

Sepsis Prevalence and Mortality Risk in ICU Patients

A MIMIC-III Database Analysis

Ernest Ceballos, Júlia Galimany, Oriol Galimany

Universitat Rovira i Virgili

January 2026

1 Introduction

2 Methods

3 Results

4 Discussion

5 Conclusion

6 Appendix

Section 1

Introduction

What is Sepsis?

- **Definition:** Life-threatening organ dysfunction caused by dysregulated host response to infection
- **Clinical Impact:**
 - Leading cause of mortality in ICUs worldwide
 - Rapid progression requires early identification
 - Affects millions globally each year
- **Challenge:** Can we predict who is at highest risk?

Research Objectives

- ➊ **Estimate sepsis prevalence** among ICU patients
- ➋ **Compare mortality rates** between sepsis and non-sepsis patients
- ➌ **Develop a predictive model** for mortality risk in sepsis patients
- ➍ **Identify key risk factors** associated with mortality

Section 2

Methods

Data Source: MIMIC-III

- **Database:** Medical Information Mart for Intensive Care III (v1.4)
- **Institution:** Beth Israel Deaconess Medical Center
- **Period:** 2001-2012
- **Content:** De-identified health data from >40,000 ICU admissions

Cohort Selection

Inclusion Criteria:

- Adult patients (age ≥ 16 years)
- At least one ICU admission recorded

Final Cohort: 61,524 unique ICU admissions

Sepsis Identification (ICD-9 codes):

Code	Description
038.x	Septicemia
99591	Sepsis
99592	Severe sepsis
78552	Septic shock

Statistical Analysis

- **Prevalence:** Proportion of ICU stays with sepsis diagnosis
- **Mortality comparison:** Chi-square test
- **Predictive modeling:** Logistic regression

Model predictors:

- Age, Gender, Ethnicity
- Insurance type
- ICD-9 diagnosis code

Section 3

Results

Sepsis Prevalence

8.24% of ICU admissions had sepsis

- **Total ICU admissions: 61,524**
- **Sepsis cases: 5,071**

Approximately 1 in 12 ICU patients has sepsis

Prevalence by Demographics

By Gender:

- **Males:** 8.81% (higher)
- **Females:** 7.51%

By Age Group:

Age Group	Prevalence
16-40	6.4%
40-60	8.2%
60-80	8.7%
80+	8.1%

Prevalence increases with age until 60-80 years

Mortality: Sepsis vs Non-Sepsis

Group	Mortality Rate
With Sepsis	24.39%
Without Sepsis	9.93%

Sepsis patients have 2.5× higher mortality risk

$p < 0.001$ (*Chi-square test*)

Predictive Model: ROC Curve

Logistic Regression Performance:

Metric	Value
AUC	0.622
Optimal Threshold	0.334
Accuracy	58.3%
Sensitivity	60.4%
Specificity	57.3%
PPV	40.6%
NPV	74.9%

Model Interpretation

Key Findings:

- AUC of 0.622 indicates **modest discriminative ability**
- Model performs better than random ($AUC > 0.5$)
- **Demographic factors alone are insufficient** for accurate risk stratification

Significant predictors:

- Age (older \rightarrow higher risk)
- Insurance type (proxy for socioeconomic status)
- Specific ICD-9 codes

Section 4

Discussion

Key Insights

① High Disease Burden

- 1 in 12 ICU patients affected
- 2.5× mortality risk

② Demographic Patterns

- Higher prevalence in males
- Increasing prevalence with age

③ Model Limitations

- Demographic predictors insufficient ($AUC = 0.622$)
- Need for clinical variables (lactate, vitals, etc.)

Limitations

- **Retrospective design:** Cannot establish causality
- **Single center:** Results may not generalize
- **ICD-9 coding:** Possible under/over-reporting
- **Independence assumption:** Same patient may have multiple admissions
- **Missing clinical variables:** Vital signs, lab values excluded

Clinical Implications

- Sepsis screening should be **universal in ICU settings**
- **Demographic factors alone cannot guide** treatment decisions
- Future models should incorporate:
 - Vital signs (temperature, heart rate, BP)
 - Laboratory values (lactate, WBC, creatinine)
 - Organ dysfunction scores (SOFA, qSOFA)

Section 5

Conclusion

Summary

Main Findings

- ① Sepsis prevalence: **8.24%** of ICU admissions
- ② Mortality risk: **2.5× higher** with sepsis
- ③ Predictive model: **AUC = 0.622** (modest performance)

Take-home message:

Sepsis remains a critical clinical challenge. Effective risk stratification requires integration of clinical parameters beyond basic demographics.

Future Directions

① Incorporate clinical variables

- Vital signs, laboratory values
- Organ dysfunction scores

② Machine learning approaches

- Random forests, neural networks
- Real-time prediction models

③ External validation

- Multi-center studies
- Prospective evaluation

Thank You!

Questions?

Contact:

- Ernest Ceballos
- Júlia Galimany
- Oriol Galimany

Electronic Health Records - Activity A3

Universitat Rovira i Virgili, January 2026

Section 6

Appendix

Code Availability

- **Analysis Notebook:** `analysis/Activity-A3.ipynb`
- **Full Report:** `paper/main.pdf`
- **Repository:** GitHub - Blanqui04/EHR-A3-Group-B

References

- ① Singer M, et al. *JAMA*. 2016;315(8):801-810. (Sepsis-3 definitions)
- ② Johnson AEW, et al. *Scientific Data*. 2016;3:160035. (MIMIC-III)
- ③ Fleischmann C, et al. *Am J Respir Crit Care Med*. 2016;193(3):259-272.