Sepsis Prediction Algorithmic Impact Assessment

Artifact 3 | NorthStar Health Network

NorthStar Health Network is a fictional regional health system comprising 12 hospitals, 70 outpatient facilities, approximately 18,000 employees, and roughly 3 million annual patient encounters. All systems, metrics, incidents, vendors, and findings in this portfolio are simulated.

System Under Assessment

Field	Assessment
System	NS-AI-001 - Sepsis Prediction Model
Intended use	Identify adult inpatients at elevated risk of sepsis and prompt clinician assessment; not an autonomous diagnosis or treatment order.
Users	Nurses, hospitalists, rapid-response clinicians, clinical operations leaders
Affected population	Adult inpatients at participating NorthStar hospitals
Decision influence	High. Alerts can accelerate assessment, labs, escalation, and treatment decisions.
Risk tier	Critical
Approval state	Conditional approval pending closure of safety, subgroup-validation, and alert-workflow findings.

Clinical Workflow

EHR data ingestion → risk score → alert threshold → nurse / physician review → clinical assessment → independent clinical decision → documented action or override.

Impact Assessment

Domain	Key question	Finding	Required action
Patient safety	Could a false negative delay care?	Yes; missed risk is the most consequential failure mode.	Sensitivity floor and false-negative review by service line.
Alert burden	Could false positives create alert fatigue?	Yes; high alert volume may reduce response quality.	Site-level PPV, alert rate, and acknowledgment-time thresholds.
Subgroup performance	Does performance vary by age, sex, race/ethnicity, language, disability, or site?	Synthetic validation indicates wider uncertainty in smaller subgroups.	Minimum sample rules; stratified performance; escalation where confidence is inadequate.
Data quality	Can missing or delayed labs distort risk?	Yes.	Freshness checks, missingness flags, downtime procedure.
Human oversight	Can clinicians understand and challenge the alert?	Partially. Current explanation is insufficiently specific.	Surface key contributing factors and clear “not a diagnosis” language.
Transparency	Are users aware of intended use and limitations?	Needs improvement.	Clinical user guide, version notice, training, update communication.
Privacy	Does the model use PHI appropriately?	Permitted healthcare operations / treatment context assumed; minimum necessary and vendor handling still require review.	Document data flows, BAA status, access, retention, and secondary-use restrictions.

Synthetic Validation Snapshot

Metric	Pilot result	Governance threshold	Status
AUROC	0.82	≥ 0.80 with stable site variation	Meets
Sensitivity	0.86	≥ 0.85	Meets
Positive predictive value	0.31	≥ 0.28	Meets

Metric	Pilot result	Governance threshold	Status
Median alert lead time	4.1 hours	≥ 3 hours	Meets
Worst-site sensitivity	0.79	≥ 0.82	Gap
Alert acknowledgment within 30 min	78%	≥ 90%	Gap

Residual Risk & Decision

Residual risk: High. NorthStar should not expand system-wide until worst-site sensitivity, alert-response workflow, explanation quality, and subgroup evidence meet the defined conditions. The system may remain in a limited controlled pilot with enhanced monitoring and rapid pause authority.