Mapping of Lung Cancer Trial Criteria to EMR Data (FHIR and Unstructured Sources)

---

1. Demographic Criteria

Rules:

- Patient must be ≥18 years old.
- Any gender is allowed; must be recorded.
- Female patients of childbearing potential must have a negative pregnancy test.

a. Structured Data (FHIR):

- **Age**: `Patient.birthDate`
- **Sex (Gender)**: `Patient.gender`
- **Pregnancy Status**: `Observation` with LOINC `2106-3`

b. Unstructured Data:

- Clinical notes and discharge summaries: Age and gender.
- Lab reports or staff notes: Pregnancy test status.

---

2. Clinical Criteria

Rules:

- ECOG Performance Status must be 0 or 1.
- No serious comorbidities (e.g., uncontrolled heart failure, recent stroke, severe COPD).
- No recent or concurrent invasive cancers (usually within 2–5 years).
- No active HIV, HBV, or HCV infection.
- No active autoimmune disease requiring systemic immunosuppression.
- No dependence on long-term supplemental oxygen.
- Life expectancy should be >3–6 months (physician judgment).

a. Structured Data (FHIR):

- **ECOG**: `Observation` with LOINC `89247-1`
- **Comorbidities**: `Condition` with SNOMED CT/ICD
- **Infections**: `Observation` with LOINC for HIV/HBV/HCV
- **Autoimmune Disorders**: `Condition` with SNOMED CT
- **Oxygen Use**: `DeviceUseStatement` or narrative Observation

b. Unstructured Data:

- Clinical and infectious disease notes
- Pathology history
- Text-based ECOG scoring

---

3. Diagnosis Criteria

Rules:

- Must have a confirmed diagnosis of lung cancer (e.g., NSCLC, SCLC).
- Disease must be at a specified stage (e.g., Stage IV, unresectable IIIB).
- No other active primary cancers.
- Cancer histology must match trial requirement (e.g., non-squamous NSCLC).

a. Structured Data (FHIR):

- **Diagnosis**: `Condition` with verification status "confirmed"
- **Stage**: `Observation` linked to `Condition`
- **Pathology Report**: `DiagnosticReport` with category `pathology`

b. Unstructured Data:

- Oncology notes, pathology reports, and tumor board summaries

---

4. Imaging Criteria

Rules:

- Brain MRI must show no active or untreated brain metastases unless stable.
- Must have measurable disease (e.g., ≥10 mm lesion on CT).
- No extensive disease preventing assessment (e.g., massive pleural effusion).
- No imaging signs of ILD (interstitial lung disease).
- No risky tumor locations (e.g., encasing blood vessels).

a. Structured Data (FHIR):

- **Imaging**: `ImagingStudy`, `DiagnosticReport`
- **Lesion Size**: `Observation`

b. Unstructured Data:

- Radiology reports and narrative scan reviews

---

5. Pathology/Biomarker Criteria

Rules:

- Must have specific mutation or expression if required by trial:
  - EGFR mutation for EGFR trials
  - ALK/ROS1/RET fusions for respective therapies
  - PD-L1 ≥1% or ≥50% for immunotherapy
  - TMB high for certain trials
- No squamous histology if contraindicated
- Pathology report must be available; tumor tissue sample must be accessible

a. Structured Data (FHIR):

- `Observation`, `DiagnosticReport` for EGFR, ALK, PD-L1, TMB, etc.

b. Unstructured Data:

- Pathology/molecular reports (scanned or text)
- Clinician notes summarizing results

---

6. Laboratory Criteria

Rules:

- **Hematologic (CBC):**
  - ANC ≥ 1.5 × 10⁹/L
  - Platelets ≥ 100 × 10⁹/L
  - Hemoglobin ≥ 9–10 g/dL

- **Liver Function:**
  - Bilirubin ≤ 1.5× ULN (or 3× if benign elevation)
  - AST/ALT ≤ 2.5× ULN (5× if liver mets)
- **Renal Function:**
  - Creatinine ≤ 1.5× ULN or CrCl ≥ 50–60 mL/min
- **Pregnancy:**
  - Negative HCG test before trial entry
- **Others:**
  - Electrolytes within normal range
  - INR/aPTT normal if not anticoagulated
  - TSH/T4 within normal range if immune-related therapy

a. Structured Data (FHIR):

- `Observation` codes:
  - ANC: LOINC `26499-4`
  - Platelets: LOINC `777-3`
  - Hemoglobin: LOINC `718-7`
  - Bilirubin: LOINC `1975-2`
  - ALT: LOINC `1742-6`, AST: LOINC `14629-0`
  - Creatinine: LOINC `2160-0`, eGFR: LOINC `2164-2`
  - Pregnancy: LOINC `2106-3`

b. Unstructured Data:

- Lab panels (PDF, scanned tables), clinician summary notes

A Term-to-Representation Mapping Table

Category	Trial Term	FHIR Resource	Code/System Used	Unstructured Source Examples
Demographics	Age	Patient	`birthDate`	"55-year-old male" in clinical notes
Demographics	Sex	Patient	`gender`	Gender mentioned in history
Clinical	ECOG Score	Observation	LOINC `89247-1`	"ECOG 1" in oncology note
Diagnosis	Lung Cancer Diagnosis	Condition	SNOMED/ICD	"biopsy proven NSCLC"
Imaging	Brain Metastases	DiagnosticReport	MRI Narrative Text	"brain mets present" or "no mets"

Pathology/Biomarker	PD-L1 Expression	Observation	LOINC, % TPS	"PD-L1 80% TPS" in pathology
Pathology/Biomarker	EGFR Mutation	Observation	Genomic variant ID/LOINC	"EGFR mutation detected" in molecular report
Laboratory	Hemoglobin	Observation	LOINC 718-7	Lab report table
Laboratory	Platelet Count	Observation	LOINC 777-3	CBC narrative
Laboratory	Creatinine	Observation	LOINC 2160-0	"Cr 1.0" in lab report
Laboratory	Creatinine Clearance	Observation/Calc	Derived or coded	"CrCl 45 mL/min – ineligible"
Laboratory	Pregnancy Test	Observation	LOINC 2106-3	"Preg test negative on mm/dd/yyyy"
Clinical	Autoimmune Condition	Condition	SNOMED CT	"history of lupus"
Clinical	Prior Malignancy	Condition	SNOMED/ICD	"prior breast cancer" in oncology summary
Imaging	Imaging Lesion Size	Observation	Tumor size (mm or cm)	"2.5 cm lesion on CT" in radiology report

This structured mapping bridges trial eligibility requirements with hospital EMR data to support accurate, efficient screening and rule-based patient matching.