**Sepsis Detection Criteria (Using Lab Data from labevents)**

Sepsis is typically defined based on **clinical criteria** like SOFA (Sequential Organ Failure Assessment) or SIRS (Systemic Inflammatory Response Syndrome). However, the most relevant **lab-based features** commonly associated with sepsis include:

**Key Indicators from Lab Data:**

1. **White Blood Cell Count (WBC)**:
   * High (>12,000 cells/mm³) or low (<4,000 cells/mm³) WBC levels can indicate infection or inflammation.
2. **Lactate**:
   * Elevated lactate (>2 mmol/L) is a hallmark of sepsis-related hypoperfusion.
3. **Platelets**:
   * Low platelet count (<100,000/µL) indicates coagulation abnormalities, often present in sepsis.
4. **Creatinine**:
   * Elevated creatinine (>1.2 mg/dL) indicates kidney dysfunction, a common sepsis complication.
5. **Procalcitonin (PCT)**:
   * High PCT levels (>2 ng/mL) suggest bacterial infections leading to sepsis.
6. **C-reactive Protein (CRP)**:
   * Elevated CRP (>10 mg/L) indicates systemic inflammation.
7. **Arterial Blood Gases** (if available):
   * Includes pH, pCO₂, and bicarbonate levels to assess metabolic acidosis (lactate-related).

**Features for Sepsis Prediction**

**Suggested Features:**

1. **Demographics**:
   * Age, sex (from patient or admissions table).
2. **Vital Signs**:
   * While not from labevents, vital signs (e.g., heart rate, blood pressure, temperature) can be included.
3. **Lab Features (from labevents)**:
   * WBC count
   * Lactate
   * Platelets
   * Creatinine
   * Procalcitonin (PCT)
   * CRP
4. **Derived Features**:
   * Sequential time trends in lab results (e.g., change in WBC over time).
   * Aggregated statistics (mean, min, max) for lab values during the hospital stay.

**Target Variable (Label):**

* Sepsis diagnosis could be based on clinical criteria or ICD codes (from diagnoses\_icd or admissions table) for sepsis or septic shock.

**Workflow to Extract Lab Features for Sepsis Prediction:**

1. **Filter Lab Test Data**: Use the d\_labitems table to identify itemid values for:
   * WBC count
   * Lactate
   * Platelets
   * Creatinine
   * Procalcitonin
   * CRP
2. **Aggregate Lab Results**: For each subject\_id and hadm\_id:
   * Calculate summary statistics (e.g., mean, max, min, trend over time).
   * Map valuenum to these lab test names (e.g., lactate, WBC).

These features are comprehensive for sepsis detection and prognosis:

* They cover **inflammation (CRP, WBC)**, **organ dysfunction (Creatinine)**, **coagulation issues (Platelet Count)**, and **metabolic abnormalities (Lactate, Bicarbonate)**.
* While **Procalcitonin (PCT)** is a specific marker for sepsis, its absence won't significantly limit your project as CRP and WBC provide similar inflammatory insights.