SHAP/LIME) | 0.87-0.99 |
| **Deep Learning** | Excellent performance, Pattern recognition | Limited interpretability, High complexity | 0.94-0.99 |
| **Hybrid Approach** | Combines accuracy with interpretability | Implementation complexity | 0.85-0.95 |

**Performance Comparison from Literature:**

• **LGBM (Lin et al.)**: AUC 0.90 using CBC data only [1].  
• **Random Forest (Zhou et al.)**: AUC 0.87 with balanced datasets [2].  
• **XGBoost (Multiple studies)**: Consistent AUCs 0.82-0.91 [3].  
• **Traditional Scores**: qSOFA (0.76), SOFA (0.73) significantly lower [4].

**Proposed methodology**: Hybrid ML approach using XGBoost with SHAP interpretability, achieving optimal balance between performance and clinical usability.

**Future Trends**

**Emerging Technologies:**

• **Federated Learning**: Multi-hospital model training while preserving privacy  
• **Edge Computing**: Real-time inference on wearable devices for immediate alerts  
• **Multimodal Integration**: Combining clinical notes, imaging, and sensor data  
• **Conformal Prediction**: Uncertainty quantification for safer clinical deployment

**Clinical Integration:**

• **EHR Integration**: Seamless integration with existing hospital information systems  
• **Mobile Applications**: Point-of-care decision support on smartphones and tablets  
• **Regulatory Compliance**: FDA approval pathways for AI-based medical devices  
• **Continuous Learning**: Models that adapt and improve with new clinical data

**References**

1. Lin, et al. (2025*). Light Gradient Boosting Machine (LGBM) for Sepsis Prediction Using CBC Data.* Journal of Medical Internet Research. [Link to paper](https://www.jmir.org/2025/1/e74940)
2. Zhou, et al. (2024). *A Random Forest Model for Sepsis Prediction on Balanced Datasets.* Preventive Medicine Reports. [Link to paper](https://www.mdpi.com/2075-4418/15/3/302)
3. Shanmugam, et al. (2025*). A Review of XGBoost Models for Sepsis Prediction. Indian Journal of Critical Care Medicine.* Journal of SepsisAI. [Link to paper](https://www.ijccm.org/doi/10.5005/jp-journals-10071-24986)
4. Gupta, A., Chauhan, R., Saravanan, G., & Shreekumar, A. (2024*). Improving sepsis prediction in intensive care with SepsisAI.* Journal of SepsisAI. [Link to paper](https://journals.plos.org/digitalhealth/article?id=10.1371/journal.pdig.0000569)
5. Yadgarov, et al. (2024*). A Comprehensive Network Meta-Analysis of Machine Learning Algorithms for Sepsis Prediction.* Frontiers in Medicine. [Link to paper](https://www.frontiersin.org/articles/10.3389/fmed.2024.1500326/full)
6. Liu, et al. (2025). *Gradient Boosting on the Large-Scale MIMIC-IV Dataset for Sepsis Prediction.* Scientific Reports. [Link to paper](https://www.nature.com/articles/s41598-025-99879-9)